

ADENILDA RIBEIRO DE MOURA

**MODELAGEM DE CRESCIMENTO E PRODUÇÃO EM PLANTIOS DE
EUCALIPTOS NO AMAPÁ: AVALIAÇÃO DE MÉTODOS CLÁSSICOS E
APRENDIZADO DE MÁQUINA**

**RECIFE
PERNAMBUCO – BRASIL
FEVEREIRO -2025**

ADENILDA RIBEIRO DE MOURA

**MODELAGEM DE CRESCIMENTO E PRODUÇÃO EM PLANTIOS DE
EUCALIPTOS NO AMAPÁ: AVALIAÇÃO DE MÉTODOS CLÁSSICOS E
APRENDIZADO DE MÁQUINA**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências Florestais da Universidade Federal Rural de Pernambuco, para obtenção do título de Doutora em Ciências Florestais.

Orientador:

Prof. PhD. José Antônio Aleixo da Silva

Coorientadores

Prof. Dr. Rinaldo Luiz Caraciolo Ferreira

Prof. Dr. Jadson Coelho de Abreu

**RECIFE
PERNAMBUCO – BRASIL
FEVEREIRO -2025**

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Bibliotecário(a): Suely Manzi – CRB-4 809

M929m Moura, Adenilda Ribeiro de.

Modelagem de crescimento e produção em plantios de eucaliptos no Amapá: avaliação de métodos clássicos e aprendizado de máquina: modelagem de crescimento e produção / Adenilda Ribeiro de Moura. – Recife, 2025.

107 f.; il.

Orientador(a): José Antônio Aleixo da Silva. Co-orientador(a): Rinaldo Luiz Caraciolo Ferreira. Co-orientador(a): Jadson Coelho de Abreu.

Tese (Doutorado) – Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Ciências Florestais, Recife, BR-PE, 2025.

Inclui referências.

1. *Eucalyptus grandis* - Crescimento. 2. Levantamentos florestais . 3. Madeiras de eucaliptos – Indústrias I. Silva, José Antônio Aleixo da, orient. II. Ferreira, Rinaldo Luiz Caraciolo, coorient. III. Abreu, Jadson Coelho de, coorient. IV. Título

ADENILDA RIBEIRO DE MOURA

**MODELAGEM DE CRESCIMENTO E PRODUÇÃO EM PLANTIOS DE
EUCALIPTOS NO AMAPÁ: AVALIAÇÃO DE MÉTODOS CLÁSSICOS E
APRENDIZADO DE MÁQUINA**

Data: 07/02/2025

Banca examinadora:

Prof. Dr. Rafael Leite Braz
(Departamento de Ciência Florestal/UFRPE)

Prof. Dr. Perseu Silva Aparício
(Universidade Estadual do Amapá- UEAP)

Prof. Dr. Fernando Henrique de Lima Gadelha
(Curso Técnico em Meio Ambiente - IFPE)

Prof. Dr. José Wesley Lima Silva (CNPq)

Orientador:

Prof. Dr. José Antônio Aleixo da Silva
(Orientador- Departamento de Ciência Florestal/UFRPE)

**RECIFE
PERNAMBUCO – BRASIL
FEVEREIRO -2025**

Ao meu querido pai Manoel Ribeiro (In memoria), cuja presença e sabedoria moldaram quem sou hoje. Sua vida foi uma fonte inesgotável de amor, força e inspiração. Embora sua ausência seja profundamente sentida, seu legado de bondade, integridade e perseverança vive de mim, que tive a sorte de o conhecer. Dedico este trabalho a ele, com eterna gratidão e amor, lembrando sempre de seu exemplo e das lições valiosas que me deixou. Que sua memória seja sempre a minha fonte de espição para prosseguir.

AGRADECIMENTOS

A Deus, pela sabedoria, força e pela trajetória de vida que ele preparou para mim, concedendo-me oportunidades únicas de enfrentar desafios, superar obstáculos e alcançar a vitória.

Ao meu Orientador Professor José Antônio Aleixo da Silva, pela confiança e crença em minha capacidade desde o início da minha jornada no doutorado. Sua disposição em me orientar, mesmo diante dos desafios que o trabalho demandava, foi fundamental para que eu acreditasse em meu potencial e pudesse superar cada obstáculo. Sou imensamente grata pelo apoio, pela paciência e pela inspiração ao longo desse caminho.

Aos meus coorientadores, Rinaldo Luiz Caraciolo Ferreira e Jadson Coelho de Abreu, pela orientação, apoio e pelo compartilhamento de conhecimentos fundamentais ao longo desta jornada. Em especial, ao Professor Jadson Coelho, cuja dedicação foi incansável, sempre disposto a oferecer contribuições valiosas para o desenvolvimento deste trabalho.

À minha família, que não mediu esforços para me acolher e apoiar nesta jornada repleta de desafios e privações. Em especial, às duas grandes guerreiras: minha filha, Dayane Hally de Moura Chaves, e minha filha de coração, Cleide Maria Luiz, cuja presença foi essencial nos momentos mais difíceis desta caminhada. Sou profundamente grata por todo o amor e força que me proporcionaram.

Ao meu filho, Bernardo de Moura Rodrigues, que com apenas seis anos, esteve ao meu lado em todos os momentos, enchendo meus dias de carinho e amor. Mesmo tão pequeno, enfrentou comigo as privações e o peso das minhas obrigações, e, com sua inocência e alegria, me fez acreditar que tudo valeria a pena. Sou imensamente grata pelo seu amor incondicional e pela força que me transmitiu ao longo desta caminhada.

Aos meus amigos, Jeneson Ferreira da Silva e Thalisson Romário Cabral Gomes, pela amizade e pelo apoio constante ao longo desta jornada.

À Amapá Florestal e Celulose (AMCEL) pela oportunidade de realização da pesquisa, com apoio e confiança durante todo o período.

A Antônio Carlos Rosa, pela confiança, apoio dado para a realização desta pesquisa.

À Universidade Federal do Amapá, pela oportunidade e pelo apoio concedido, permitindo-me a liberação para a realização do Doutorado e, assim, viabilizando mais um passo importante na minha trajetória acadêmica e profissional.

Ao Programa de Pós-Graduação em Ciências Florestais, pela oportunidade de cursar o Doutorado, pela formação acadêmica de excelência e pelo ambiente de aprendizado que me permitiu ampliar conhecimentos, enfrentar desafios e construir novas perspectivas profissionais e científicas.

Aos professores, Dr. José Wesley da Silva, Dr. Perseu da Silva Aparício, Dr. Fernando Henrique de Lima Gadelha e Dr. Rafael Leite Braz pelas contribuições e ensinamentos aplicados a este trabalho.

MOURA, ADENILDA RIBEIRO DE. Modelagem de crescimento e produção em plantios de eucaliptos no Amapá-Brasil: avaliação de métodos clássicos e aprendizado de máquina. 2025. Orientador: José Antônio Aleixo da Silva. Coorientadores: Rinaldo Luiz Caraciolo Ferreira, Jadson Coelho de Abreu.

RESUMO

O *Eucalyptus urograndis* é amplamente cultivado no Brasil devido ao seu rápido crescimento e alto potencial para a produção de madeira, na Região Norte, 5% das florestas plantadas do gênero *Eucalyptus*, 1% está no Amapá. Neste contexto, o objetivo da pesquisa foi avaliar a eficácia e a robustez de diferentes modelos probabilísticos e aprendizado de máquina na previsão do crescimento e produção de florestas de rápido crescimento, com ênfase em povoamentos de *Eucalyptus urograndis* no estado do Amapá. Os dados foram obtidos por meio do Inventário Florestal Contínuo, abrangendo 40 parcelas e 733 árvores medidas entre dois e sete anos, com 165 árvores cubadas pelo método de Smalian. Dentre os modelos probabilísticos usados para a classificação de sítios, o modelo de Chapman-Richards apresentou o melhor desempenho, destacando-se no Critério de Informação de Akaike ($AIC = 830,102$), na Raiz do Erro Médio Quadrático ($REMQ = 2,1026$) e no Erro Médio Absoluto ($EMA = 1,640$), diferenciando-se estatisticamente do modelo Weibull a 5% de significância. Na modelagem volumétrica em função da idade, o modelo Clutter-Jones apresentou os menores erros ($REMQ = 35,985$), garantindo estimativas mais precisas. Além disso, a análise do Incremento Médio Anual (IMA) e da Idade Técnica de Corte (ITC) indicou que sítios de maior qualidade ($S=30$) atingem maior produtividade, com produção estimada de $293,459 \text{ m}^3.\text{ha}^{-1}$ aos 6 anos, permitindo ajustes estratégicos no manejo florestal. Na modelagem volumétrica em função do diâmetro e altura, o modelo Schumacher-Hall Não Linear se destacou com o menor $REMQ$ (8,81) e o menor Valor Ponderado do Escore Estatístico ($VPEE = 4$), sendo o mais preciso, seguido pelo Schumacher-Hall Log-linear (8,83). A análise de correlação indicou alta similaridade entre Schumacher-Hall e Spurr, enquanto o modelo Brody Não Linear apresentou menor correlação, sugerindo previsões mais discrepantes. Os modelos de aprendizado de máquina apresentaram melhores ajustes em relação aos probabilísticos tradicionais, com a Rede Neural Artificial (RNA) registrando os menores valores de $REMQ$ (8,180 no treino e 9,116 na validação), seguida pelo Schumacher-Hall, enquanto o Clutter-Jones obteve os maiores erros. O teste de Kruskal-Wallis confirmou diferenças estatisticamente significativas entre os modelos. Além disso, diâmetro (30,6%) e altura total (29,26%) foram as variáveis mais influentes na modelagem, seguidas pela idade (16,12%) e pelos códigos de clones. Os modelos treinados com 100% dos dados, especialmente RNA e Schumacher-Hall, apresentaram maior estabilidade e precisão, enquanto aqueles treinados com 70% mostraram maior variabilidade nos erros. O modelo Schumacher-Hall 70/30 demonstrou maior consistência em relação à RNA 70/30, sugerindo que pode ser mais robusto quando há menos dados disponíveis para treinamento. O teste ANOVA confirmou diferenças estatisticamente significativas entre os modelos, reforçando que o treinamento com 100% dos dados melhora a confiabilidade das previsões. A aplicação do Algoritmo Genético (AG) possibilitou a otimização dos modelos RNA, Máquina de Vetores de Suporte (SVM), Random Forest e Schumacher-Hall, reduzindo a variabilidade dos erros. Embora diferenças numéricas tenham sido observadas entre os modelos, os testes estatísticos indicaram que não houve diferenças significativas ao nível de 5% após a otimização. Assim, tanto os modelos probabilísticos quanto os de aprendizado de máquina são viáveis para estimar o volume comercial de *Eucalyptus urograndis*, cabendo à empresa florestal optar entre simplicidade e confiabilidade (Schumacher-Hall e Clutter-Jones) ou maior flexibilidade e precisão (RNA, SVM e Random Forest).

Palavras-chave: *Eucalyptus grandis* - Crescimento. 2. Levantamentos florestais . 3. Madeiras de eucaliptos – Indústrias .

MOURA, ADENILDA RIBEIRO DE. Growth and Yield Modeling in Eucalyptus Plantations in Amapá, Brazil: Evaluation of Classical Methods and Machine Learning. 2025. Orientador: José Antônio Aleixo da Silva. Co-supervisors: Rinaldo Luiz Caraciolo Ferreira, Jadson Coelho de Abreu.

ABSTRACT

Eucalyptus urograndis is widely cultivated in Brazil due to its rapid growth and high potential for wood production. In the Northern Region, *Eucalyptus* plantations account for 5% of the total planted forests, with only 1% located in Amapá. In this context, the objective of this research was to evaluate the effectiveness and robustness of different probabilistic and machine learning models in predicting the growth and production of fast-growing forests, with an emphasis on *Eucalyptus urograndis* plantations in the state of Amapá. Data were obtained through Continuous Forest Inventory, covering 40 plots and 733 trees measured between two and seven years, with 165 trees scaled using the Smalian method. Among the probabilistic models used for site classification, the Chapman-Richards model showed the best performance, standing out in the Akaike Information Criterion (AIC = 830.102), Root Mean Square Error (RMSE = 2.1026), and Mean Absolute Error (MAE = 1.640), differentiating statistically from the Weibull model at a 5% significance level. In volumetric modeling as a function of age, the Clutter-Jones model had the lowest errors (RMSE = 35.985), ensuring more accurate volume estimates. Additionally, the analysis of Mean Annual Increment (MAI) and Technical Cutting Age (TCA) indicated that higher-quality sites (S=30) achieve greater productivity, with an estimated production of 293.459 m³.ha⁻¹ at six years, allowing for strategic adjustments in forest management. In volumetric modeling based on diameter and height, the Schumacher-Hall Non-Linear model stood out with the lowest RMSE (8.81) and Weighted Statistical Score Value (WSSV = 4), making it the most accurate, followed by the Schumacher-Hall Log-linear model (8.83). Correlation analysis indicated high similarity between Schumacher-Hall and Spurr, while the Brody Non-Linear model showed lower correlation, suggesting more divergent predictions. Machine learning models achieved better adjustments compared to traditional probabilistic models, with the Artificial Neural Network (ANN) registering the lowest RMSE values (8.180 in training and 9.116 in validation), followed by Schumacher-Hall, while Clutter-Jones had the highest errors. The Kruskal-Wallis test confirmed statistically significant differences among models. Additionally, diameter (30.6%) and total height (29.26%) were the most influential variables in modeling, followed by age (16.12%) and clone codes.

Models trained with 100% of the data, especially ANN and Schumacher-Hall, showed greater stability and accuracy, while those trained with 70% exhibited higher error variability. The Schumacher-Hall 70/30 model demonstrated greater consistency compared to ANN 70/30, suggesting that it may be more robust when less training data is available. The ANOVA test confirmed statistically significant differences among the models, reinforcing that training with 100% of the data improves prediction reliability. The application of the Genetic Algorithm (GA) allowed for the optimization of ANN, Support Vector Machine (SVM), Random Forest, and Schumacher-Hall models, reducing error variability. Although numerical differences were observed among the models, statistical tests indicated no significant differences at the 5% level after optimization. Thus, both probabilistic models and machine learning approaches are viable for estimating the commercial volume of *Eucalyptus urograndis*, allowing forest companies to choose between simplicity and reliability (Schumacher-Hall and Clutter-Jones) or greater flexibility and precision (ANN, SVM, and Random Forest).

Keywords: *Eucalyptus grandis* - Growth. 2. Forest surveys. 3. *Eucalyptus* wood – Industries.

LISTA DE FIGURAS

Figura 1- Incremento corrente anual.....	20
Figura 2- Arquitetura de uma rede MLP.....	32
Figura 3- Representação Gráfica do Hiperplano.....	34
Figura 4- Árvore de decisão.....	35
Figura 5- Fluxograma do Algoritmo Genético.....	36
Figura 6- Localização da área de estudo, plantio experimental contínuo de clones Eucaliptos pela Empresa AMCEL S/A- Amapá Florestal e Celulose, localizado município de Porto Grande, AP, Brasil.....	38
Figura 7- Fluxograma aplicado na metodologia da pesquisa.....	40
Figura 8- Etapas metodológicas para a avaliação de diferentes modelos de estimativa volumétrica.....	41
Figura 9 - Arquitetura da rede neural artificial com pesos sinápticos para estimativa do volume <i>Eucalyptus urograndis</i>	48
Figura 10 - Modelos de crescimento ajustados em função da idade de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	57
Figura 11 - Distribuição dos resíduos testados na estimativa das variáveis altura dominante em função das idades de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	58
Figura 12 - Histograma dos resíduos percentuais das variáveis altura dominante em função das idades de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	58
Figura 13- Curvas de crescimento da altura dominante (h_{dom}) em função da idade, ajustadas para diferentes classes de sítio em povoamentos de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	59
Figura 14- Dispersão entre volumes observados e estimados para a prognose do volume (m^3) de <i>Eucalyptus urograndis</i> implantados no município de Porto Grande- AP.....	63
Figura 15 - Dispersão dos resíduos para a prognose do volume (m^3) de <i>Eucalyptus urograndis</i> implantados no município de Porto Grande- AP.....	63
Figura 16 - Distribuição da frequência dos resíduos do volume (m^3) de <i>Eucalyptus urograndis</i> implantados no município de Porto Grande- AP.....	64
Figura 17- Idade técnica de corte (ITC), definida pela interseção das curvas de incremento médio anual (IMA) e incremento corrente anual (ICA) e as classes produtivas de <i>Eucalyptus urograndis</i> implantados no município de Porto Grande- AP....	65
Figura 18 - Ranking final dos modelos com base nos critérios estatísticos (REMQ, AIC IA _c).....	69
Figura 19- Dispersão entre os volumes observados e previstos pelos modelos testados na estimativa do volume comercial de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	70
Figura 20- Distribuição residual entre os volumes observados e previstos pelos modelos probabilísticos testados na estimativa do volume comercial de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	71
Figura 21- Histograma de frequência dos resíduos do volume comercial ($m^3 ha^{-1}$) de <i>Eucalyptus urograndis</i> testados pelos modelos probabilísticos no município de Porto Grande-AP.....	72
Figura 22- Matriz de correlação utilizada na estimativa dos volumes comerciais de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	73
Figura 23- Representação esquemática da arquitetura da RNA.....	75
Figura 24- Distribuição do REMQ na modelagem volumétrica de <i>Eucalyptus urograndis</i>	76

Figura 25 - Gráficos de predições, resíduos e histogramas para os modelos Clutter e Schumacher-Hall nos conjuntos de treino e teste, utilizados na estimativa do volume comercial de <i>Eucalyptus urograndis</i> em Porto Grande, AP.....	78
Figura 26 - Importância das variáveis na estimativa do volume comercial de <i>Eucalyptus urograndis</i> com base no aumento percentual do erro MSE.....	79
Figura 27 -Comparação dos modelos RNA e Schumacher-Hall para estimativa volumétrica com base em REMQ, MAE e IAaj.....	88
Figura 28- Comparação entre volumes observados e previstos pelos modelos RNA, SVM, Random Forest e Schumacher-Hall para estimativa de volume comercial de <i>Eucalyptus urograndis</i>	96
Figura 29 - Distribuição gráfica dos resíduos e histogramas dos modelos RNA, SVM, Random Forest e Schumacher-Hall na validação para estimativa do volume comercial de <i>Eucalyptus urograndis</i>	97

LISTA DE TABELAS

Tabela 1- Modelos matemáticos para construção das curvas de Índice de Sítio.....	41
Tabela 2- Modelos não lineares ajustados para estimativa do volume de <i>Eucalyptus urograndis</i> no município do Amapá, Brasil.....	43
Tabela 3- Modelos não lineares para a estimativa do volume de madeira ($m^3 \text{ arvore}^{-1}$) em povoamento comercial de clones de <i>Eucalyptus urograndis</i>	45
Tabela 4- Arquitetura das redes neurais artificiais selecionadas na fase de treinamento para estimativa do volume de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	46
Tabela 5- Modelos probabilísticos utilizados para estimativa do volume de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	46
Tabela 6- Estatística descritiva dos valores médios das variáveis biométricas D, H e V do povoamento de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	55
Tabela 7- Coeficientes e estatísticas dos modelos testados para construção das curvas de sítio.....	56
Tabela 8- Sumário do teste de Scott-Knott de comparação de média dos modelos testados para construção das curvas de sítio.....	57
Tabela 9- Limites das classes de sítio da relação entre a idade e a altura dominante (h_{dom}) em diferentes classes de sítio, refletindo a produtividade dos povoamentos de <i>Eucalyptus urograndis</i>	60
Tabela 10- Equações obtidas para os seis modelos avaliados para a prognose do volume (m^3) de <i>Eucalyptus urograndis</i> implantados no município de Porto Grande- AP	61
Tabela 11- Estatísticas e coeficientes das equações para a prognose do volume (m^3) de <i>Eucalyptus urograndis</i> implantados no município de Porto Grande- AP.....	62
Tabela 12- Prognose do volume ($m^3 \text{ ha}^{-1}$) para a produção estimada e Incremento Médio Anual (IMA) para diferentes classes de sítio no modelo de Clutter e Jones.....	66
Tabela 13- Sumário do teste de Scott-Knott de comparação de média do volume <i>Eucalyptus urograndis</i> em função do diâmetro e altura.....	67
Tabela 14- Valores dos parâmetros estimados, os índices de ajustes e os erros padrões das estimativas obtidos a partir da modelagem do crescimento em volume de <i>Eucalyptus urograndis</i>	68
Tabela 15- Especificações dos modelos utilizados para estimativa do volume comercial de <i>Eucalyptus urograndis</i>	74
Tabela 16- Análise das configurações e estatísticas de qualidade do Perceptron Multicamadas (MLP) usando a das funções de ativação tangente hiperbólica aplicado a determinação do volume comercial de <i>Eucalyptus urograndis</i> no município de Porto Grande-AP.....	74
Tabela 17- Parâmetros ajustados e métricas de desempenho dos modelos testados para a estimativa do volume comercial de <i>Eucalyptus urograndis</i> em Porto Grande, AP.....	76
Tabela 18- Importância e sensibilidade das variáveis no modelo Random Forest com base no aumento percentual do erro MSE e pureza dos nós.....	80
Tabela 19- Estatísticas descritivas das métricas de avaliação dos modelos testados de aprendizado de máquina e probabilísticos com amostras aleatórias na estimativa do volume de comercial de <i>Eucalyptus urograndis</i>	81
Tabela 20- Parâmetros estimados do modelo Schumacher-Hall log-linear em 20 Simulações para a estimativa do volume de <i>Eucalyptus urograndis</i>	82

Tabela 21- Estatísticas das 20 simulações independentes para os modelos RNA e Schumacher-Hall nas Configurações de Treino (70% e 100%) e Validação (30%) para a estimativa do volume de <i>Eucalyptus urograndis</i>	84
Tabela 22- Estatísticas da validação cruzada: RMSE, MAE e IAaj para os modelos testados de aprendizado de máquina e probabilísticos com amostras aleatórias na estimativa do volume de comercial de <i>Eucalyptus urograndis</i>	86
Tabela 23- Análise de variância (ANOVA) para comparação de modelos ajustados de RNA, e Schumacher-Hall na estimativa do volume comercial de <i>Eucalyptus urograndis</i>	86
Tabela 24- Resultados do Teste de Scott-Knott para comparação do desempenho dos modelos testados de aprendizado de máquina e probabilísticos com amostras aleatórias na estimativa do volume de comercial de <i>Eucalyptus urograndis</i>	87
Tabela 25- Desempenho dos modelos preditivos (RNA, MVS, Random Forest e Schumacher-Hall) em treino e validação, com parâmetros ajustados pelo Algoritmo Genético (GA).....	89
Tabela 26- Comparação dos modelos RNA, SVM, Random Forest e Schumacher-Hall na estimativa do volume comercial de <i>Eucalyptus urograndis</i> em Porto Grande-AP, com parâmetros ajustados e métricas de desempenho.....	91
Tabela 27- Análise de variância (ANOVA) para comparação de modelos ajustados de .RNA, SVM, Random Forest e Schumacher-Hall na estimativa do volume comercial de <i>Eucalyptus urograndis</i>	94
Tabela 28- Sumário do teste de Scott-Knott de comparação de média dos modelos RNA, SVM, Random Forest e Schumacher-Hall na estimativa do volume comercial de <i>Eucalyptus urograndis</i>	95

SUMÁRIO

RESUMO	
ABSTRACT	
LISTA DE FIGURAS	
LISTA DE TABELAS	
1 INTRODUÇÃO	15
2 OBJETIVOS	16
2.1 Objetivo geral	16
2.2 Objetivo específicos	16
3 HIPÓTESE	17
4 REVISÃO DE LITERATURA	17
4.1 FLORESTAS DE RÁPIDO CRESCIMENTO	17
4.2 CAPACIDADE PRODUTIVA	18
4.3. CRESCIMENTO E PRODUÇÃO FLORESTAL	19
4.4 MODELOS DE CRESCIMENTO E PRODUÇÃO	21
4.4.1 Modelo de Bertalanffy	22
4.4.2 Modelo de Chapman-Richards	24
4.4.3 Modelo de Brody	25
4.4.4 Modelo de Weibull	26
4.4.5 Modelo de Spurr	27
4.4.6 Modelo de Silva- Bailey	28
4.4.7 Modelo de Clutter E Jones	29
4.4.8 Modelo de Schumacher e Hall	30
4.5 POTENCIALIDADES E DESAFIOS NO USO DA INTELIGÊNCIA ARTIFICIAL NA MODELAGEM FLORESTAL	30
4.5.1 Redes Neurais Artificiais (RNA)	31
4.5.2 Máquina de Vetor de Suporte (MVS)	33
4.5.3 Randon Forest	35
4.5.4 Algoritmo Genético	36
5 MATERIAL E MÉTODOS	38
5.1 ÁREA DE ESTUDO	38
5.2 DESCRIÇÃO DOS DADOS	38
5.3 METODOLOGIA UTILIZADA NA MODELAGEM FLORESTAL	41
5.3.1 Modelagem para classificação de sítios florestais	41
5.3.2 Modelagem volumétrica em função da idade	42
5.3.3 Modelagem volumétrica em função do diâmetro (D) e altura (H)	44
5.3.4 Algoritmos de aprendizagem de máquina	45
5.3.4.1 Redes Neurais Artificiais	45
5.3.4.2 Simulações com amostras aleatórias na estimativa do volume de <i>Eucalyptus urograndis</i> comparando modelos computacionais e probabilísticos	47
5.3.4.3 Algoritmo Genético	49
6 MÉTRICAS PARA AVALIAÇÃO DOS MODELOS	50
7 RESULTADOS E DISCUSSÃO	55
7.1 Classificação de sítios de povoamentos de <i>Eucalypto urograndis</i>	55
7.2 Modelagem volumétrica de <i>Eucalyptus urograndis</i> em função da idade	61
7.3 Modelagem volumétrica de <i>Eucalyptus urograndis</i> em função do diâmetro e altura	65

7.4	Potencial de algoritmos de aprendizado de máquina na modelagem volumétrica de <i>Eucalyptus urograndis</i>	74
7.5	Impacto de simulações com amostras aleatórias na estimativa do volume de <i>Eucalyptus urograndis</i> comparando modelos computacionais e tradicionais.....	81
7.6	Otimização de modelos preditivos com algoritmo genético para estimativa do volume de <i>Eucalyptus urograndis</i>	89
8	CONCLUSÃO	98
9	REFERÊNCIAS BIBLIOGRÁFICAS	99

1. INTRODUÇÃO

O *Eucalyptus urograndis*, híbrido resultante do cruzamento entre *Eucalyptus grandis* e *Eucalyptus urophylla*, destaca-se no setor florestal brasileiro devido ao seu rápido crescimento e alto potencial para a produção de madeira, sendo amplamente empregado na indústria de celulose (IBÁ, 2023). No entanto, na Região Norte do Brasil, as florestas plantadas com o gênero *Eucalyptus* representam apenas 5% do total, indicando uma menor participação em comparação com outras regiões do país. Dentro desse percentual, apenas 1% da área destinada a florestas plantadas encontra-se no estado do Amapá.

Apesar desta participação reduzida, o cultivo de *Eucalyptus urograndis* tem se consolidado economicamente na região, impulsionando a produção de madeira de rápido crescimento. Empresas como a Amapá Florestal e Celulose S.A. (AMCEL) têm investido na formação de povoamentos clonais dessa espécie, garantindo sua adaptação às condições edafoclimáticas locais e contribuindo significativamente como matéria-prima para as indústrias de celulose e papel, tanto no Brasil quanto no exterior (Pereira *et al.*, 2016).

Desta forma a modelagem do crescimento e da produção florestal é essencial para o manejo sustentável e a tomada de decisões estratégicas no setor (Campos e Leite, 2017). Modelos matemáticos permitem a estimativa do volume comercial de madeira com base em variáveis dendrométricas, como diâmetro e altura das árvores (Leite *et al.*, 2011). Estas estimativas possibilitam o planejamento eficiente das colheitas, a otimização do uso dos recursos florestais e o aumento da produtividade.

Com os avanços tecnológicos, novas abordagens computacionais foram desenvolvidas para aprimorar a precisão das estimativas volumétricas. Entre essas técnicas, destaca-se o aprendizado de máquina, que se diferencia pela capacidade de capturar padrões complexos nos dados e reduzir o impacto de outliers (Cunha Neto *et al.*, 2021). O uso dessas ferramentas tem se mostrado promissor na predição do volume florestal (Bonete *et al.*, 2019; Abreu *et al.*, 2020).

Estudos demonstram que modelos baseados em aprendizado de máquina podem ser usados com alternativas na modelagem florestal, por apresentaram bons ajustes e precisão das estimativas (Binoti *et al.*, 2014; Martins *et al.*, 2015; Gao *et al.*, 2018; Pandit *et al.*, 2018; Cunha Neto *et al.*, 2021). Pereira *et al.* (2016) utilizaram Redes Neurais Artificiais (RNA) para ajustar modelos volumétricos de simples e dupla entrada, obtiveram os melhores resultados na predição do volume de *Eucalyptus urograndis* no município de Porto Grande, Amapá. Da mesma forma, Garcia Neto *et al.* (2012) verificaram que técnicas como Máquinas de Vetores de Suporte (SVM) e Redes Perceptron Multicamadas (MLP) apresentaram maior precisão na estimativa do volume de casca de *Eucalyptus globulus* quando comparadas aos modelos tradicionais.

Estudos recentes também destacam o uso das Redes Neurais Artificiais na modelagem volumétrica. Cordeiro *et al.* (2022) identificaram a RNA como a técnica mais eficiente para estimar o volume individual com casca em *Eucalyptus urophylla*. Em contrapartida, Santana *et al.* (2023) apontaram o modelo de Floresta Aleatória (Random Forest - RF) como o mais preciso para prever o volume de madeira de eucalipto, embora a inclusão de variáveis categóricas, como espécie, não tenha melhorado significativamente a precisão das estimativas.

Além das RNAs, os Algoritmos Genéticos (AGs) têm sido explorados na ciência florestal como alternativa para otimizar modelos de crescimento e produção. Esta técnica é baseada em princípios de evolução biológica, apresentam vantagens na solução de problemas complexos e na busca por múltiplas soluções simultâneas (Kivinen, 2004; Souza *et al.*, 2018).

Diante do exposto, este estudo tem como objetivo avaliar a eficácia e a robustez de diferentes modelos matemáticos na previsão do crescimento e produção de florestas de rápido crescimento, com ênfase em povoamentos de *Eucalyptus urograndis* no estado do Amapá. A pesquisa busca contribuir para a otimização da modelagem volumétrica, permitindo um melhor planejamento e manejo sustentável das plantações florestais.

2. OBJETIVOS

2.1 GERAL

Avaliar a eficácia e a robustez de diferentes modelos probabilísticos e de aprendizado de máquina na previsão do crescimento e produção de florestas de rápido crescimento, com ênfase em povoamentos de *Eucalyptus urograndis* no estado do Amapá.

2.2 ESPECÍFICOS

- Avaliar o desempenho de modelos não lineares para classificação de sítios em povoamentos de *Eucalyptus urograndis*, com base em parâmetros estimados, estatísticas de precisão e curvas de crescimento;
- Avaliar o desempenho de modelos probabilísticos na estimativa do volume de *Eucalyptus urograndis* em função da idade, visando identificar a equação mais precisa e consistente para a prognose volumétrica e a definição da idade técnica de corte (ITC).
- Avaliar e comparar o desempenho de diferentes modelos probabilísticos (lineares, não lineares e exponenciais) na estimativa do volume comercial de *Eucalyptus urograndis* em função do diâmetro e da altura.

- Comparar o desempenho de Redes Neurais Artificiais (MLP) e modelos probabilísticos clássicos, incluindo Schumacher-Hall, na estimativa do volume comercial de *Eucalyptus urograndis*, utilizando simulações com amostras aleatórias.
- Utilizar o Algoritmo Genético (AG) para otimizar o desempenho de modelos preditivos (RNA, SVM, RF e Schumacher-Hall) nas estimativas dos volumes comerciais de *Eucalyptus urograndis*.

3. HIPÓTESE

Os modelos de aprendizado de máquina (RNA, SVM e RF) e os modelos probabilísticos apresentam desempenhos semelhantes na previsão de crescimento e produção de *Eucalyptus urograndis*, no estado do Amapá.

4. REVISÃO DE LITERATURA

4.1 FLORESTAS DE RÁPIDO CRESCIMENTO

As florestas plantadas desempenham um importantíssimo papel na cadeia produtiva e os benefícios extraídos por elas, geram renda para população, alavancando a economia do País a partir dos produtos, subprodutos, derivados, serviços e insumos relativos com grande potencial de expansão no mercado (Brasil, 2019).

As florestas de rápido crescimento são encontradas em todo o mundo, mas são mais comuns em países com climas tropicais e subtropicais. Dentre as principais espécies de árvores utilizadas nessas florestas incluem as do gênero *Eucalyptus* com uma série de vantagens sobre as florestas naturais, desde que sejam bem planejadas antes da sua implantação (Moreira *et al.*, 2017, Fritzsos; Parron, 2017). Podendo ser plantadas e colhidas em um período relativamente curto, principalmente devido à alta adaptabilidade das espécies a diferentes condições de solo e clima e elevada produtividade em biomassa o que as torna uma fonte de madeira mais sustentável (Ribeiro *et al.*, 2017).

No Brasil até o princípio do século XX, os eucaliptos foram plantados como árvores decorativas. O gênero *Eucalyptus* não era visto como potencial econômico e ambiental, apenas era enfatizado as suas propriedades sanitárias. Pouquíssimas foram às plantações com fins industriais e caráter florestal. O desfecho maior para a introdução dos eucaliptos no Brasil se deve à Companhia Paulista de Estradas de Ferro e ao Dr. Edmundo Navarro de Andrade com a expansão da eucaliptocultura, por meio da sistematização da cultura e com o incentivo de

plantios em áreas desmatadas, por ser um gênero de crescimento rápido, que poderia suprir a demanda da madeira em um intervalo curto de tempo em relação as espécies nativas (Moreira *et al.*, 2017).

Desta forma, a Companhia Paulista de Estradas de Ferro e Fluvial, começou a adquirir terras ao longo das linhas férreas, com o propósito de estimular a cultura florestal no estado de São Paulo, a partir de 1903. Em 1904, Edmundo Navarro de Andrade instalou os primeiros experimentos de natureza silvicultural para determinar uma essência capaz de fornecer madeira e lenha combustível para suprir as necessidades da ferrovia. Essa expansão estava relacionada com a pressão sobre as florestas naturais que aos poucos foi mudando para as florestais plantadas, com propósito comercial em reduzir a busca por produtos florestais naturais, em contrapartida suprir as necessidades madeireiras (Moreira *et al.*, 2017).

Esse impulso inicial para a formação de florestas plantadas no Brasil, motivado pelas necessidades da ferrovia e pela preservação das florestas nativas, abriu caminho para o avanço da silvicultura no país. Com o tempo, a adoção de práticas silviculturais aprimoradas e o desenvolvimento genético das espécies cultivadas resultaram em ganhos expressivos na produtividade das plantações florestais, como evidenciado pelo incremento médio anual (IMA) do eucalipto, que tem se destacado no cenário global (IBÁ, 2023).

O incremento médio anual (IMA) dos plantios de eucalipto no Brasil é superior ao de outras regiões do mundo. Em 1970, o IMA era de $10\text{m}^3.\text{ha}^{-1}$. Em 2021 chegou a $38,9\text{m}^3.\text{ha}^{-1}$. Já o IMA do pinus, outra espécie muito importante no plantio florestal, foi de $29,7\text{m}^3.\text{ha}^{-1}$ em 2021 (IBÁ, 2022). Estes altos IMAs são resultado de diversos fatores, como a adoção de técnicas modernas de plantio e manejo, o desenvolvimento de novas variedades de árvores e o clima favorável ao crescimento das florestas plantadas no Brasil. A alta produtividade do eucalipto brasileiro é um importante diferencial competitivo do País no mercado internacional de madeira. Além disto, as florestas plantadas contribuem para a geração de empregos, renda e impostos, e ajudam a preservar o meio ambiente.

4.2 CAPACIDADE PRODUTIVA

A capacidade produtiva de um local, conforme definido por Campos e Leite (2017), refere-se à quantidade máxima de madeira ou outros produtos florestais que uma determinada área pode produzir para uma espécie ou clone específico. Para que esta definição seja aplicável, o espaço geográfico deve ser homogêneo em termos de características físicas e biológicas, sendo assim classificado como um "sítio" (Skovsgaard; Vanclay, 2008).

No contexto do manejo florestal, a qualidade de sítio se refere ao potencial produtivo de madeira de uma determinada área, levando em consideração uma espécie ou o tipo de povoamento que será estabelecido em um determinado local (Clutter *et al.*, 1983). Assim, é possível classificar ou mensurar o sítio florestal com base em sua capacidade produtiva. De acordo com Husch *et al.* (1982), a qualidade de sítio está diretamente relacionada à produtividade do povoamento florestal, sendo avaliada por meio do crescimento de variáveis intrínsecas à árvore ou ao povoamento ao longo de um período.

Este conceito é fundamental no manejo florestal, pois oferece informações críticas para a tomada de decisões estratégicas e a otimização da produção (Husch *et al.*, 1982). A análise da capacidade produtiva do sítio permite um entendimento mais profundo dos fatores que influenciam a produtividade, como o clima, o solo e as características genéticas das árvores, o que é essencial para o planejamento de práticas de manejo florestal sustentável. Para quantificar a capacidade produtiva de povoamentos equiâneos, o método mais usual em manejo florestal é a estimativa de índices de local, calculados em função de dados médios de alturas dominantes e idade das árvores (Campos; Leite, 2017).

A altura das árvores dominantes é considerada a variável mais eficiente para avaliar a qualidade do sítio (Santos *et al.* 2018). Esta métrica é determinada pela altura média das árvores dominantes em diferentes idades do povoamento. Utilizar a altura dominante como indicador do índice de sítio apresenta vantagens práticas, como sua ampla aplicabilidade e a possibilidade de comparar taxas de crescimento entre plantações da mesma idade, mesmo que estejam submetidas a diferentes densidades e práticas de manejo (Prodan *et al.*, 1997; Campos; Leite, 2017).

4.3. CRESCIMENTO E PRODUÇÃO FLORESTAL

Para planejar o manejo florestal, é importante saber como as árvores e florestas crescem. Isto significa estimar a quantidade de madeira que elas produzirão agora e no futuro (Prodan *et al.*, 1968). Os estudos de crescimento e produção florestal utilizam dados de pontos como base para a modelagem (Avery; Burkhart, 1983). Estes dados podem ser coletados por meio de parcelas permanentes.

É importante utilizar um conjunto de dados consistente para garantir a precisão das estimativas. As variáveis utilizadas na modelagem devem ser relevantes para o objetivo do estudo e fáceis de serem medidas (Abreu, 2000). No contexto atual de mudanças climáticas e crescente demanda por recursos florestais, a modelagem do crescimento e da produção florestal é uma ferramenta essencial para auxiliar os gestores florestais na tomada de decisões mais

precisas e eficientes, garantindo a produção de madeira de forma sustentável e minimizando os impactos ambientais.

A produção de um povoamento florestal pode ser expressa matematicamente por uma equação de produção. A derivada da equação define o incremento corrente anual (ICA), que é o aumento da produção em um determinado período, geralmente um ano (Scolforo, 2006, Campos; Leite, 2017).

O ICA é um indicador importante da produtividade de uma floresta e ajuda os gestores a tomarem decisões sobre o manejo florestal. No contexto atual de mudanças climáticas e crescente demanda por recursos florestais, a medição da produção florestal e a utilização de indicadores como o ICA e a taxa média de crescimento representada pelo incremento médio anual (IMA) se constituem ferramentas importantes e reveladoras para a caracterização dos povoamentos florestais equiâneos e é, normalmente, usada nos estudos de dinâmica de povoamentos florestais (Campos; Leite, 2017). Contudo, a medição da produção florestal e a utilização de indicadores como o ICA e IMA são ferramentas essenciais para o manejo florestal sustentável no contexto atual.

A (Figura 1) ilustra a relação entre o Incremento Corrente Anual (ICA) e o Incremento Médio Anual (IMA) ao longo do tempo, indicando o ponto em que ambos se igualam, conhecido como Idade Técnica de Corte (ITC).

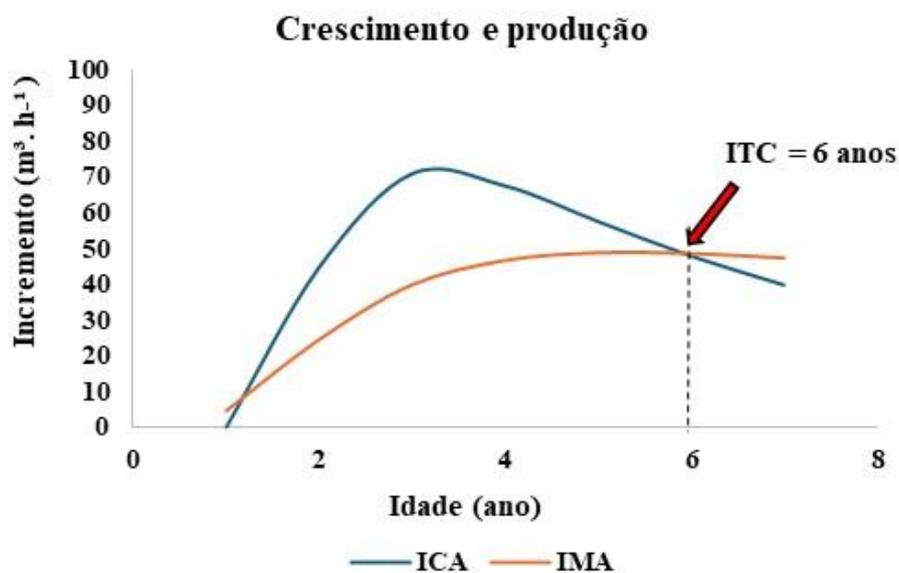


Figura 1- Incremento corrente anual. Fonte: elaborada pela autora (2024).

O ITC é um marco fundamental para a decisão de manejo, pois representa a idade ótima para a colheita, maximizando a produtividade florestal sem comprometer o crescimento. A construção das curvas de incremento médio anual (IMA) e incremento corrente anual (ICA), são representadas de acordo com as equações 1 e 2:

$$IMA = V_t / t \quad (1)$$

$$ICA = V_t - V_{t-1} \quad (2)$$

Em que:

V_t = é o volume do povoamento ($m^3 \cdot ha^{-1}$) na idade em meses;

t = tempo.

No manejo florestal, o ICA reflete a capacidade de crescimento anual em termos volumétricos, enquanto o IMA é a média do crescimento acumulado ao longo dos anos. A interpretação conjunta destes indicadores permite identificar a eficiência do povoamento em converter recursos em biomassa ao longo do tempo. O equilíbrio entre estes dois incrementos, conforme mostrado no gráfico, é determinante para planejar colheitas que otimizem a produção de madeira, considerando aspectos econômicos e ecológicos.

4.4 MODELOS DE CRESCIMENTO E PRODUÇÃO

Para gerenciar as florestas plantadas é essencial compreender os processos e fatores que influenciam o crescimento e a produção de povoamentos. Isto ocorre porque as decisões de manejo são baseadas em projeções de produção futuras, que são feitas a partir de informações correntes (Campos; Leite, 2017; Cosenza *et al.*, 2017). Desta forma, os modelos de crescimento se tornam ferramentas importantíssimas para representar a dinâmica natural de um povoamento florestal (Maestri *et al.*, 2013).

Assim, a empregabilidade do melhor modelo não é uma tarefa fácil, a consulta à literatura é fundamental. Diversos trabalhos abordam o uso de modelos probabilísticos e computacionais com vasta aplicação na mensuração florestal.

Binoti *et al.*, (2014), utilizaram modelos de distribuição diamétrica e Redes Neurais Artificiais em povoamentos equiâneos. Castro *et al.* (2013) usaram o modelo de Clutter (1963) para avaliar crescimento e produção em plantios comerciais de eucaliptos. Bhering *et al.* (2015) aplicaram o modelo de RNA para previsão de volume em eucalipto. Melo e Miguel (2016) fizeram uso da aplicação das (RNA) para predição da qualidade do aglomerado.

Pelo fato do crescimento das árvores geralmente não ser linear, os modelos não lineares passaram a ser mais estudados na mensuração florestal bem como em diversas áreas. Com

relação aos modelos não lineares, baseados no crescimento de povoamentos florestais em função do tempo

Vários modelos não lineares são descritos na literatura para descrever a curva de crescimento tais como: Chapman-Richards, Bertalanffy, Brody, Silva-Bailey, Mitscherlich, Weibull, Clutter e Jones, entre outros, apresentam ajustes eficientes (Brito *et al.*, 2007; Schneider; Schneider; Souza, 2009; Barros, 2010; Lima Filho *et al.*, 2012; Sales *et al.*, 2015; Gouveia *et al.*, 2015; Silva *et al.*, 2017; Frühauf *et al.*, 2022).

Desta forma, estes modelos não lineares tem sido aplicada na modelagem de crescimento, sendo considerados como soluções de sistemas de equações diferenciais que permitem obter interpretações práticas ou biológicas sobre tal fenômeno. Estas interpretações são obtidas a partir das relações entre os parâmetros, que podem ser interpretados como variáveis que influenciam o fenômeno.

4.4.1 Modelo de Bertalanffy

O modelo de crescimento proposto por Ludwig Von Bertalanffy (Bertalanffy, 1938) consiste em uma formulação matemática que descreve o desenvolvimento de organismos vivos. Fundamentado na premissa de que o crescimento é um processo contínuo e orientado para um ponto de maturidade, o modelo tem sido amplamente aplicado no campo da ecologia, especialmente na modelagem do crescimento de peixes e crustáceos ao longo do tempo. A evolução do crescimento animal, conforme o modelo de Bertalanffy, pode ser descrita por uma equação diferencial que captura as mudanças ao longo do tempo. A hipótese de crescimento assume que a dinâmica do crescimento animal pode ser expressa por meio de uma equação específica que captura essa evolução temporal, sendo expressa por:

$$\frac{dP}{dt} = \alpha P^m - \beta P^n$$

Em que P representa o peso do organismo em função do tempo t, α e β são parâmetros associados aos processos de anabolismo e catabolismo, respectivamente, e os expoentes m e n refletem as proporções entre o peso e a potência correspondente.

Com base em fundamentos fisiológicos, Bertalanffy, concluiu que, para uma vasta classe de organismos, a relação alométrica entre a área superficial (seja a área total ou a área absorvente) assume que o expoente n é igual a 1, enquanto m tem o valor de 2/3, derivado da

correlação entre o peso do animal e sua área corporal. Assim, o modelo revisado por Bertalanffy (1957) pode ser formalmente expresso da seguinte maneira:

$$\frac{dP}{dt} = \alpha P^{\frac{2}{3}} - \beta P$$

Em que:

$P = P(t)$ é a variável resposta;

t é o tempo;

α , e β = parâmetros do modelo.

Conforme, proposto por Bertalanffy o tamanho do organismo cresce exponencialmente com o tempo e, conseqüentemente, a taxa de crescimento diminui e o tamanho final do organismo é determinado pelo valor assintótico (P_{∞}). Desta forma, a equação diferencial proposta pelo autor é que a taxa de anabolismo é proporcional à área de superfície corporal, enquanto o catabolismo é proporcional ao peso corporal. Neste modelo o ponto de inflexão é fixo em aproximadamente 0,30 do valor de α sendo representado pelas expressões 3 e 4 :

Que pode ser expresso como:

$$W(t) = \alpha [1 - \beta e^{-kt}]^3 + \varepsilon_i \quad (3)$$

$$W_i = W_{\infty} [1 - \exp(-\kappa.t_i)]^3 + \varepsilon_i \quad (4)$$

Em que:

W_i = tamanho inicial do organismo, podendo ser expresso em várias unidades, como peso, altura, volume.

W_{∞} = tamanho máximo que o organismo pode atingir (valor assintótico);

κ = parâmetro do modelo;

ε_i = erro aleatório.

Clutter (1963) desenvolveu uma metodologia na qual o valor do W_{∞} pode ser eliminado do modelo, pelo uso do método das diferenças de equações.

Considerando que no tempo i o modelo possui as seguintes expressões (5 e 6):

$$W_i = W_{\infty} [1 - \exp(-\kappa.t_i)]^3 \quad (5)$$

e no tempo $i+1$:

$$W_{i+1} = W_{\infty} [1 - \exp(-\kappa \cdot t_{i+1})]^3 \quad (6)$$

Encontrando o valor de W_{∞} em (5) e substituir em (6), a estrutura do modelo muda

para:

$$W_{i+1} = W_i \left[\frac{1 - \exp(-k \cdot t_{i+1})}{1 - \exp(-k \cdot t_i)} \right]^3 + \varepsilon_i$$

Esta forma permite que se estime um valor futuro no tempo $i+1$ pelo conhecimento do valor em um tempo i . Este modelo possui as propriedades de compatibilidade e invariância (Clutter, 1963). Portanto, o modelo de Bertalanffy é uma ferramenta valiosa para os biólogos que estudam o crescimento de organismos vivos.

4.4.2 Modelo de Chapman-Richards

Richards (1959) e Chapman (1961), ao investigarem o crescimento de plantas e peixes, respectivamente, identificaram que o expoente alométrico de $2/3$ proposto por Bertalanffy era limitado para descrever o crescimento de organismos de diferentes espécies e tamanhos.

Este expoente, conforme Silva (1986), pode variar de acordo com a natureza da população em estudo. A generalização de Chapman-Richards ao modelo de crescimento de Bertalanffy configura uma formulação matemática aprimorada, concebida para fornecer uma interpretação mais ampla e holística dos fatores intrínsecos que modulam o crescimento de organismos vivos.

Esta abordagem inovadora confere aos parâmetros do modelo um significado biológico ou fisiológico, o que amplia substancialmente sua aplicabilidade a uma vasta gama de espécies e condições ecológicas distintas (Pienaar; Turnbull, 1981). Desta forma, ao incorporar tal modificação, obtém-se uma ferramenta de análise mais robusta e adaptável.

Assim sendo, fazendo esta modificação, e por exemplo, mensurando o crescimento em altura (H), obtém-se:

$$\frac{dH}{dt} = \alpha H^m - \gamma H$$

Após resolução desta equação diferencial se encontra a forma final do modelo expresso por (7 e 8):

$$H = U \{1 - \exp[-k(t - t_0)]\}^{\frac{1}{1-\mu}} + \varepsilon_i \quad (7)$$

Quando $t_0 = 0$

$$H = U \{1 - \exp[-kt]\}^{\frac{1}{1-\mu}} + \varepsilon_i \quad (8)$$

ou:

$$W_{i+1} = W_i \left[\frac{1 - \exp(-k \cdot t_{i+1})}{1 - \exp(-k \cdot t_i)} \right]^{\frac{1}{1-\mu}} + \varepsilon_i$$

Este modelo, nomeado como modelo de crescimento de Chapman-Richards por Turnbull (1963), tem sido usado com frequências em pesquisas florestais para estimar o crescimento e a produção de árvores.

Desta forma, o modelo de Chapman-Richards pode ser usado para estimar o volume de madeira produzido por uma árvore em uma determinada idade. Para isto, é necessário medir o diâmetro e/ou a altura da árvore na idade inicial e na idade final. Os valores medidos são então usados para estimar os parâmetros do modelo. Uma vez que os parâmetros são estimados, é possível usar a equação resultante para estimar o volume de madeira produzido pela árvore em qualquer idade.

4.4.3 Modelo de Brody

Modelo proposto por Brody (1945) para descrever o crescimento da população humana. Ele prevê que a população irá crescer exponencialmente até atingir a capacidade de suporte do ambiente. Após atingir a capacidade de suporte, a população irá crescer mais lentamente, até atingir um equilíbrio.

Desta forma, o modelo parte de um modo geral de crescimento que resulta da conjunção de dois modelos, a monomolecular (exponencial côncavo) e o exponencial (convexo), para obter o padrão de crescimento ao longo do tempo. A primeira curva sendo uma função exponencial crescente, com sua assíntota infinita, porém esta curva só pode ser aplicada em intervalos limitados e não pode ser extrapolada. A segunda é uma função exponencialmente decrescente (Brody 1945).

O modelo é baseado pela seguinte equação diferencial:

$$\frac{W}{t} = \alpha \cdot [1 - \beta e^{-kt}]$$

Após a solução da equação diferencial e aplicação do método das diferenças de equações resultou em:

$$W_f = W_i - \beta_0 [e^{\beta_1 T_f} - e^{\beta_1 T_i}] + \varepsilon_i$$

Em que:

W_i = variável resposta;

$W_\infty = \alpha$ = valor assintótico (tamanho máximo atingido);

β_0, β_1 = parâmetros do modelo;

t_i = tempo;

ε_i = erro aleatório.

4.4.4 Modelo de Weibull

A função Weibull é uma distribuição de probabilidade contínua amplamente utilizada para modelar diversos fenômenos, especialmente aqueles relacionados a eventos de crescimento e falha. No contexto florestal, é frequentemente aplicada para modelar o crescimento de árvores e a distribuição de diâmetro em uma floresta (Bailey; Dell, 1973). Sua flexibilidade permite que seja ajustada para diferentes padrões de crescimento ao longo do tempo. Em estatística, a função Weibull é uma das mais usadas, ficando atrás apenas das distribuições Normal e Gama em termos de popularidade devido à sua capacidade de se adaptar a diferentes formas de dados (Percontini *et al.*, 2014).

Esta função tendo sido aplicada principalmente em trabalhos que visam construir modelos de crescimento e produção de madeira. Segundo Binoti *et al.* (2010), a função Weibull é a mais utilizada no setor florestal, devido ao fato de os plantios equiâneos apresentarem distribuições diamétricas assimétricas à direita (Campos; Leite, 2017). Isto ocorre porque as árvores maiores apresentam maiores taxas de crescimento que as menores (suprimidas). Isto gera uma configuração que coincide com as condições ideais para o ajuste dessa função.

A função proposta por Weibull é dada por:

$$W_f = W_i - \beta \left(e^{-\kappa \cdot T_f^\phi} - e^{-\kappa \cdot T_i^\phi} \right) + \varepsilon_i$$

Em que:

W_f = variável resposta no tempo final;

W_i = variável resposta no tempo inicial;

α , κ e ϕ = parâmetros do modelo;

t = idade em anos;

ε_i = erro aleatório.

A escolha do método mais apropriado para ajustar a função Weibull depende das especificidades e objetivos de cada estudo. De forma geral, o ajuste de modelos estatísticos é frequentemente realizado utilizando o Método dos Mínimos Quadrados Não-Lineares, especialmente para modelos não-lineares. Este método estabelece uma relação entre a distribuição da variável dependente aleatória e seu preditor determinístico por meio de uma função de ligação, que conecta a média das observações à parte sistemática do modelo, garantindo maior precisão no ajuste (Carvalho *et al.*, 2011).

4.4.5 Modelo de Spurr

O modelo de Spurr (1952) é amplamente utilizado na estimativa volumétrica de árvores individuais, sendo uma das equações mais empregadas em inventários florestais devido à sua simplicidade e eficiência. A estrutura matemática do modelo é dada por:

$$V_i = \beta_0 + \beta_1 \cdot D_i^2 \cdot H_i + \varepsilon_i$$

Em que:

V_i = volume em m³ da árvore i

D_i = diâmetro a altura do peito (1,3m de altura)

H_i = Altura total da árvore (m) da árvore i

β_0 e β_1 = parâmetros do modelo

ε_i = erro aleatório

O modelo de Spurr é uma equação volumétrica que relaciona o volume de uma árvore ao seu diâmetro e altura, sendo frequentemente utilizado devido à sua fácil aplicação e interpretação biológica. Uma das principais vantagens desse modelo é a consideração do termo $D^2 \cdot H$, que reflete a relação geométrica do volume cilíndrico, permitindo uma boa aproximação do volume real da árvore.

Além da forma original linear, o modelo pode ser transformado pela aplicação do logaritmo neperiano (ln), resultando em uma equação linearizável que melhora a normalidade

dos resíduos e reduz a heterocedasticidade dos dados. Essa transformação facilita a estimação dos parâmetros por meio de regressão linear múltipla.

Diversos estudos têm demonstrado a aplicabilidade do modelo de Spurr em diferentes contextos florestais. Por exemplo, em uma análise de modelos volumétricos, Andrade *et al.* (2019) observaram que o modelo de Spurr foi um dos mais utilizados, presente em 65% dos trabalhos analisados sendo frequentemente preferido pela sua simplicidade e facilidade de aplicação.

4.4.6 Modelo de Silva- Bailey

Silva (1986) usou a teoria que Patrick Leslie desenvolveu em 1945, para estudar o crescimento de árvores de *Pinus elliottii*. A teoria diz que a estrutura da população de árvores pode ser expressa na forma de um sistema de equações lineares. Silva observou que a estrutura da população muda de acordo com a fonte de nutrientes. A teoria de Leslie diz que, para qualquer distribuição arbitrária de árvores em uma população, a estrutura da população pode ser expressa na forma de um sistema de equações lineares. Essas equações descrevem a mudança na distribuição da população ao longo do tempo.

O modelo de Silva-Bailey expressa o tamanho da população (W_{i+1}) no tempo $i+1$ como uma função do tamanho da população no tempo i (W_i), da taxa de sobrevivência (S) e do número médio de recrutamentos (F), conforme o modelo.

$$W_{i+1} = W_i(S.F)^n$$

Em populações florestais, o recrutamento (nascimento) e a migração (movimento de indivíduos de uma população para outra) são geralmente desprezíveis. A mortalidade, por outro lado, pode ser modelada com relativa facilidade.

Como a migração são desprezíveis em populações florestais, o modelo de Leslie pode ser simplificado para uma população fechada. Neste caso, o parâmetro F , que representa o número médio de recrutamentos, pode ser retirado do modelo. Desta forma, Silva (1986) considerou o plantio de *Pinus* como uma população fechada e o parâmetro F pode ser retirado do modelo. O modelo apresenta a seguinte forma:

$$W_{i+1} = W_i \cdot e^{\beta \cdot (k^{t_{i+1}} - k^{t_i})} + \varepsilon_i$$

Este modelo simplificado foi utilizado na área florestal por Silva e Bailey (1991), sendo válido para muitos plantios florestais, nos quais as árvores são plantadas em um local e mantidas até a colheita.

4.4.7 Modelo de Clutter e Jones

Clutter e Jones (1980) desenvolveram um modelo matemático para prever o número de árvores em plantações desbastadas e não desbastadas de *Pinus elliottii* em uma idade t . Este modelo se baseia em uma equação diferencial que descreve a dinâmica de crescimento e produção florestal, sendo amplamente utilizado em estudos de manejo e prognose florestal.

A equação diferencial básica do modelo é expressa como:

$$\frac{dW}{dt} = k_1W - k_2W^\beta$$

Em que:

W = representa a variável dependente, como volume em função da idade t ;

t = é a idade do povoamento;

k_1 e k_2 = são constantes de crescimento, específica do sítio e das condições de manejo;

β = é um parâmetro que controla a taxa de crescimento.

A solução desta equação diferencial resulta em uma forma funcional que permite estimar o volume ou biomassa ao longo do tempo.

O modelo ajustado pode ser escrito como:

$$W_{i+1} = \left[W_i^\beta + \alpha (t_{i+1}^\gamma - t_i^\gamma)^{\frac{1}{\beta}} \right] + \varepsilon_i$$

Em que:

W_i = representa o peso (volume) no tempo t_i ;

W_{i+1} = representa o peso estimado no tempo t_{i+1} ;

β = parâmetro que define a forma do crescimento;

$(t_{i+1}^\gamma - t_i^\gamma)$ = A diferença entre os tempos t_{i+1} e t_i , elevada à potência γ . Este termo reflete como o crescimento acumulado varia com o tempo.

γ = parâmetro que define como a idade da árvore influencia o incremento de crescimento.

$(t_{i+1}^\gamma - t_i^\gamma)^{\frac{1}{\beta}}$ = Ajusta o impacto da idade com base na forma do crescimento (β);

ε_i = erro aleatório.

4.4.8 Modelo de Schumacher e Hall

Diversos modelos são empregados para estimar o volume de madeira com base no diâmetro e na altura das árvores, e um dos mais amplamente utilizados na pesquisa florestal é o modelo de Schumacher e Hall que foi desenvolvido em 1933 por F. X. Schumacher e F. D. Hall. Eles propuseram um modelo no qual apresentaram uma forma logarítmica para estimar o volume de árvores com base em variáveis dendrométricas, como diâmetro medido a 1,3 metros de altura (D) e altura total da árvore (H). Este modelo Schumacher e Hall em 1933. Esse modelo se destaca por proporcionar estimativas precisas e estáveis, sendo uma das abordagens mais comuns devido à sua capacidade de gerar resultados não tendenciosos em termos estatísticos.

A forma clássica do modelo logarítimo é expressa como:

$$\ln(V_i) = \ln \beta_0 + \beta_1 \ln(D_i) + \beta_2 \ln(H_i) + \ln \varepsilon_i$$

Ou, de forma exponencial

$$V_i = \beta_0 \cdot D_i^{\beta_1} \cdot H_i^{\beta_2} \cdot \varepsilon_i$$

Em que:

V = volume da árvore;

D = diâmetro medido a 1,3 m;

H = altura total da árvore;

β_0 = intercepto do modelo. Representa o valor logarítmico do volume quando D=1e H=1;

β_1 = parâmetro (logaritmo do D);

β_2 = parâmetro (logaritmo da altura total);

ε = erro aleatório.

4.5. POTENCIALIDADES E DESAFIOS NO USO DA APRENDIZADO DE MÁQUINA NA MODELAGEM FLORESTAL

A Inteligência Artificial (IA) é uma tecnologia promissora para a modelagem do crescimento e da produção florestal. Mesmo que as pesquisas indiquem um aumento nos números de publicações sobre a (IA) ainda precisa ser mais explorada na mensuração e manejo florestal, principalmente quando usadas para realizar previsões dendrométricas com ênfase no crescimento e na produção (Costa Filho *et al.*, 2019; Reis *et al.*, 2018).

Dentre as potencialidades da IA voltada ao manejo florestal, tem-se a utilização das técnicas que permitem combinar variáveis numéricas com categóricas (Binoti, 2010), fornecendo estimativas mais precisas do volume de árvores (Binoti *et al.*, 2014), eficiência em prever o diâmetro ao longo do tronco em eucaliptos (Costa Filho *et al.*, 2019; Mendonça *et al.*, 2015), podendo ser indicadas como ferramentas para prever o crescimento de forma eficaz e não destrutiva em eucaliptos pelas empresas do setor florestal (Monteiro *et al.*, 2023).

As Redes Neurais Artificiais (RNAs) têm demonstrado resultados promissores na previsão da produção volumétrica de povoamentos florestais, evidenciando seu grande potencial para aplicações na mensuração e manejo florestal. A escolha da configuração ideal da RNA e das variáveis de entrada envolve uma ampla variedade de combinações de parâmetros, o que possibilita ajustes finos e a otimização dos modelos propostos (Binoti *et al.*, 2014). Dessa forma, a realização de uma análise detalhada dos dados torna-se essencial para aprimorar a eficácia dos sistemas de Inteligência Artificial e assegurar que os resultados obtidos apoiem decisões estratégicas de maneira confiável.

4.5.1 Redes Neurais Artificiais (RNA)

Uma alternativa na predição da produção florestal com resultados satisfatórios é o uso das técnicas de Redes Neurais Artificiais (RNAs) (Campos; Leite, 2017). O crescente interesse pelas RNAs é atribuído as variadas características úteis e sua habilidade à adaptação podendo ser retreinada a partir de modificações ajustando seus pesos em tempo real, aprendizagem generalizada com os dados de treinamento, não-linearidade e identidade neurobiológica (Haykin, 2001).

As RNAs podem ser definidas com técnicas computacionais que apresentam um modelo inspirado na estrutura das redes biológicas (neurônios artificiais) e que adquirem conhecimentos por meio de experiência (Valença, 2015). Exemplo de uso de RNAs foram desenvolvidos por Alcântara (2015), quando propôs uma nova metodologia para prognose do crescimento e da produção com uso de redes neurais artificiais em povoamento de eucaliptos com seis e sete anos de idade no estado de Minas Gerais. Outras pesquisas também contextualizam as redes neurais artificiais com contribuições significantes para a ciência florestal (Binoti, *et al.*, 2014; Campos; Leite 2017).

A propriedade mais importante das redes neurais é a capacidade de aprender e com isto melhorar seu desempenho. Isto é obtido por meio de um processo iterativo de ajustes aplicado a seus pesos, o treinamento. O aprendizado ocorre quando a rede neural alcança uma solução generalizada para uma classe de problemas (Haykin, 2001). O algoritmo mais conhecido para

o treinamento de redes do tipo multicamadas é o *Backpropagation* (retropropagação do erro) que procura melhorar o desempenho da rede por meio do ajuste de pesos visando minimizar o erro (Valença, 2015).

A (Figura 2) apresenta a arquitetura de uma rede MLP composta por três camadas: a camada de entrada com dois neurônios que representa as variáveis independentes altura total (Ht) e Diâmetro medido a 1,3 metros do solo (D) e um neurônio para representar o limiar (bias); uma camada intermediária com 16 neurônios na camada escondida e mais um neurônio representando o (*bias*) $X_0 = +1$ e um único neurônio na camada de saída representado o volume.

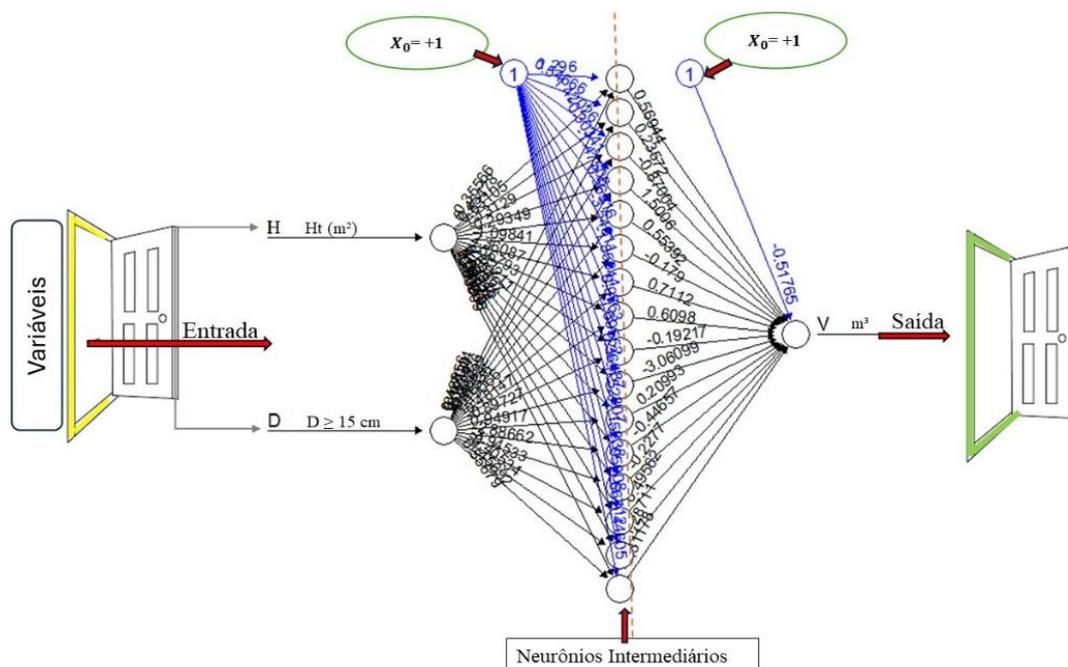


Figura 2- Arquitetura de uma rede MLP. Fonte: elaborada pela autora (2024)

A variação é calculada para os pesos conectados as unidades de processamento da camada de saída. Prosseguindo, usa-se a informação do erro para mudar os pesos chegando às unidades da penúltima camada e assim por diante até chegar à primeira camada, o treinamento ocorre em duas fases, *forward* e *backward*. Como padrão, incorpora-se o limiar diferente de zero a função de propagação (Valença, 2015).

$$\sum_i w_{ji} x_i > T_i \Rightarrow \sum_i w_{ji} x_i - T_i > 0$$

Em que w_{ji} são os pesos, x_i os padrões de entrada e T_i o limiar.

O erro de um neurônio de saída j na iteração n é definido por:

$$e_j(n) = y_d^j(n) - y_j(n)$$

sendo a soma dos erros quadráticos de todos os neurônios de saída na iteração n ao longo do ciclo em todo processo de treinamento (Valença, 2015). Neste caso o erro médio quadrático para o primeiro ciclo é dado por:

$$\varepsilon(n) = \frac{1}{2} \sum_{j=1}^n e_j^2(n)$$

Em que:

$\varepsilon(n)$ representa o erro médio quadrado no instante n , ou seja, a medida da diferença entre a previsão do modelo e os valores reais;

$e_j(n)$ o erro individual para a amostra j no instante n , ou seja, a diferença entre a saída esperada e a saída da rede;

$\sum_{j=1}^n$: indica a soma dos erros ao quadrado para todas as amostras até n ;

$\frac{1}{2}$: o fator $1/2$ é usado para simplificar a derivada do erro durante o processo de treinamento, facilitando a atualização dos pesos.

Esta fórmula reflete o custo total a ser minimizado no treinamento da rede.

No setor florestal, como em outros setores, as análises vêm evoluindo, exigindo novas abordagens para alcançar resultados mais precisos e viáveis. Embora o uso de (RNAs) seja relativamente recente na pesquisa florestal, essas redes já têm mostrado uma variedade de aplicações bem-sucedidas. Entre essas aplicações se pode citar os trabalhos de Leal *et al.* (2020) ao ajustar o modelo de Schumacher-Hall e aplicar redes neurais para estimativa de volume e altura em *Eucalyptus urophylla* S. T. Blake, concluiu que o modelo de Schumacher-Hall apresenta bons ajustes. Leal *et al.* (2015) observaram que independente da técnica utilizada (regressão ou RNA), ocorre a redução dos custos dos inventários florestais em plantios de eucaliptos.

4.5.2 Máquina de Vetor de Suporte (MVS)

Concomitante com as Redes Neurais Artificiais, outra técnica promissora na modelagem matemática em sistemas complexos é as Máquinas de Vetores de Suporte (MVS). As Máquinas de Vetores de Suporte foram desenvolvidas por Vapnik (1995), são algoritmos de aprendizado de máquina baseados na teoria de aprendizado estatístico. Segundo Karthik *et al.* (2016), as Máquinas de Vetores de Suporte (MVS) foram inicialmente desenvolvidas como algoritmos de

classificação binária e posteriormente estendidas para tarefas de regressão. A (Figura 3) representa uma MVS ajustada com um kernel linear, ilustrando como o modelo separa os dados em duas classes (-1 e 1) utilizando um hiperplano de decisão em que os pontos mais próximos a esse plano são conhecidos como vetores de suporte (*Support Vector Regressor - SVR*) e desempenham um papel crucial na definição do hiperplano.

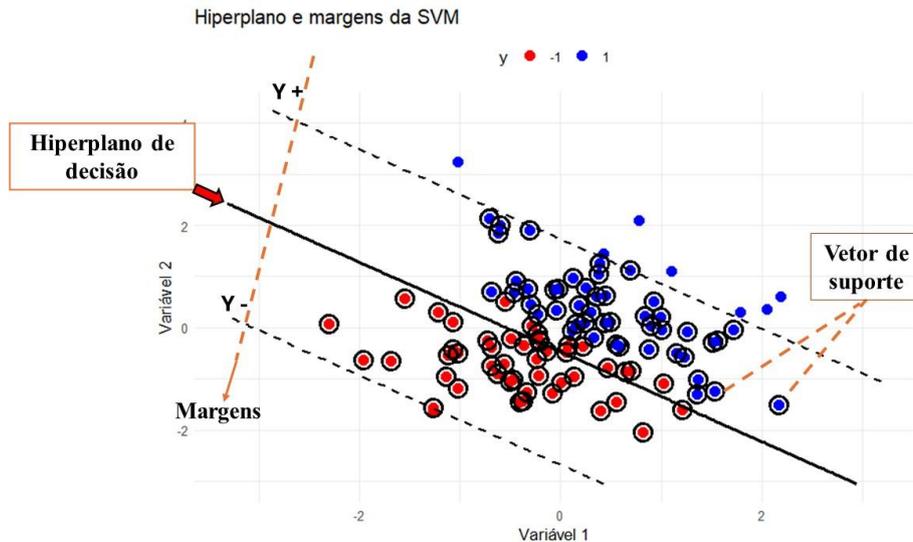


Figura 3- Representação Gráfica do Hiperplano. Fonte: adaptado de Rodriguez; Bajorath (2022).

Segundo Carneiro (2017), para lidar com uma regressão não linear, a abordagem consiste em mapear os dados para um espaço de maior dimensionalidade, onde uma função Kernel é aplicada para realizar a transformação e permitir uma separação linear nesse novo espaço. Esta característica faz da MVS uma técnica poderosa e robusta, amplamente utilizada em diversos contextos, como classificação de imagens, reconhecimento de padrões e análise preditiva em dados estruturados. Porém, este método é ainda pouco empregado em Ciências Florestais na estimação do volume de árvores, o que justifica o aprofundamento no método para estudos desta natureza (Abreu, 2019, Cordeiro *et al.*, 2022).

Sendo assim, as MVS em comparação com as RNA, têm a vantagem de conduzir a uma solução global num espaço definido. No entanto, encontrar um modelo final de MVS pode apresentar complexidade computacional, porque exige a solução de um modelo de programação quadrática e a solução de um conjunto de equações não lineares. Neste viés, estudos desta natureza, buscam por melhores estimativas, testando novas ferramentas metodológicas com o objetivo de descrever o perfil do povoamento florestal aderindo as técnicas (RNA, MVS e

modelos de Regressão), capazes de assimilar e modelar diversas variáveis e suas relações não lineares, para melhor assimilação do conhecimento (Cordeiro, *et al.*, 2015).

4.5.3 Randon Forest

Floresta Aleatória é uma técnica de aprendizado de máquina desenvolvida por Leo Breiman em 2001, foi introduzida, como uma extensão do método de "bagging" (Bootstrap Aggregating). Ela combina várias árvores de decisão para formar um conjunto robusto, o que melhora a precisão das previsões e reduz o risco de sobreajuste (overfitting) (Breiman, 2001).

Na regressão o método de florestas aleatórias (*Random Forest*) funciona como um conjunto de árvores de decisão, onde cada árvore é treinada com uma amostra aleatória dos dados. A árvore de decisão segmenta os dados em grupos com base nos valores das variáveis 1 e 2 representada por x e y a partir do nó raiz, conforme a (Figura 4) apresenta claramente as regras de decisão e as proporções das classes em cada nó, facilitando a interpretação do modelo.

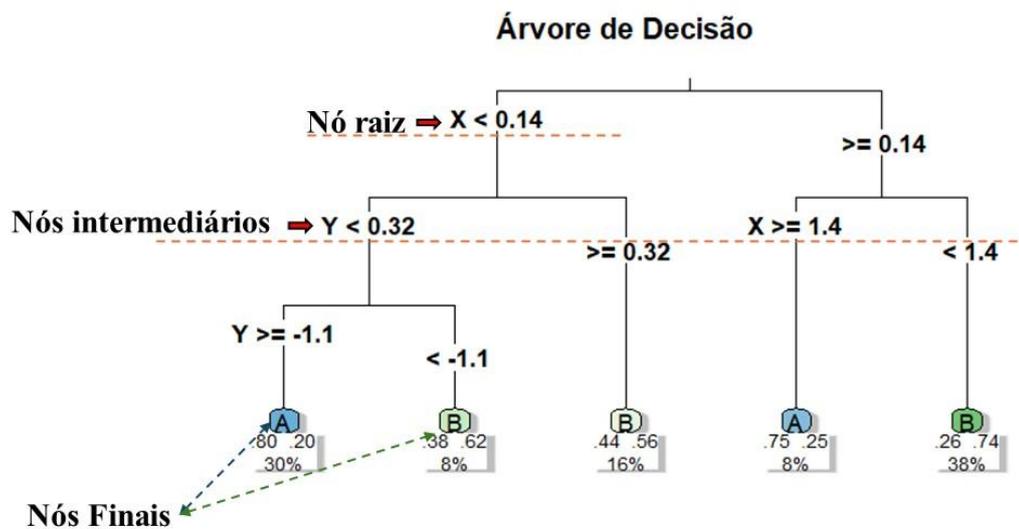


Figura 4- Árvore de decisão. Fonte: elaborada pela autora (2024)

Em cada nó, um subconjunto aleatório de variáveis de entrada, definido pelo parâmetro $mtry$, é selecionado, e a melhor divisão é determinada exclusivamente dentro deste subconjunto. Para assegurar que as árvores da floresta cresçam até o seu pleno potencial, não é realizado nenhum processo de poda (Breiman, 2001; Genuer; Poggi; Tuleau-Malot, 2010). O algoritmo combina as previsões de todas as árvores, suaviza as variações específicas de cada árvore, diminuindo a variância e, conseqüentemente, mitigando o overfitting.

O Random Forest é amplamente estudado e utilizado em aprendizado supervisionado, sendo aplicado tanto para classificação quanto para regressão. Uma das características mais

notáveis da floresta aleatória é sua capacidade de calcular a importância das variáveis, permitindo que os usuários identifiquem quais características têm maior influência nas previsões do modelo. Além disso, o uso de amostras aleatórias para treinar cada árvore garante que o modelo seja menos sensível a variações específicas dos dados, aumentando sua generalização (Gorgens *et al.*, 2015; Genuer *et al.*, 2015).

4.5.4 Algoritmo Genético

Os Algoritmos Genéticos (AG) foram inicialmente desenvolvidos por John Holland na década de 1960 (formalizados em 1975), com o objetivo de explorar mecanismos de adaptação e seleção natural, inspirados nos princípios evolutivos da teoria da seleção natural. Holland foi pioneiro na aplicação de Algoritmos Genéticos (AG) para estudar e compreender processos adaptativos presentes na natureza. Ele adaptou estes conceitos para solucionar problemas complexos, baseando-se nos princípios de seleção natural e "sobrevivência do mais apto".

Neste contexto, uma população inicial de indivíduos é submetida a operadores genéticos, como seleção, cruzamento (crossover) e mutação. Estes operadores introduzem variações nos indivíduos, permitindo que a população evolua progressivamente em direção a regiões mais promissoras no espaço de busca (Yu; Xu, 2014; Rahman; Saleh, 2018). O funcionamento do AG pode ser demonstrado de acordo com o fluxograma apresentado na (Figura 5).

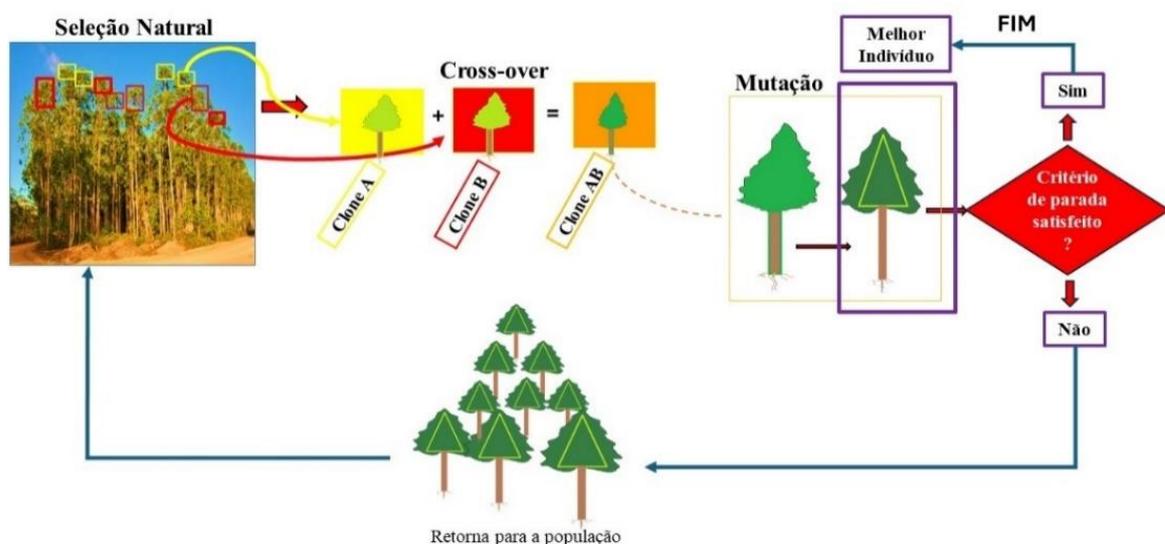


Figura 5- Fluxograma do Algoritmo Gen tico. Fonte: elaborada pela autora (2024).

Estes algoritmos aplicam conceitos de evolu o biol gica ao ambiente computacional, buscando imitar a forma como as esp cies evoluem para encontrar solu es mais eficientes.

Posteriormente, os AGs passaram a ser amplamente aplicados em diversas áreas de otimização e inteligência artificial, conforme explorado por Mitchell (1997), que destacou a versatilidade dos AGs na resolução de problemas complexos.

De acordo com Rodrigues *et al.* (2004), os Algoritmos Genéticos (AGs) demonstram ser altamente eficazes na resolução de diversos desafios relacionados ao planejamento florestal, devido às características complexas desses problemas. Entre os exemplos de aplicações se destacam: a regulação florestal, que requer a escolha de um regime de manejo específico para cada talhão; o manejo de restrições de adjacência em áreas destinadas à colheita; a otimização de rotas para o transporte de madeira; o seccionamento de toras; e atividades industriais, como o corte em empresas de móveis e fábricas de papel.

Estes problemas geralmente envolvem múltiplas variáveis e restrições, o que faz dos AGs uma ferramenta eficaz para lidar com a complexidade e as interações entre diferentes elementos do planejamento florestal. Diversos estudos na literatura confirmam a eficácia dos Algoritmos Genéticos (AGs) para resolver problemas complexos de otimização. Falcão; Borges (2001) e Gomide; Arce; Silva (2009) destacam o uso dos AGs no planejamento florestal e em outras áreas, nas quais soluções eficientes são necessárias para problemas combinatórios. Além disto, técnicas como o Enxame de Partículas (Nascimento *et al.*, 2012) têm se mostrado promissoras na resolução de problemas semelhantes.

Pode-se afirmar que o uso de Algoritmos Genéticos na mensuração florestal ainda é um campo com lacunas significativas de conhecimento, indicando a necessidade de mais estudos e aplicações para explorar plenamente seu potencial nesta área. Apesar de seu sucesso em outros aspectos do planejamento florestal, como otimização e manejo, a aplicação específica na mensuração florestal permanece pouco investigada e documentada, limitando o entendimento de suas vantagens e limitações nesse contexto.

Porém, para obter bons resultados em Algoritmos Genéticos (AG) depende diretamente da definição adequada dos parâmetros de probabilidade de crossover e mutação que serão usados. A correta configuração dessas probabilidades é crucial para garantir a diversidade da população e o equilíbrio entre a exploração do espaço de busca, promovendo assim um processo eficiente de evolução e otimização durante as iterações do algoritmo (Karine; Karem, 2019).

5. MATERIAIS E MÉTODOS

5.1 ÁREA DE ESTUDO

A área estudada pertence à empresa Amapá Florestal e Celulose (AMCEL S/A), situada no município de Porto Grande, próximo ao município de Ferreira Gomes (Figura 6) nas (Latitude: 0.85376, Longitude: -51.1941/ 0° 51' 14" Norte, 51° 11' 39" Oeste– Altitude de 4m - Classificação climática de Köppen-Geiger: Am). O clima do estado do Amapá, segundo a classificação de Köppen-Geiger, é do tipo Am (Silva *et al.*, 2020).

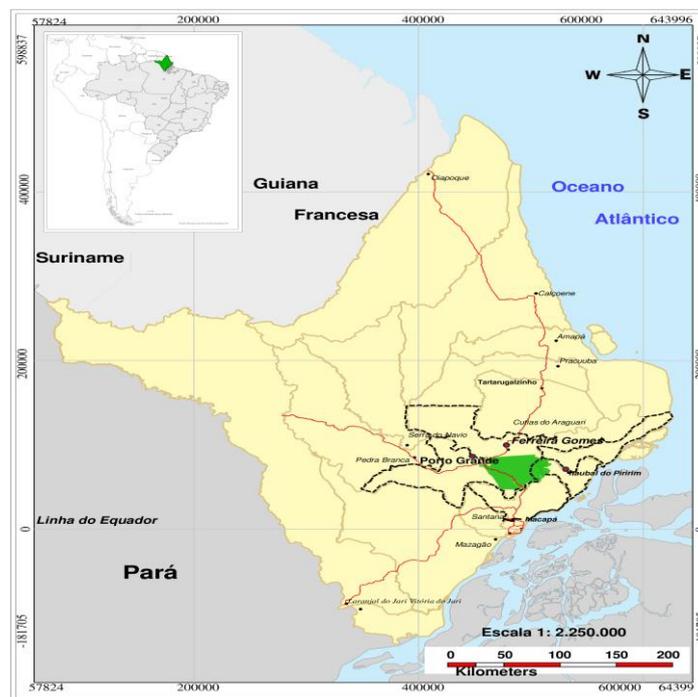


Figura 6- Localização da área de estudo (área na cor verde), plantio experimental de clones de *Eucalyptus urograndis* pela Empresa AMCEL S/A- Amapá Florestal e Celulose, localizado entre os municípios de Porto Grande, Ferreira Gomes, Itaubal e Macapá-AP, Brasil. Fonte: AMCEL (2023).

5.2 DESCRIÇÃO DOS DADOS

A pesquisa foi conduzida em um plantio permanente de clones de *Eucalyptus urograndis* entre os anos de 2014 a 2020, plantados no espaçamento 3 m x 2 m, localizado no município de Porto Grande, município do Amapá, pertencendo ao principal bloco de plantio da AMCEL S/A- Amapá Florestal e Celulose.

Os dados foram obtidos por meio de um Inventário Florestal Contínuo (IFC), com uma intensidade amostral sistemática de uma parcela a cada 10 hectares, realizadas anualmente em árvores com idades de 2 a 7 anos. Foram avaliadas 40 parcelas de 15 m x 30 m (430 m²), totalizando uma área amostral de 18.000 m² (1,80 ha). No total, foram mensuradas 733 árvores

com $D \geq 15$ cm, sendo registradas as alturas totais (H) e comerciais (h_m), além de medir o diâmetro com casca de todos os fustes.

Foi realizada a cubagem rigorosa de 165 árvores-amostra, nas quais seus respectivos diâmetros foram mensurados ao longo do fuste em posições fixas, nas seguintes alturas: 0 cm (Base); 0,30 m, 0,50 m, 0,70 m, 0,90 m, 1,10 m, 1,30 m, 1,50 m, 1,70 m e, a partir de 2,0 m, em intervalos de 2,0 m até atingir o diâmetro mínimo de 4,0 cm, sendo o volume de cada secção determinado pelo método de Smalian e o volume total com casca da árvore determinado pelo somatório sucessivo de todas as secções mensuradas em cada árvore.

$$V = \sum_i^n v_i \quad \text{sendo } v = \frac{l}{2}(g_b + g_t)$$

Em que:

v_i = volume da secção i em m^3 ;

V = volume total da árvore em m^3 ;

l = altura ou comprimento da secção i em m;

g_b = área transversal da base da secção i em m^2 ;

g_t = área transversal do topo da secção i em m^2 .

Os volumes das árvores, nas idades de 2 a 7 anos, foram calculados com base no fator de forma igual a 0,55 (Cordeiro, 2022).

$$Volume = \left(\frac{\pi}{4}\right) \times (D)^2 \times H \times 0,55$$

Em que:

$\left(\frac{\pi}{4}\right)$ é uma constante (para o cálculo da área da secção transversal do tronco, assumindo uma forma circular);

D é o diâmetro à altura a 1,30 cm do solo;

H é a altura total da árvore (m).

O fator de forma médio (f_i) foi calculado pela razão entre o volume real da árvore e o volume de um cilindro com o mesmo diâmetro e altura:

$$f_i = \frac{Volume \text{ ral}}{Volume \text{ cilindro}}$$

A seguir é apresentado na (Figura 7) o fluxograma das etapas metodológicas que inicia com o inventário florestal até a avaliação dos modelos propostos no estudo.

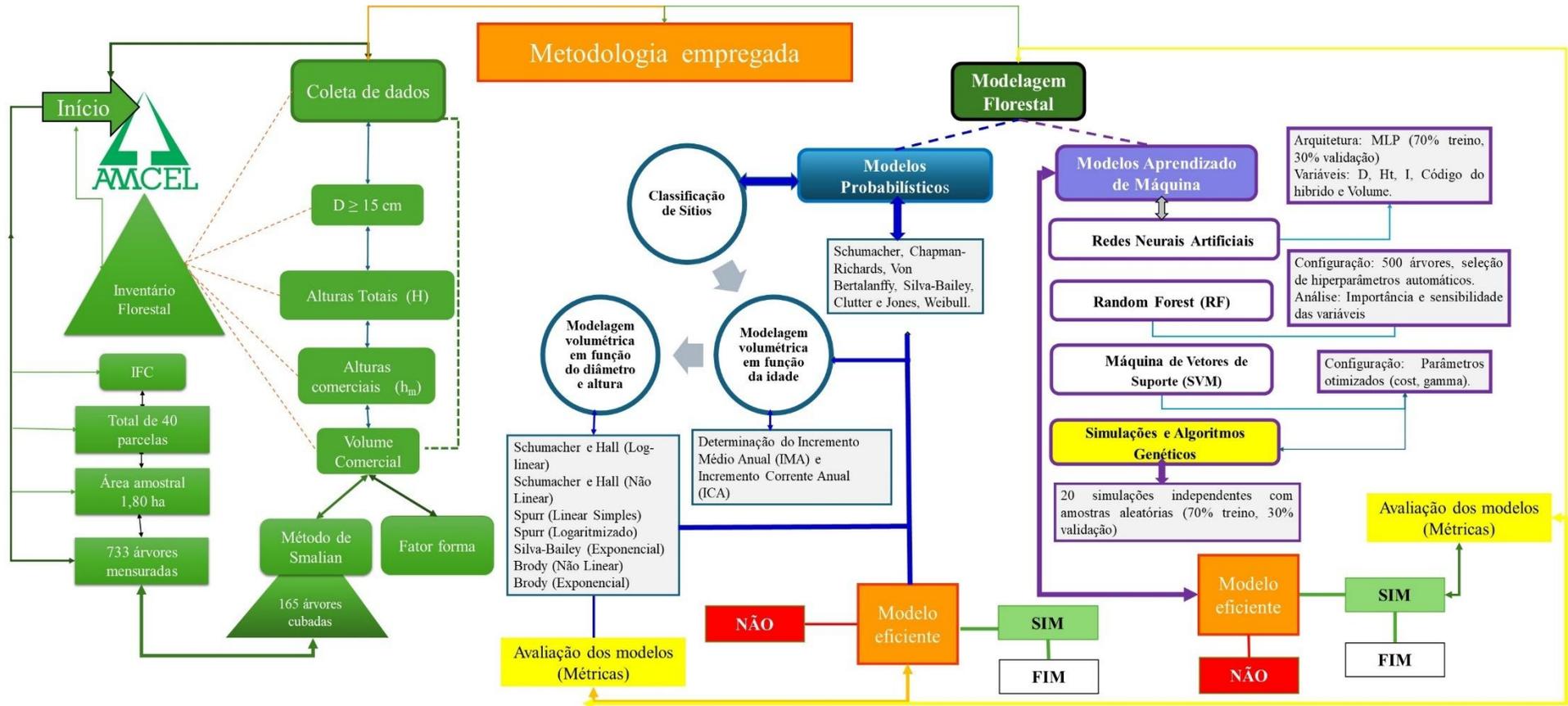


Figura 7- Fluxograma aplicado na metodologia da pesquisa. Fonte: elaborado pela autora (2024)

5.3 METODOLOGIA UTILIZADA NA MODELAGEM FLORESTAL

A metodologia foi estruturada em quatro principais etapas da modelagem, conforme ilustrado na Figura 8.

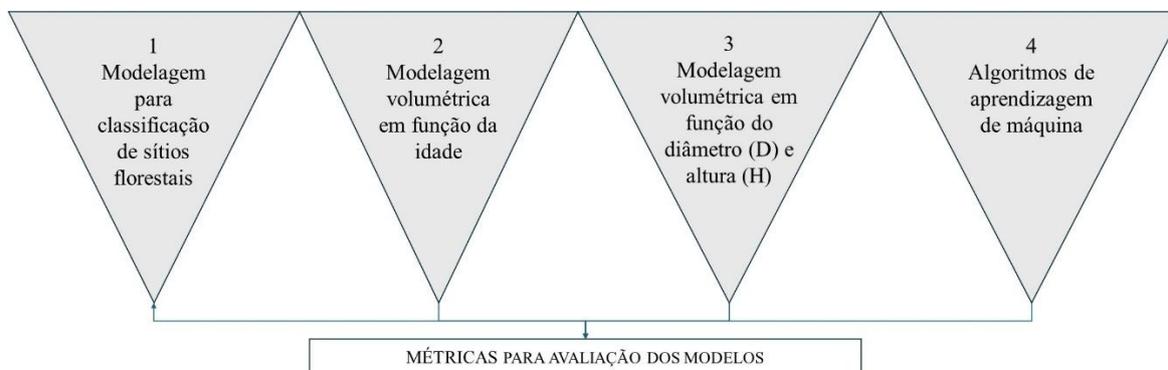


Figura 8 - Etapas metodológicas para a avaliação de diferentes modelos de estimativa volumétrica.

5.3.1 Modelagem para classificação de sítios florestais

Para a capacidade produtiva foram testados os modelos de Schumacher; Chapman-Richards; Bertalanffy; Silva-Bailey; Clutter e Jones e Weibull na forma não linear, conforme (Tabela 1). A aplicação de modelos não lineares na classificação da capacidade produtiva é amplamente justificada devido à sua habilidade de capturar o comportamento intrinsecamente não linear das variáveis envolvidas, além de possibilitar uma interpretação clara e objetiva de seus parâmetros (Regazzi, 2003). Esta abordagem permite maior precisão na análise de dados complexos, sendo essencial em estudos que envolvem relações dinâmicas, como o crescimento florestal.

Tabela 1- Modelos matemáticos para construção das curvas de Índice de Sítio

Autor	Modelo	Modelo Guia
Schumacher	$h_{dom} = \beta_0 \cdot e^{\left(\frac{\beta_1}{I}\right)} \varepsilon_i$	$h_{dom} = S \cdot e^{\beta_1 \left(\frac{1}{I} - \frac{1}{I_i}\right)} \varepsilon_i$
Chapman- Richards	$h_{dom} = \beta_0 \cdot [1 - e^{(-\beta_1 I)}]^{\beta_2} \varepsilon_i$	$h_{dom} = S \cdot \left[\frac{[1 - e^{(-\beta_1^I)}]^{\beta_2}}{1 - e^{(-\beta_1^I)}} \right] \varepsilon_i$
Bertalanffy	$h_{dom} = \beta_0 (1 - e^{-\beta_1 \cdot I})^3$	$h_{dom} = S \cdot \left(\frac{1 - e^{-\beta_1 \cdot I}}{1 - e^{-\beta_1 \cdot I_i}} \right)^3 \varepsilon_i$
Silva-Bailey	$h_{dom} = \beta_0 \cdot e^{(\beta_1 \cdot \beta_2^I)} \varepsilon_i$	$h_{dom} = S \cdot e^{[\beta_1 (\beta_2^I - \beta_2^I_i)]} \varepsilon_i$

Clutter e Jones	$h_{dom} = \beta_0 [1 + \beta_1 I^{\beta_2}]^{\beta_3} \varepsilon_i$	$h_{dom} = S \cdot \left[\frac{1 + \beta_1 I^{\beta_2}}{1 + \beta_1 \cdot I_i^{\beta_2}} \right]^{\beta_3} \varepsilon_i$
Weibull	$h_{dom} = \beta_0 [1 - e^{-\beta_1 I^{\beta_2}}] \varepsilon_i$	$h_{dom} = S \cdot \left[\frac{1 - e^{-\beta_1 I^{\beta_2}}}{1 - e^{-\beta_1 I_i^{\beta_2}}} \right] \varepsilon_i$

h_{dom} = altura dominante (m); S = índice de sítio na idade de referência (anos); I_i = idade de referência ou idade índice (anos); I = idade em (ano); e = base do logaritmo neperiano; β_0 , β_1 , β_2 e β_3 = parâmetros dos modelos a serem estimados; ε_i = erro aleatório

Método da curva-guia

O método da curva-guia, foi empregado para construir as curvas de sítio em todos os modelos analisados. Este método utiliza o modelo matemático não linear para descrever a relação entre a idade (I) e a altura dominante (h_{dom}), permitindo uma representação fiel do crescimento florestal. Com base na equação derivada da curva-guia, foi possível ajustar a equação do índice de sítio, de forma que, ao considerar a idade do povoamento igual à idade de referência (I_r), a altura dominante (h_{dom}) obtida equivale ao índice de sítio (S), expressando matematicamente a relação $h_{dom} = S \leftrightarrow I = I_r$ (Fiorentin *et al.* 2017).

Para determinar os limites dos sítios florestais, foram calculados os fatores de proporção a partir da razão entre as alturas dominantes observadas e estimadas. Estes fatores permitiram estabelecer os limites inferior e superior das classes de sítio, com uma amplitude de 4 metros.

Por fim, cada parcela foi classificada com base nos limites determinados, e foram geradas as curvas de sítio para representar as seis classes identificadas. Adicionalmente, foram analisados graficamente os resíduos e construídas curvas de sítio utilizando o método da curva-guia.

Ajustes dos modelos e métricas avaliadas

Após a classificação do sítio, a relação entre o índice de sítio e a distribuição dos pares de observações de idade e altura dominante (h_{dom}) foram analisadas utilizando modelos matemáticos não lineares. Também foi aplicado o teste de agrupamento de Scott-Knott (Scott; Knott, 1974) para classificar os modelos em grupos homogêneos.

5.3.2 Modelagem volumétrica em função da idade

Após o processamento dos dados de volume das árvores, sete modelos não lineares foram ajustados. As estimativas de volume com casca foram obtidas pelo ajuste dos modelos apresentados na (Tabela 2), na qual foi estimado o volume comercial das árvores (m^3) que melhor

se adequa ao plantio de eucalipto. Ressalta-se que os coeficientes dos modelos foram obtidos pelo método dos mínimos quadrados ordinários não lineares (MQONL).

Tabela 2 - Modelos não lineares ajustados para estimativa do volume de *Eucalyptus urograndis* no município do Amapá, Brasil

Autor	Modelo
Chapman-Richards	$V_f = V_i \cdot \left(\frac{1 - e^{-\beta_0 \cdot t_f}}{1 - e^{-\beta_0 \cdot t_i}} \right)^{\beta_1} + \varepsilon$
Von Bertalanffy	$V_f = V_i \cdot \left(\frac{1 - e^{-\beta_0 \cdot t_f}}{1 - e^{-\beta_0 \cdot t_i}} \right)^3 + \varepsilon$
Brody	$V_f = V_i - \beta_0 (e^{\beta_1 \cdot t_f} - e^{\beta_1 \cdot t_i}) + \varepsilon$
Silva-Bailey	$V_f = V_i \cdot e^{\beta_0 (\beta_1^{t_f} - \beta_1^{t_i})} + \varepsilon$
Weibull	$V_f = V_i - \beta_0 \left[e^{-\beta_1 \cdot t_f^{\beta_2}} - e^{-\beta_1 \cdot t_i^{\beta_2}} \right] + \varepsilon$
Clutter-Jones	$V_f = \left[V_i^{\beta_0} + \beta_1 (t_f^{\beta_2} - t_i^{\beta_2})^{\frac{1}{\beta_0}} \right] + \varepsilon$

Em que: V_f = volume final (m^3); V_i = volume inicial (m^3); t_f = tempo final; t_i = tempo inicial; $\beta_0, \beta_1, \beta_2$ = parâmetros do modelo; ε_i = erro aleatório.

Modelagem de crescimento e produção

Para a modelagem de crescimento e produção volumétrica foi utilizado o modelo de Clutter (Clutter, 1963) para o ajuste da curva de sítio. Estes ajustes foram realizados em dois estágios, aplicando o procedimento estatístico dos mínimos quadrados. Na primeira etapa, os parâmetros iniciais foram estimados entre o área basal inicial na idade inicial (G_1) e a área basal na idade final (G_2), considerando a influência das idades, inicial e futura do povoamento (I_1 e I_2), bem como a variável do sítio (S).

$$\ln(G_2) = \ln(G_1) \cdot \left(\frac{I_1}{I_2} \right) + \alpha_0 \cdot \left(1 - \frac{I_1}{I_2} \right) + \alpha_1 \cdot \left(1 - \frac{I_1}{I_2} \right) \cdot S + \varepsilon_i \quad (1).$$

Com os valores de G_2 obtidos no primeiro estágio, passa-se a modelar o volume de madeira (V_2). Nesta fase, o volume é relacionado ao índice I_2 , ao fator do sítio (S) e ao crescimento já estimado (G_2), através da seguinte equação:

$$\ln(V_2) = \beta_0 + \frac{\beta_1}{I_2} + \beta_2 \cdot S + \beta_3 \cdot \ln(G_2) + \varepsilon_i \quad (2).$$

Em que:

V_2 = volume, em $m^3 \cdot ha^{-1}$, na idade I_2 ;

G_1 = área basal, em $m^2 \cdot ha^{-1}$, na idade I_1 ;

G_2 = área basal, em $m^2 \cdot ha^{-1}$, na idade I_2 ;

I_1 = Idade atual do povoamento em anos;

I_2 = Idade futura do povoamento em anos;

S = índice de sítio, em metros;

$\beta_0, \beta_1, \beta_2,$ e β_3 = parâmetros dos modelos.

Construção das curvas de produção

Com base no modelo de crescimento e produção de Clutter foram construídas as curvas de incremento médio anual (IMA) e incremento corrente anual (ICA), usando os seguintes cálculos:

$$IMA = V_t / t$$

$$ICA = V_t - V_{t-1}$$

V_t = é o volume do povoamento ($m^3 \cdot ha^{-1}$) na idade em meses;

t = idade em meses.

Idade Técnica de Corte- ITC

A idade técnica de corte (ITC) foi definida no ponto em que se cruzam o IMA e ICA, ou no ponto onde o IMA foi máximo. A partir das tabelas e gráficos, foi observada a tendência das estimativas geradas e o comportamento da ITC com o aumento da capacidade produtiva.

5.3.3 Modelagem volumétrica em função do diâmetro (D) e altura (H)

Os modelos mencionados na (Tabela 3), como Schumacher-Hall, Spurr, Silva-Bailey e Brody, na modelagem volumétrica foram selecionados para o estudo, considerando aspectos como a natureza dos dados, o comportamento esperado do crescimento das árvores e a facilidade de interpretação dos parâmetros (Silva, 2015).

Tabela 3 - Modelos não lineares para a estimativa do volume de madeira (m^3 árvore $^{-1}$) em povoamento comercial de clones de *Eucalyptus urograndis*

Descrição	Autor(es)
$\ln V_i = \ln \beta_0 + \beta_1 \ln D_i + \beta_2 \ln H_i + \ln \varepsilon_i$	Schumacher e Hall (Log-linear)
$V_i = \beta_0 D_i^{\beta_1} H_i^{\beta_2} \varepsilon_i$	Schumacher e Hall (Não Linear)
$V_i = \beta_0 + \beta_1 (D^2 H)_i + \varepsilon_i$	Spurr (Linear Simples)
$\ln V_i = \ln \beta_0 + \beta_1 \ln (D^2 H)_i + \ln \varepsilon_i$	Spurr (Logaritmizado)
$V_i = \beta_0 \cdot \exp [\beta_1 (\beta_2^D + \beta_3^H)] + \varepsilon_i$	Silva-Bailey (Exponencial)
$V_i = \beta_0 + \beta_1 (-\beta_2 D_i - \beta_3 H_i) + \varepsilon_i$	Brody (Linear)
$V_i = \beta_0 + \beta_1 \cdot \exp (-\beta_2 D_i - \beta_3 H_i) + \varepsilon_i$	Brody (Exponencial)

Em que: V_i = Volume da árvore i em m^3 ; D = diâmetro medido a 1,3 metros do solo em cm; H = altura total da árvore i em m; β_0 , β_1 , β_2 e β_3 = parâmetros dos modelos e ε_i = erro aleatório.

Para agrupar os modelos de acordo com seus desempenhos estatísticos em relação à métrica de avaliação, identificando quais modelos apresentam diferenças estatisticamente significativas entre si, foi utilizado o teste de Scott-Knott.

Para verificar a similaridade entre os modelos ajustados, foi calculada a matriz de correlação. A matriz foi construída utilizando o coeficiente de correlação de Pearson para complementar a análise, variando de -1 a 1, onde:

- Valores próximos de 1 indicam uma correlação positiva forte, ou seja, os modelos produzem previsões muito semelhantes.
- Valores próximos de -1 indicariam uma correlação negativa forte (o que não foi observado).
- Valores entre -0,4 e 0,4 indicam ausência de correlação.

5.3.4 Algoritmos de aprendizagem de máquina

5.3.4.1 Redes Neurais Artificiais

A metodologia empregada no desenvolvimento deste estudo envolveu a aplicação de diferentes etapas analíticas para avaliar a importância das variáveis e o desempenho de modelos de aprendizado de máquina no ajuste do volume comercial de clones de *Eucalyptus urograndis*. Desta forma, a rede treinada foi a do tipo Perceptron de múltiplas camadas, MLP (Multilayer Perceptron). Os dados foram divididos em dois subconjuntos contendo 70% dos dados, utilizado para ajustar os modelos e 30% dos dados, utilizado para avaliar o desempenho preditivo dos modelos.

A base de dados utilizada foi gerada a partir de inventários florestais contínuos, contendo variáveis dendrométricas e genéticas, incluindo diâmetro (D), altura total (H), idade (I) e volume comercial com casca. Os híbridos foram classificados por códigos (Cod.1 a Cod.7), conforme a (Tabela 4).

Tabela 4 – Arquitetura das redes neurais artificiais selecionadas na fase de treinamento para estimativa do volume de *Eucalyptus urograndis* no município de Porto Grande-AP

RNA	Arquitetura	Variáveis		Função de Ativação
		Entrada	Saída	
MLP	(10-3-1)	Cod. 1 ao Cod.7	Volume	Tangente Hiperbólica
		D I H		

Em que: Cod. 1 ao Cod 7 são os híbridos avaliados, D = diâmetro a 1,30 m do solo (cm); H = altura total (m) e I = idade em anos.

Posteriormente a RNA foi comparada com os modelos probabilísticos descritos na Tabela 5.

Tabela 5 – Modelos probabilísticos utilizados para estimativa do volume de *Eucalyptus urograndis* no município de Porto Grande-AP

Autores	Modelos
Clutter e Jones	$V = \beta_0 \cdot (D^2)^{\beta_1} \cdot H^{\beta_2} \cdot e^{(\beta_3 \cdot I)} + \varepsilon$
Schumacher-Hall	$V = e^{(\beta_0 + \beta_1 \cdot \log(D) + \beta_2 \cdot \log(H) + \beta_3 \cdot D^2) + \varepsilon}$

Em que: V = volume da árvore com casca (m³); log = logaritmo natural; D = diâmetro a 1,30 m do solo (cm); H = altura total (m), e β_0 , β_1 , β_2 e β_3 = coeficientes dos modelos a serem estimados.

Para avaliar a diferença nas raízes dos erros médios quadráticos (REMQ) entre os grupos analisados foi realizado o teste de Kruskal-Wallis. Este teste é apropriado para verificar diferenças significativas entre grupos independentes quando os dados não atendem aos pressupostos de normalidade e homogeneidade de variâncias.

Para a análise de importância das variáveis independentes (Híbridos = Cod1 a Cod.7; D; H e I) na estimativa do volume foi utilizado o algoritmo de Random Forest. Este método consiste em construir um conjunto de árvores de decisão para modelar a relação entre a variável dependente (V) e as variáveis preditoras. O modelo foi ajustado com 500 árvores (ntree=500) e a ativação da avaliação de importância das variáveis (importance = TRUE) para garantir a robustez dos resultados. A análise foi realizada no software R, utilizando a biblioteca randomForest.

Foram consideradas duas métricas de importância das variáveis para identificar a contribuição de cada variável na precisão do modelo %PAEMQ (Porcentual de aumento do erro médio quadráticos). Esta métrica avalia o impacto da permutação aleatória dos valores da variável no (EMQ) do modelo. Valores mais altos indicam maior importância para a variável. A segunda métrica foi IPN (Incremento na pureza dos nós). Variáveis com maior pureza dos nós são consideradas mais relevantes.

Posteriormente os resultados foram ordenados com base na métrica %PAEMQ, permitindo identificar as predictoras mais importantes para o modelo. O foco da análise foi determinar quais variáveis tinham maior influência no desempenho do modelo na previsão do volume comercial.

As análises gráficas foram realizadas com o pacote ggplot2. A sensibilidade das variáveis foi derivada dos índices de importância, representando a contribuição percentual relativa de cada variável no modelo.

5.3.4.2 Simulações com amostras aleatórias na estimativa do volume de *Eucalyptus urograndis* comparando modelos de aprendizado de máquina e probabilísticos

Os dados utilizados no estudo foram provenientes de medições de 733 árvores de *Eucalyptus urograndis*, incluindo as variáveis de entrada diâmetro medido a 1,3 metros do solo (D) e altura total (H), e a variável de saída volume (V). O conjunto de dados foi dividido aleatoriamente em 70% para treino e 30% para validação, garantindo que a maior parte dos dados fosse utilizada para ajuste dos modelos.

É uma prática comum dividir o banco de dados em duas partes, uma para ajuste e outra para validação. Entretanto, esta divisão se baseia no princípio de que uma amostra de 30% é representativa da população para validação do modelo. Em termos amostrais, tal justificativa não pode ser considerada como sendo verdadeira.

Para o conjunto de dados utilizado (733 árvores) foram retirada 30% (aproximadamente 220 árvores) para validação. Calculando-se quantas amostras de 220 árvores existem em uma população de 733 árvores (C_{220}^{733}) se encontra um valor extremamente grande de amostras, e assumir que a retirada de uma única amostra representa a população é aceitar a ocorrência de uma probabilidade de $p = \frac{1}{C_{220}^{733}} = 9,88 \cdot 10^{-194}$ que é um número extremamente pequeno. Para este estudo foram retiradas aleatoriamente 20 amostras de 220 árvores e ajustados pelo modelo de Schumacher e Hall e RNA.

Os ajustes das (RNA) foram realizados utilizando a função `nnet ()` da biblioteca `nnet` do software R. Os principais hiperparâmetros configurados foram: `size`, que define o número de neurônios na camada oculta, sendo 10 no exemplo; `decay`, que corresponde à taxa de regularização, fixada em 0,01; `linout = TRUE`, que especifica a função de ativação linear na camada de saída; e `maxit`, que determina o número máximo de iterações para o treinamento, estabelecido em 1000.

Os pesos sinápticos resultantes estão apresentados na (Figura 9), conectando as variáveis de entrada (D e H) aos cinco neurônios da camada oculta e os neurônios da camada oculta à camada de saída, representada pelo Volume.

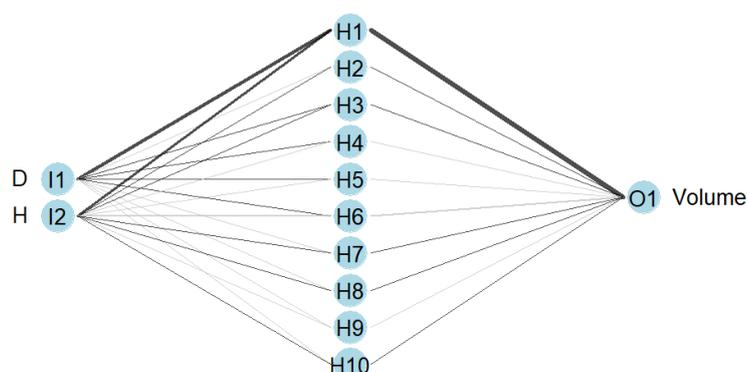


Figura 9- Arquitetura da rede neural artificial com pesos sinápticos para estimativa do volume *Eucalyptus urograndis*.

Para a estimativa do volume comercial, foi aplicado o modelo probabilístico de Schumacher-Hall log-linear. Foram realizadas 20 simulações utilizando uma divisão aleatória.

Foi aplicada uma validação cruzada holdout (Vapnik, 1995) com amostragem aleatória, onde 70% dos dados foram usados para treino e 30% para validação, repetindo esse processo 20 vezes (iterações).

Os modelos RNA e Schumacher e Hall foram ajustados e avaliados com base em métricas de desempenho. Posteriormente foi utilizado o teste Scott-Knott para avaliar o desempenho dos modelos testados. O teste de Scott-Knott foi aplicado às métricas REMQ, EMA e IAaj, a fim de identificar diferenças significativas entre os modelos avaliados e determinar quais deles apresentam desempenhos estatisticamente similares ou superiores.

5.3.4.3 Algoritmo Genético

Para a modelagem preditiva as estimativas volumétricas das árvores foram abordadas por meio de quatro técnicas de aprendizado de máquina e um modelo probabilísticos: RNA, MVS, Random Forest, e o modelo de Schumacher-Hall. Estas técnicas foram escolhidas devido à sua ampla aplicação em modelagem preditiva e à sua capacidade de lidar com dados complexos e não-lineares, características comuns em estudos florestais.

Para a análise dos modelos preditivos, utilizaram-se os seguintes pacotes no software R, que fornecem ferramentas para criação, treinamento e avaliação dos modelos:

GA. Para otimização estocástica, explicando como algoritmos genéticos podem ser aplicados para resolver problemas de otimização complexos, tanto para funções contínuas quanto discretas.

Caret é utilizado para particionar os dados e auxiliar na modelagem preditiva (Kuhn, 2008).

RandomForest, ajusta do modelo Random Forest (Liaw; Wiener 2002).

O pacote (e1071) do software R, implementa o algoritmo MVS (Máquina de Vetor de Suporte) (Meyer *et al.*, 2019).

Nnet, utilizado para treinar redes neurais artificiais (RNA) de múltiplas camadas (MLP) (Ripley, 1996).

Metrics, fornece funções para calcular métricas de erro, como REMQ e EMA (Hodson, 2022).

Metrics, foi usado para calcular o REMQ em cada iteração do modelo de Schumacher e Hall, auxiliando na otimização dos parâmetros para o modelo não-linear

Modelos Preditivos

Redes Neurais Artificiais (RNA): Utilizou-se o pacote nnet para criar redes de uma camada escondida. Foram testadas 10 diferentes configurações de redes, variando o número de neurônios na camada oculta e o parâmetro de decaimento.

Máquina de Vetor de Suporte (MVS): Utilizando o pacote e1071, testaram-se 10 diferentes configurações de parâmetros, variando o custo (cost) e o coeficiente da função de kernel radial (gamma).

Random Forest: O modelo foi ajustado utilizando o pacote randomForest. Variou-se o número de árvores (ntree) e o número de variáveis selecionadas em cada nó (mtry) em 10 diferentes combinações.

O modelo de **Schumacher-Hall** é amplamente utilizado na modelagem de crescimento e volume de árvores, devido à sua flexibilidade e capacidade de capturar a relação não-linear entre o diâmetro (D), altura total (Ht) e volume (V). O modelo baseia-se na equação:

$$\ln V_i = \ln \beta_1 + \beta_2 \ln D_i + \beta_3 \ln H_i + \ln \varepsilon_i$$

Em que: β_1 , β_2 e β_3 são parâmetros do modelo que influenciam diretamente a previsão de volume; ε_i corresponde ao erro aleatório.

6 MÉTRICAS PARA AVALIAÇÃO DOS MODELOS

6.1 Raiz do Erro Médio Quadrático

A raiz do erro médio quadrático (REMQ) é uma métrica para avaliar a qualidade de um modelo de previsão. É similar ao MSE (Erro Quadrático Médio), mas possui a vantagem de ter a mesma unidade dos dados originais, facilitando a interpretação dos resultados.

$$REMQ(Y, \hat{Y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

Em que:

Y = valores observados da variável analisada;

\hat{Y}_i = valores estimados para a variável analisada com base em um modelo;

n = quantidade total de observações referentes a y e/ou \hat{y} ;

i = cada observação específica de Y e/ou \hat{Y}_i .

6.2 Erro Médio Absoluto (EMA)

O EMA é uma métrica útil para avaliar a performance de modelos de regressão, especialmente quando a presença de outliers é uma preocupação.

$$EMA(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

Em que:

n = número de observações;

Y_i = valor observado da i -ésima amostra;

\hat{Y}_i = valor predito da i -ésima amostra;

$Y_i - \hat{Y}_i$ = Diferença absoluta entre o valor observado (Y_i) e o valor predito (\hat{Y}_i).

6.3 Índice de Ajuste

Para se obter o IAc se faz necessário o cálculo do índice de ajuste que é obtido a partir da fórmula a seguir (Schlaegel, 1981):

$$IA = 1 - \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

Em que:

IA = Índice de ajuste;

Y_i = valor observado (real);

\hat{Y}_i = valor estimado pela equação de regressão;

\bar{Y} = média dos valores observados.

Em seguida, o Índice de ajuste corrigido (IAc) foi calculado por:

$$IAc = 1 - \left[\frac{k-1}{n-k} \right] \cdot (1 - IA)$$

Em que:

IAc = Índice de ajuste corrigido

IA = índice de ajuste;

k = número de coeficientes da equação;

n = número total de observações;

6.4 Erro Padrão da Estimativa em Percentual (EPE%)

$$EPE\% = \left(\frac{100}{\bar{Y}} \right) \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-p}}$$

Em que:

\bar{Y} = média dos valores observados;

Y_i = valor observado da i -ésima amostra;

\hat{Y}_i = valor predito da i -ésima amostra;

n = número de observações;

p = número de parâmetros no modelo.

6.5 Média Quadrática dos Resíduos (MQR)

$$MQR = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - p}}$$

Y_i = valor observado;

\hat{Y} = valor predito;

n = número de observações;

p = número de parâmetros no modelo.

6.7 Coeficiente de Variação em Percentual (CV%)

$$CV\% = \left[\frac{\sqrt{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}}{(\sqrt{n - p})\bar{Y}} \right] 100$$

Em que:

\bar{Y} = média dos valores observados;

Y_i = valor observado;

valor predito;

n = número \hat{Y} = de observações;

p = número de parâmetros no modelo.

6.8 Índice de Furnival (IF) (FURNIVAL, 1961)

$$IF = \sqrt{MSR} \exp \left[-\frac{(n - 2)}{2n} \right]$$

Em que:

MQR = média quadrática dos resíduos;

n = número de observações.

6.9 Critério de Informação de Akaike (AIC) (AKAIKE, 1974)

O Critério de Informação de Akaike (AIC, Akaike Information Criterion), desenvolvido por Akaike em 1974, é uma ferramenta estatística para comparar e selecionar modelos de regressão linear. O melhor modelo é aquele que apresentar o menor valor de AIC (Akaike, 1974):

$$AIC = n \ln \left[\frac{SQRes}{n} \right] + 2p$$

Em que:

AIC = critério de informação de Akaike;

n = Número total de observações no conjunto de dados.

$SQRes$ = Mede a soma das diferenças ao quadrado entre os valores observados (Y_i) e os valores preditos (\hat{Y}) pelo modelo.

$$SQRes = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

p = Número de parâmetros estimados no modelo, incluindo o intercepto (b_0) e os coeficientes das variáveis independentes.

\ln = Logaritmo natural, usado para calcular a penalização associada à soma dos resíduos.

6.10 Critério de Informação de Akaike Corrigido (AICc) (Bozdogan, 1987)

$$AIC_c = AIC + \frac{2p(p+1)}{n-p-1}$$

Em que:

AIC = critério de Akaike;

p = número de parâmetros no modelo;

n = número de observações.

6.12 Valor Ponderado dos Escores Estatísticos (VPEE)

O VPEE atribui valores a cada equação por teste, em que a melhor equação em um teste recebe peso 1, a segunda recebe peso 2, e assim por diante. O VPEE é expresso por:

$$VPEE = \sum_{i=1}^n (Nr_i)(P_i)$$

Em que:

Nr_i = número de ocorrências na i -ésima classificação;

P_i = peso da i -ésima classificação.

O gráfico de resíduos versus valores estimados apresenta a diferença entre os valores observados e os valores preditos pelo modelo, plotada em função das previsões.

$$Res (\%) = \left(\frac{v_{obs} - v_{est}}{v_{est}} \right) 100$$

Em que:

$Res (\%)$ = resíduo em porcentagem;

v_{obs} = volume observado (m³);

v_{est} = volume estimado (m³).

Além do gráfico de dispersão dos resíduos, foi construído um histograma dos resíduos para avaliar a distribuição deles. Esta visualização auxilia na verificação da normalidade dos resíduos, que é uma das suposições fundamentais para a validade dos modelos.

6.13. Teste de Scott-Knott

Para avaliar a eficiência dos modelos ajustados, foi utilizado o teste de agrupamento de médias Scott-Knott, um método estatístico que permite identificar grupos homogêneos com desempenhos estatisticamente semelhantes. A métrica adotada para essa comparação foi a Raiz do Erro Médio Quadrático (REMQ), calculada individualmente para cada modelo e para cada grupo identificado.

Os modelos foram agrupados com base em diferenças estatísticas significativas nos valores de REMQ, adotando-se níveis de significância de 5% e 1%. Esses agrupamentos foram representados por categorias distintas (letras de "a" a "f"), onde modelos que não apresentaram diferenças significativas entre si foram classificados na mesma categoria, enquanto aqueles com desempenhos estatisticamente distintos foram separados em grupos diferentes.

6.14 Teste de Kruskal-Wallis

O teste de Kruskal-Wallis foi utilizado para verificar se existem diferenças estatisticamente significativas entre os desempenhos dos modelos preditivos analisados. Esse teste não paramétrico é indicado para comparar múltiplos grupos independentes quando a suposição de normalidade dos dados não pode ser garantida, sendo uma alternativa robusta à análise de variância (ANOVA). Para essa análise, foram consideradas as métricas de desempenho dos modelos, com destaque para REMQ.

7 RESULTADOS E DISCUSSÃO

Os resultados da análise estatística para Diâmetro (D), Altura (H) e Volume (V) revelam padrões distintos de variabilidade e dispersão entre as variáveis medidas em uma amostra de $n = 733$ indivíduos conforme a (Tabela 6).

Tabela 6 - Estatística descritiva dos valores médios das variáveis biométricas D, H e V do povoamento de *Eucalyptus urograndis* no município de Porto Grande-AP

Estatística	Variáveis		
	D (cm)	H (m)	V (m ³)
Min	5,909756	7,014771	8,581162
Q1	9,891265	14,384846	54,776601
Mediana	11,22196	17,75596	90,26343
Q3	12,75363	20,65962	129,77660
Média	11,29004	17,54098	97,06017
Max	16,95390	29,44689	318,89529
Desvio Padrão	2,052513	4,035959	52,220346
CV%	18,17985	23,00874	53,80203
Intervalo de confiança a 1% de significância	11,29004±0,19556	17,54098±0,38460	318,89529±4,97631
Tamanho da amostra	733	733	733
Erro de amostragem	1,73%	2,19%	1,56%

Q1 e Q2 = quartis, CV% = Coeficiente de variação, D = diâmetro, H = altura e V= volume.

Observa-se que os erros de amostragem foram muito baixos configurando que a amostra de 733 árvores foi suficiente para as três variáveis utilizadas nesta pesquisa, no caso, (D) e volume.

O diâmetro apresentou baixa variabilidade em relação à sua média, evidenciando uma distribuição relativamente homogênea entre as árvores avaliadas. Em contrapartida, a altura demonstrou uma maior variação em comparação ao diâmetro, o que pode estar associado a diferentes condições de crescimento dentro do povoamento. O volume, por sua vez, apresentou a maior variação relativa (CV% = 53,8%), indicando uma elevada heterogeneidade entre os indivíduos. A ampla variação no volume pode ser atribuída à influência combinada do diâmetro e da altura na determinação do volume, bem como à dinâmica de crescimento do povoamento.

7.1 Classificação de sítios de povoamentos de *Eucalypto urograndis*

Na (Tabela 7) os modelos de Chapman Richards e Weibull apresentaram o melhor desempenho, indicando alta correlação entre os valores observados e estimados. Nas métricas de

ajuste, o Chapman-Richards e o Weibull obtiveram os menores valores de AIC (830,102 e 829,608, respectivamente), sugerindo melhor equilíbrio entre qualidade do ajuste e simplicidade do modelo. Além disto, estes modelos apresentaram os menores erros médios quadráticos (REMQ), com valores de 2,1026, e os menores erros médios absolutos (EMA), com valores de 1,640 (Chapman-Richards) e 1,646 (Weibull).

Tabela 7- Coeficientes e estatísticas dos modelos testados para construção das curvas de sítio

Autores	Coeficientes				Estatística									
	b ₀	b ₁	b ₂	b ₃	AIC	AIC _c	REMQ	EMA	IAC	EPE	MQR	CV%	IF	VPEE
Chapman Richards	30,799	0,250	1,003	-	826,609	829,738	2,103	1,646	0,930	13,347	4,421	13,374	0,763	0,691
Weibull	30,828	0,248	1,001	-	829,609	829,738	2,103	1,646	0,930	13,347	4,421	13,374	0,763	0,671
Silva-Bailey	27,981	-1,919	0,659	-	830,005	830,134	2,105	1,647	0,930	13,361	4,430	13,361	0,763	0,671
Schumacher-Hall	32,754	-2,038	-	-	830,103	830,167	2,116	1,640	0,929	13,435	4,479	13,435	0,768	0,670
Clutter e Jones	1,000	7,169	0,665	1,000	836,502	830,718	2,130	1,688	0,927	13,520	4,536	13,520	0,773	0,663
Bertalanffy	22,857	0,793	-	-	845,190	849,254	2,202	1,684	0,924	13,979	4,850	13,979	0,799	0,657

Coeficientes da equação, b₀, b₁, b₂, b₃.

O índice de ajuste corrigido (IAC), que mede a adequação do modelo aos dados, variou de 0,924 (Bertalanffy) a 0,930 (Chapman-Richards), confirmando o melhor desempenho deste último. O erro padrão da estimativa (EPE) e o coeficiente de variação (CV%) foram menores para os modelos Chapman-Richards e Weibull, indicando maior precisão e menor variabilidade nas predições. Ambos também obtiveram os menores valores para o índice de Furnival (FI), sendo 0,7626, reforçando sua adequação.

Por outro lado, o modelo Bertalanffy apresentou os maiores valores de REMQ (2,202), EMA (1,684), CV% (13,98) e índice de ajuste inferior (0,924), demonstrando menor capacidade de representar os dados observados. O modelo Schumacher-Hall teve um desempenho intermediário, com métricas próximas às dos melhores modelos, mas ainda inferior ao Chapman-Richards e ao Weibull.

Com base nas análises realizadas, os modelos Chapman-Richards e Weibull foram os que apresentaram relativamente melhores desempenhos na maioria das métricas avaliadas. Apesar de os modelos Chapman-Richards e Weibull apresentarem métricas de desempenho estatístico muito próximas (como REMQ, EMA e IAC), também foram avaliados pelo teste de médias de Scott-Knott com base na raiz do erro quadrático médio (REMQ), para agrupar os modelos com desempenhos estatisticamente semelhantes (Tabela 8).

Ao nível de significância de 5% e 1%, os modelos foram agrupados em diferentes categorias (letras de "a" a "f"), indicando que as diferenças nas métricas de REMQ foram estatisticamente significativas entre os modelos.

Tabela- 8 Sumário do teste de Scott-Knott de comparação de média dos modelos testados para construção das curvas de sítio

Autores	Equações	Média REMQ	5%	1%
Bertalanffy	$h_{dom} = 22,85769. (1 - e^{-0,79351.I})^3$	2,202	a	a
Clutter Jones	$h_{dom} = 1,0000. [1 + 7,1692. I^{0,6651}]^{1,0000}$	2,129	b	a
Schumacher	$h_{dom} = 32,75472. e^{\left(\frac{-2,03842}{I}\right)}$	2,116	c	a
Silva- Bailey	$h_{dom} = 27,98114. e^{(-1,91912.0,65947^I)}$	2,104	d	a
Weibull	$h_{dom} = 30,82874. [1 - e^{(-0,24856.I)^{1,00133}}]$	2,102	e	a
Chapman Richards	$h_{dom} = 30,7998. [1 - e^{(-0,2502.I)}]^{1,0037}$	2.102	f	a

O modelo Chapman-Richards apresentou melhor desempenho nas métricas de REMQ, com diferenças de desempenho muito pequenas em relação ao modelo de Weibull. No nível de significância de 5%, o teste de Scott-Knott conseguiu separar os dois modelos em grupos distintos, mas, ao nível de 1%, todos os modelos foram considerados estatisticamente semelhantes.

A (Figura 10) apresenta os gráficos de crescimento ajustados para diferentes modelos. Os modelos Chapman-Richards e Weibull destacaram-se pelo ajuste consistente e pela boa representatividade da tendência de crescimento ao longo do tempo, evidenciando uma distribuição equilibrada dos dados observados em relação às curvas estimadas.

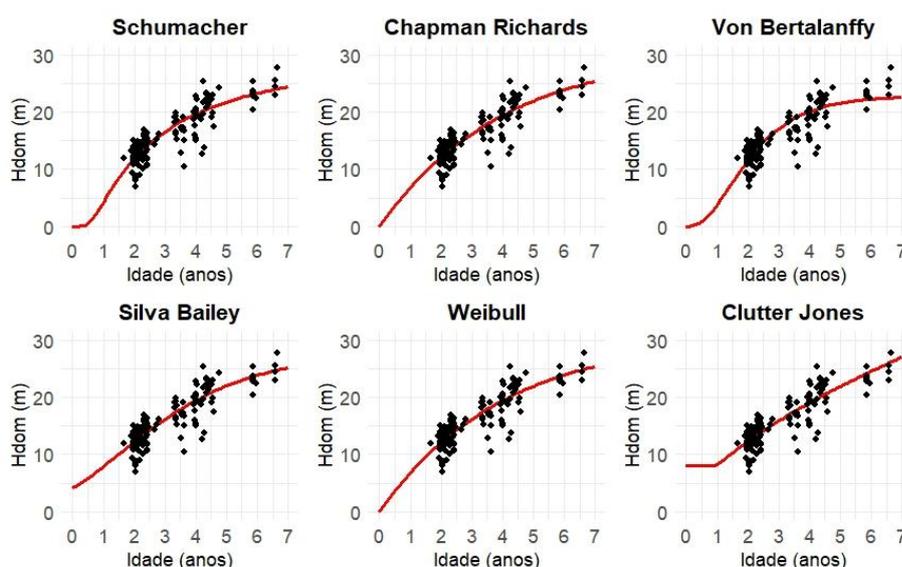


Figura – 10 Modelos de crescimento ajustados em função da idade de *Eucalyptus urograndis* no município de Porto Grande-AP.

Assim também foram os resultados da distribuição dos resíduos (Figura 11), confirmando seu desempenho superior nas métricas de qualidade de ajustes.

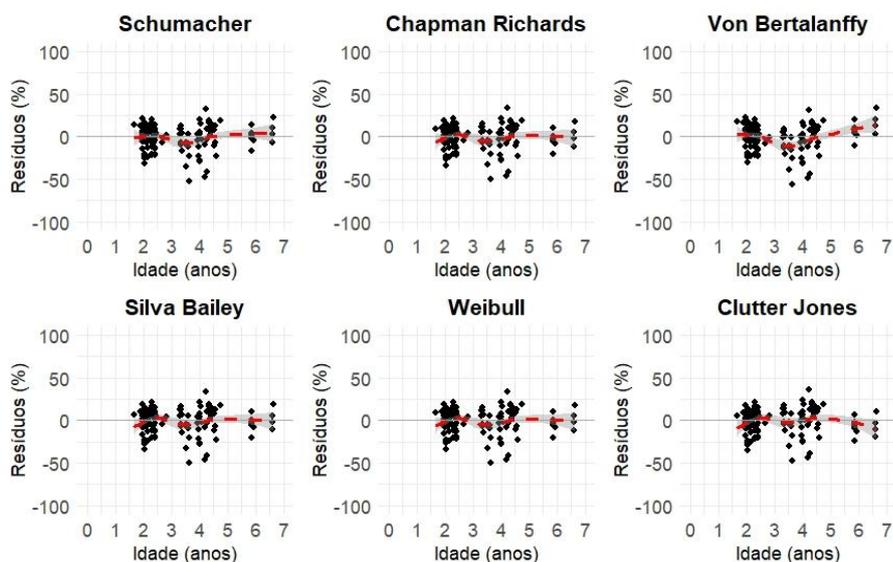


Figura 11- Distribuição dos resíduos testados na estimativa das variáveis altura dominante em função das idades de *Eucalyptus urograndis* no município de Porto Grande-AP.

Concomitante com o desempenho dos resíduos, os histogramas dos resíduos (Figura 12) apresentaram uma distribuição aproximadamente normal, centrada em torno de zero, indicando ausência de viés sistemático nos ajustes dos modelos. Esse comportamento sugere que os erros estão distribuídos de maneira homogênea, reforçando a adequação dos modelos ajustados para a estimativa do crescimento em altura dominante.

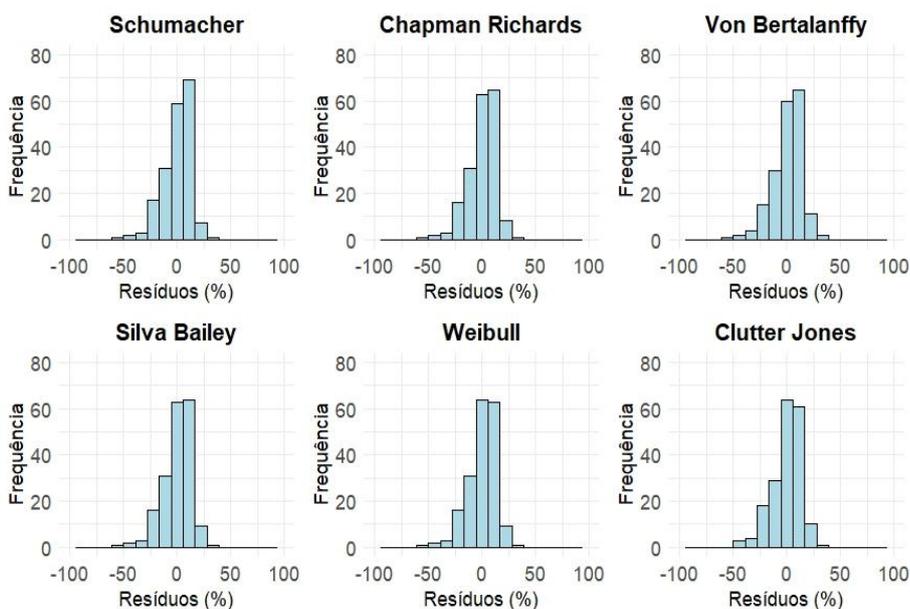


Figura 12- Histograma dos resíduos percentuais das variáveis altura dominante em função das idades de *Eucalyptus urograndis* no município de Porto Grande-AP.

Estes resultados reforçam a precisão desses modelos probabilísticos na estimativa da altura dominante em função das idades (1 a 7 anos) para o *Eucalyptus urograndis*.

Dando continuidade aos resultados dos modelos avaliados, a (Figura 13) ilustra as curvas de crescimento da altura dominante (h_{dom}) em função da idade, estratificadas em diferentes classes de sítio ($S = 12, 16, 20, 24, 28$ e 32). Estas curvas representam a capacidade produtiva dos sítios e permitem identificar a heterogeneidade das condições de crescimento para o *Eucalyptus urograndis* no intervalo de idades de 1 a 7 anos.

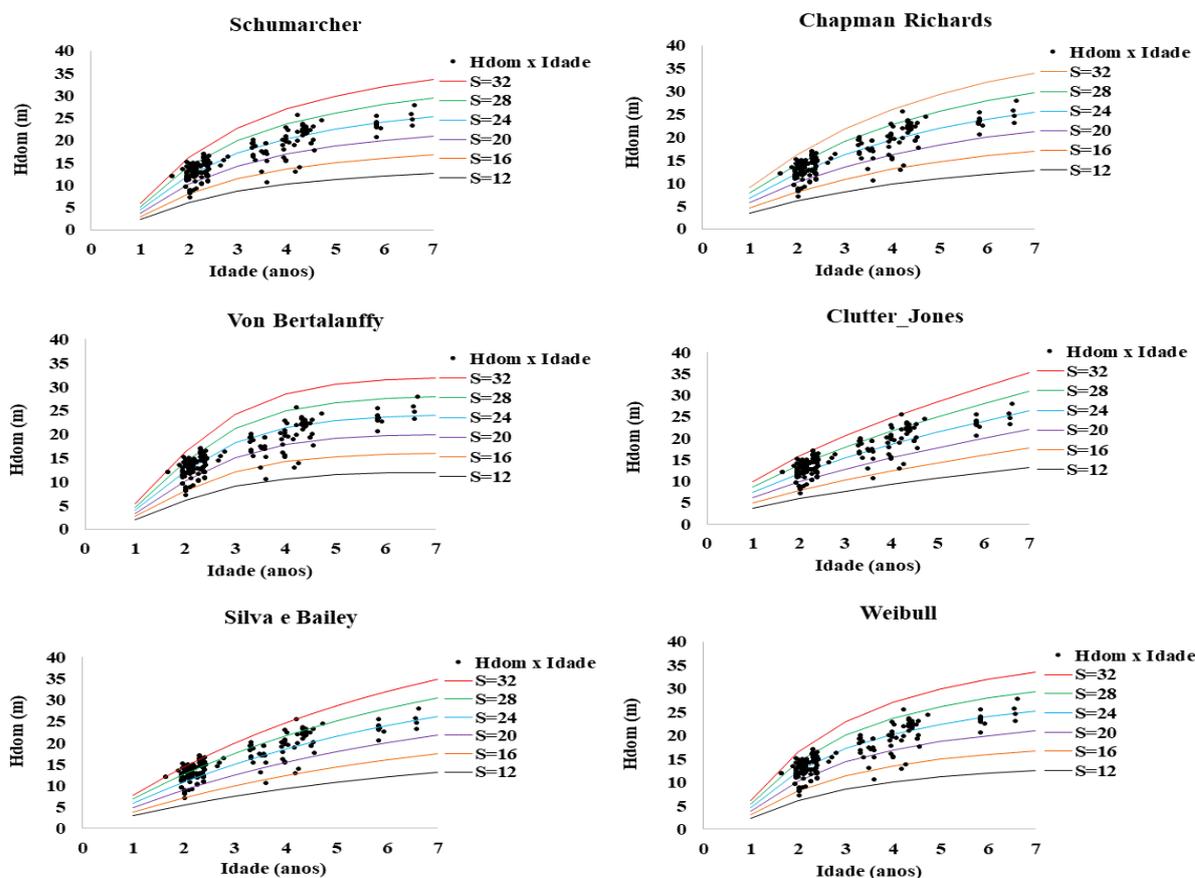


Figura 13- Curvas de crescimento da altura dominante (h_{dom}) em função da idade, ajustadas para diferentes classes de sítio em povoamentos de *Eucalyptus urograndis* no município de Porto Grande-AP.

A classificação dos sítios foi realizada com base nos parâmetros ajustados de cada modelo, gerando curvas que refletem as produtividades do povoamento de *Eucalyptus urograndis*. Observa-se que os modelos Chapman-Richards e Weibull continuam a demonstrar um bom ajuste às diferentes classes de sítio, com as curvas de crescimento apresentando coerência em relação aos valores observados (pontos pretos). Em contrapartida, modelos como

o Von Bertalanffy mostram maior dispersão em alguns pontos, especialmente nas classes de sítio superiores ($S = 28$ e $S = 32$), sugerindo limitações no ajuste.

Estes resultados diferem dos resultados obtidos por Mojena *et al.* (2019) em que o modelo de Schumacher-Hall destacou com os melhores resultados para a construção das curvas índice de sítio pelo método da curva-guia e com maior superioridade no teste de estabilidade das parcelas.

O modelo Chapman-Richards se destacou pelo bom ajuste geral, com curvas bem definidas, embora subestime ligeiramente o crescimento em classes intermediárias para idades avançadas. De acordo com Retslaff *et al.* (2015), o modelo biológico de Chapman-Richards se destacou por apresentar melhor para projeções das curvas e do índice de sítio. As curvas geradas foram consideradas apropriadas para a classificação de sítios de *Eucalyptus grandis* nos Campos Gerais, Paraná.

Com base nestes resultados a (Tabela 9) apresenta os limites das classes de sítios para o modelo Chapman Richards por ter apresentado o melhor desempenho.

Tabela 9 - Limites das classes de sítio da relação entre a idade e a altura dominante (h_{dom}) em diferentes classes de sítio, refletindo a produtividade dos povoamentos de *Eucalyptus urograndis*

Idade	Índice de Sítio					
	S=12	S=16	S=20	S=24	S=28	S=32
1	3,402	4,536	5,670	6,804	7,938	9,072
2	6,064	8,085	10,107	12,128	14,149	16,170
3	8,140	10,853	13,566	16,280	18,993	21,706
4	9,758	13,010	16,263	19,515	22,768	26,021
5	11,018	14,691	18,364	22,036	25,709	29,382
6	12,000	16,000	20,000	24,000	28,000	32,000
7	12,765	17,020	21,275	25,529	29,784	34,039

S= Sítio.

Com base nos resultados, o modelo Chapman Richards se destacou, como a melhor escolha para descrever o crescimento em diferentes condições de produtividade. Este resultado fornece subsídios importantes para a seleção da equação mais adequada no planejamento e manejo de povoamentos florestais, garantindo maior precisão na estimativa da produtividade em diferentes condições de sítio.

As curvas de sítio permitem prever a produção de madeira em diferentes idades e locais, garantindo o atendimento às demandas de produção de celulose, cavaco, biomassa e outros produtos oriundo das florestas plantas para suprir a demanda de mercado.

Desta forma, a aplicação deste modelo permitirá a empresa classificar de forma mais precisa os sítios produtivos, facilitando a identificação de áreas de maior e menor potencial de crescimento.

7.2 Modelagem volumétrica de *Eucalyptus urograndis* em função da idade

Para compreender de forma mais detalhada a produtividade dos povoamentos na estimativa de volume de eucaliptos, foram avaliados seis modelos probabilísticos que relacionam o volume e idade em função do tempo. A (Tabela 10) apresenta as equações utilizadas para estimar o volume florestal juntamente com os valores de REMQ e VPEE, que refletem a precisão e a eficiência de cada modelo.

Tabela 10 - Equações obtidas para os seis modelos avaliados para a prognose do volume (m³) de *Eucalyptus urograndis* implantados no município de Porto Grande- AP

Modelos	Equações	REMQ	VPEE
Chapman-Richards	$\hat{V}_f = V_i \cdot \left(\frac{1 - e^{-0,36700 \cdot t_f}}{1 - e^{-0,36700 \cdot t_i}} \right)^{2,56240}$	39,2296	42
Von Bertalanffy	$\hat{V}_f = V_i \cdot \left(\frac{1 - e^{-0,435785 \cdot t_f}}{1 - e^{-0,435785 \cdot t_i}} \right)^3$	39,4047	51
Brody	$\hat{V}_f = V_i - 5,215e+04 (e^{5,798e-04 \cdot t_f} - (e^{5,798e-04 \cdot t_i}))$	39,3757	58
Silva-Bailey	$\hat{V}_f = V_i \cdot e^{-4,24585 (0,61255^{t_f} - 0,61255^{t_i})}$	39,2343	45
Weibull	$\hat{V}_f = V_i - 5,235e + 03 \left[e^{-2,479e-02 \cdot t_f^{4,957e-01}} - e^{-2,479e-02 t_f^{4,957e-01}} \right]$	39,3187	55
Clutter-Jones	$\hat{V}_f = \left[V_i^{4,766e-01} + 1,055e + 03 (t_f^{4,858e-03} - t_i^{4,858e-03}) \frac{1}{4,766e-01} \right]$	35,9854	22

Em que: \hat{V}_f = Volume final da árvore em m³; V_i = Volume inicial das árvores m³; Raiz do Erro Quadrático Médio (REMQ), Valor Ponderado das Métricas Estatísticas (VPEE).

Em síntese, as equações apresentadas na (Tabela 10), utilizadas para modelar o volume em função da idade e do tempo, evidenciam que, entre os modelos analisados, o de Clutter-Jones apresentou o menor REMQ (35,985), indicando maior precisão nas estimativas de volume. Além disso, o VPEE corroborou a eficiência desse modelo para a prognose do volume de *Eucalyptus urograndis*. A (Tabela 11) apresenta os e métricas de desempenho das equações ajustadas para

cada modelo o AIC, AICc, REMQ, EMA, IAc, EPE, MQR, CV%, FI e VPPE, utilizadas para avaliar a precisão e a eficiência dos modelos.

Tabela 11 - Estatísticas e coeficientes das equações para a prognose do volume (m³) de *Eucalyptus urograndis* implantados no município de Porto Grande- AP

Autores	Coeficientes			Estatística									
	b ₀	b ₁	b ₂	AIC	AIC _c	REMQ	EMA	IAc	EPE	MQR	CV%	IF	VPPE
Clutter e Jones	4,77E+02	1,06E+06	4,86E+00	1367,582	1367,712	35,985	26,677	0,953	23,695	36,272	23,695	3,672	22
Chapman Richards	-0,367	2,5624	---	1398,384	1398,448	39,229	27,966	0,990	25,763	39,437	25,763	3,829	42
Silva-Bailey	-4,2459	0,6126	---	1398,430	1398,494	39,234	27,788	0,984	25,766	39,442	25,766	3,829	45
Von Bertalanffy	-0,4358	---	---	1398,076	1398,097	39,404	27,992	0,942	25,809	39,508	25,809	3,832	51
Weibull	5,24E+00	2,18E+01	4,96+E02	1401,246	1401,375	39,318	29,339	0,494	25,890	39,632	25,590	3,838	55
Brody	-2,15E+03	5,80E-01	---	1399,797	1399,797	39,375	29,344	0,509	25,859	39,584	25,589	3,836	58

Coeficientes do modelo, b₀, b₁, b₂.

O modelo Clutter-Jones apresentou o melhor desempenho geral, com menores valores de erro (REMQ, EMA), maior precisão (EPE, CV%) e os menor critério de informação (AIC). No entanto, o modelo Chapman-Richards também se destacou pelo IAc elevado, sendo uma alternativa para aplicação prática do modelo.

Estes resultados sugerem que, para estimativas de crescimento e produção de *Eucalyptus urograndis*, o modelo Clutter-Jones é a escolha mais adequada, com o Chapman-Richards como uma opção complementar de alto desempenho.

No campo da modelagem florestal, Clutter *et al.* (1983) caracterizaram os modelos de crescimento e produção que fornecem estimativa do volume por unidade de área como sistemas de produção explícitos. Esta categoria de modelos é particularmente útil para estimar o volume de madeira em florestas, por sua simplicidade e aplicabilidade em diversas situações de manejo florestal.

Observa-se que nos modelos de Chapman Richards, Bertalanffy, Silva-Bailey e Clutter e Jones as linhas apresentadas no modelo tendem a ficar sobrepostas. Isto indica que existe uma forte relação linear entre os valores observados e valores estimados, significando que, os valores estimados pelo modelo são próximos aos valores reais observados em campo mostrando de uma melhor forma as tendências de crescimento em volume (m³) conforme (Figura 14).

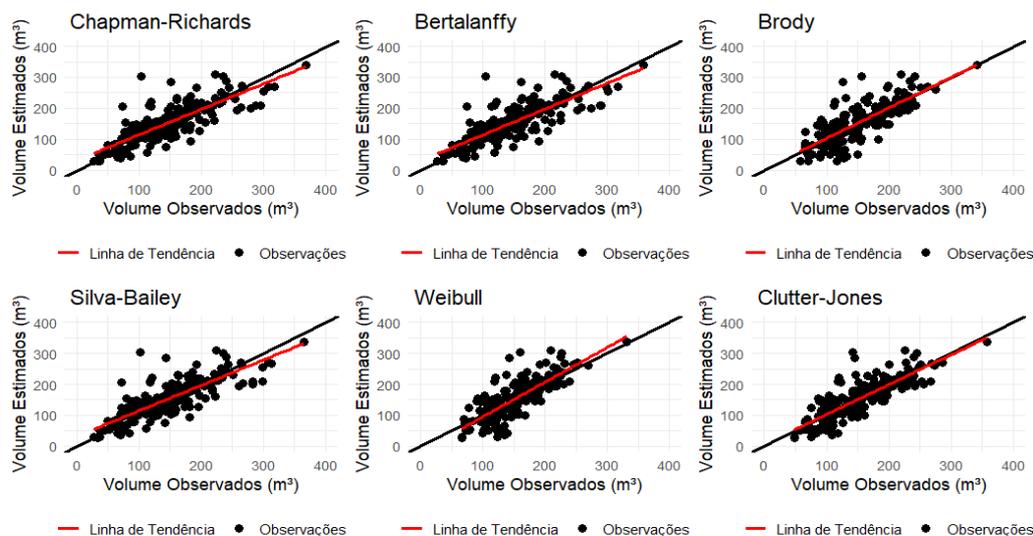


Figura 14 – Dispersão entre volumes observados e estimados para a prognose do volume (m^3) de *Eucalyptus urograndis* implantados no município de Porto Grande- AP.

Os modelos de Chapman-Richards, Bertalanffy e Silva-Bailey, em termos de capacidade preditiva, quando comparados com os demais, apresentaram maior variabilidade nos resíduos, conforme observado no gráfico de resíduos (Figura 15).

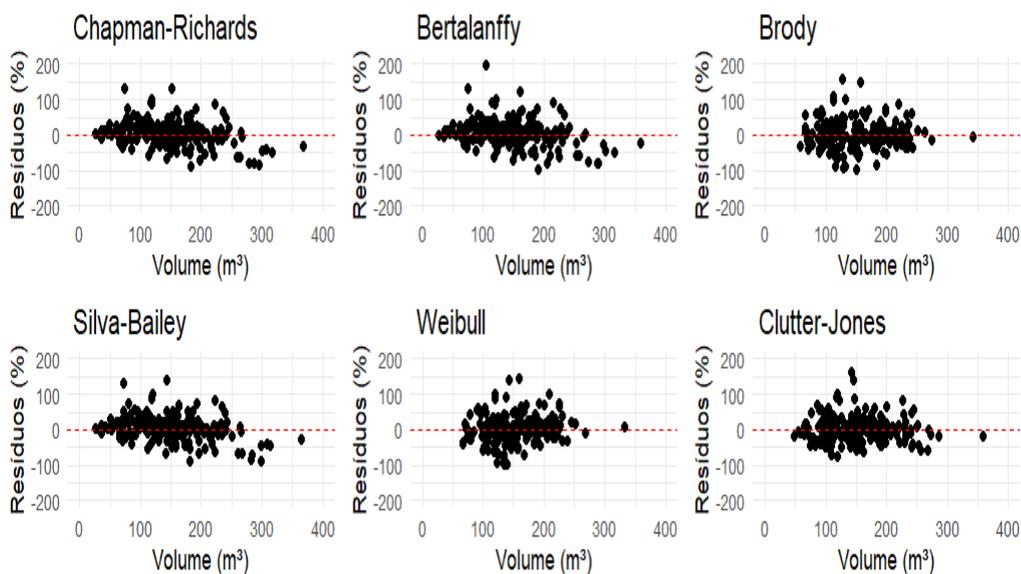


Figura 15- Dispersão dos resíduos para a prognose do volume (m^3) de *Eucalyptus urograndis* implantados no município de Porto Grande- AP.

O modelo Clutter-Jones demonstrou o melhor desempenho, com resíduos bem distribuídos e baixa dispersão. Os modelos de Chapman-Richards e Weibull também apresentaram bons desempenho, mas com ligeira tendência de erro em alguns pontos.

O modelo Silva-Bailey, embora os resíduos estejam concentrados em torno de zero, há alguns valores mais afastados para volumes baixos e altos. O modelo apresenta tendência de erro em valores extremos. Diferentemente dos modelos Bertalanffy e Brody que apresentaram maior variação e dispersão em toda a faixa de volumes, sugerindo um ajuste menos preciso em comparação aos outros modelos, resultados semelhantes foram obtidos por Silva *et al.*, (2017) ao ajustarem diferentes modelos de crescimento volumétrico tradicionais para clones de *Eucalyptus urophylla* na Região do Polo Gesseiro do Araripe, Pernambuco no qual o modelo de Clutter e Jones apresentou a melhor distribuição gráfica dos resíduos.

Desta forma, concomitante com a distribuição gráfica e residual a distribuição da frequência dos resíduos do volume estimado representa a dispersão dos valores residuais em torno de zero (Figura 16). Esta análise permite verificar se os resíduos se distribuem aleatoriamente e se seguem uma distribuição normal e quanto mais próximo concentrado do eixo zero melhor é a estimativa do modelo, bem representado pelo modelo de Clutter e Jones.

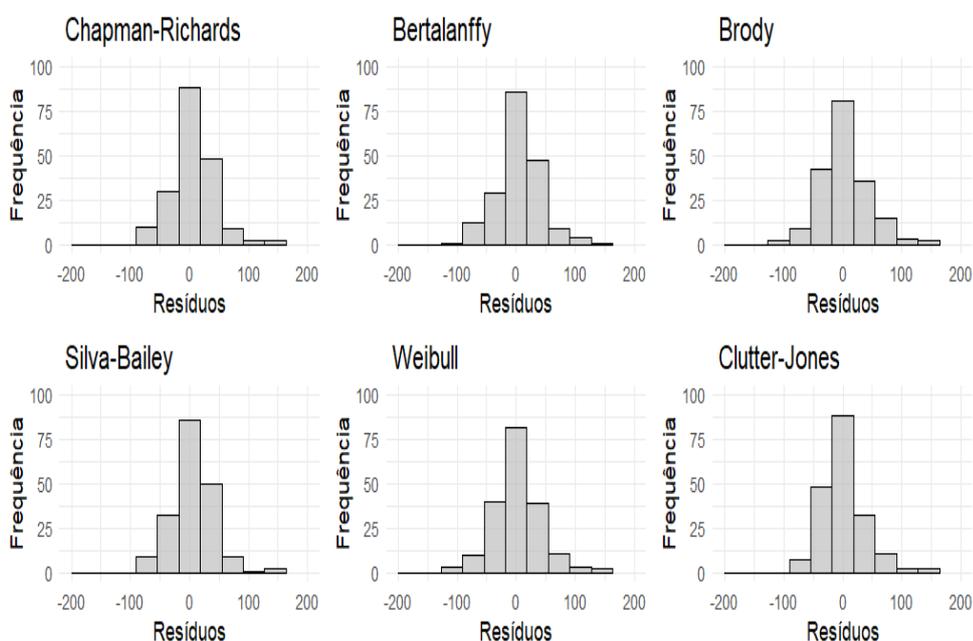


Figura 16- Distribuição da frequência dos resíduos do volume (m^3) de *Eucalyptus urograndis* implantados no município de Porto Grande- AP.

Além do mais, compreender a distribuição da frequência dos resíduos do volume estimado é crucial para avaliar os ajustes e confiabilidade dos modelos de estimativa de volume. Esta análise fornece insights valiosos sobre a variabilidade dos dados e a existência de outliers, elementos essenciais para a interpretação correta dos resultados.

O estudo de Penido *et al.* (2020) avaliou três categorias de modelos de crescimento e produção para a estimativa volumétrica em povoamentos comerciais de eucaliptos, concluindo que tais modelos proporcionam estimativas confiáveis de volume em diferentes idades e classes de produtividade, auxiliando na definição de regimes de manejo mais eficientes. Porém, as simulações em diferentes idades e classes de produtividade precisam ser confiáveis para a projeção futura do volume em povoamentos comerciais de eucaliptos.

Desta forma, a (Figura 17) apresenta o modelo de Clutter e Jones ajustado para a modelagem do volume em função da idade com respectivas curvas de sítio.

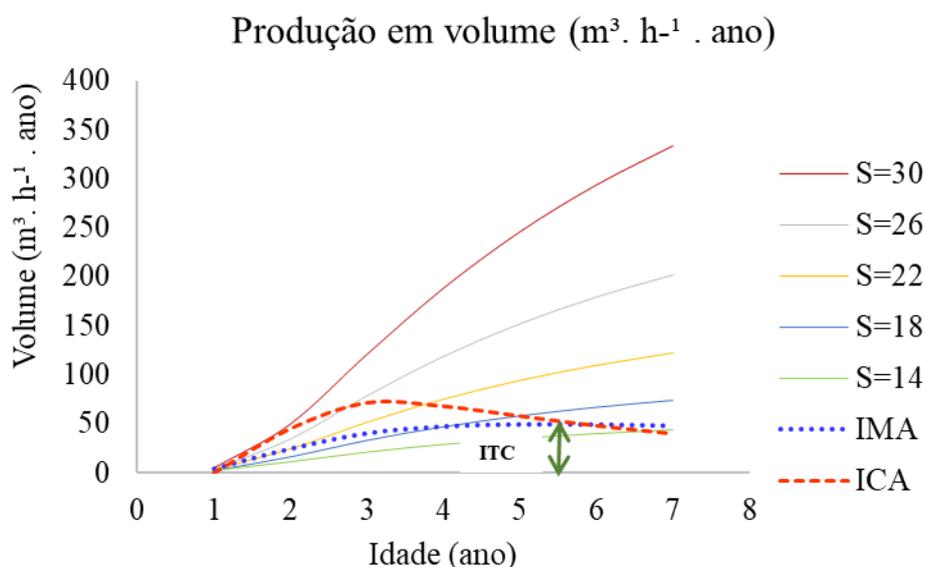


Figura 17- Idade técnica de corte (ITC), definida pela interseção das curvas de incremento médio anual (IMA) e incremento corrente anual (ICA) e as classes produtivas de *Eucalyptus urograndis* implantados no município de Porto Grande- AP.

As curvas representam diferentes classes de produtividade ($S=30$, $S=26$, $S=22$, $S=18$, e $S=14$), indicando a variação na produção volumétrica ao longo do tempo. Adicionalmente, as linhas pontilhadas representam o Incremento Médio Anual (IMA) e o Incremento Corrente Anual (ICA), evidenciando o comportamento do crescimento e a estimativa de pico de produção. O gráfico ilustra como a idade influencia diretamente o volume estimado e as diferenças associadas às condições de sítio no povoamento de *Eucalyptus urograndis* implantados no município de Porto Grande- AP, entre as idades de um a sete anos.

Para cada classe de sítio (Tabela 12), são exibidos o Incremento Médio Anual (IMA), a Produção Total Estimada ao longo do ciclo (Prod. Estimada) e a Idade Técnica de Corte (ITC), definida como o momento em que o IMA atinge o seu valor máximo. Os dados indicam que sítios de maior qualidade ($S=30$) apresentam maior volume e produtividade em relação aos sítios

de menor qualidade ($S=14$), evidenciando a importância da classe de sítio na definição do potencial produtivo de um povoamento.

Tabela 12 - Prognose do volume ($m^3 ha^{-1}$) para a produção estimada e Incremento Médio Anual (IMA) para diferentes classes de sítio no modelo de Clutter e Jones

	S=14	S=18	S=22	S=26	S=30	Produção	
ITC (ano)	2	3	4	5	6		
Prod. Estimada m^3	29,1145	46,9068	94,1742	152,3552	245,5106		
IMA ($m^3/há^{-1}/ano$)	7,2786	11,7267	18,8348	30,4710	49,1021		
Idade (ano)	V (m^3/ha)	IMA	ICA				
1	1,7919	2,3593	3,0024	3,7248	4,5309	4,5299	0,0000
2	10,9937	16,3467	23,8956	34,4892	49,2908	24,6361	44,7424
3	21,0436	33,0091	51,1935	78,7251	120,2700	40,0728	70,9463
4	29,1145	46,9068	74,9312	118,9401	187,8680	46,9457	67,5643
5	35,3755	57,9158	94,1742	152,3552	245,5106	49,0791	57,6127
6	40,2807	66,6557	109,6758	179,6980	293,4597	48,8865	47,9236
7	44,1957	73,6950	122,2874	202,1844	333,3420	47,5971	39,8608

Os dados supracitados formem informações valiosas em definir a idade de corte ideal que agregue a empresa de produtos florestais maior retorno econômico, como exemplificado na acima, em que um ITC de 6 anos o $S=30$ apresenta uma produção estimada de $293,459 (m^3 \cdot ha^{-1})$, um IMA de $48,886 (m^3 \cdot ha^{-1})$, por vezes pode ser necessário antecipar ou adiar a colheita em relação a ITC para 5 anos no $S=30$ a produção estimada passar para $245,510 (m^3 \cdot ha^{-1})$, um IMA de $49,102 (m^3 \cdot ha^{-1})$, dependendo da finalidade a idade de corte pode variar. Isto ocorre, por exemplo, em virtude de oscilações no preço da madeira.

No Brasil, o manejo florestal de eucalipto é realizado em ciclos curtos, com colheita entre 6 e 8 anos (Gonçalves *et al.*, 2013; Scolforo, 2006). A escolha da espécie para plantios comerciais se baseia no volume e no IMA, indicadores do potencial produtivo da árvore (Tonini *et al.*, 2006).

Os resultados deste estudo destacam o modelo Clutter e Jones como a alternativa mais precisa e eficiente para a modelagem volumétrica de *Eucalyptus urograndis* em função da idade, demonstrando sua superioridade na previsão do volume de madeira ao longo do tempo. A relevância prática desse resultado para a empresa florestal está na otimização da projeção da produção de madeira, possibilitando um planejamento mais estratégico da colheita e do manejo florestal. Além disso, a aplicação desse modelo contribui para minimizar incertezas nas projeções de crescimento e produção a longo prazo, garantindo maior eficiência e sustentabilidade no manejo dos povoamentos.

7.3 Modelagem volumétrica de *Eucalyptus urograndis* em função do diâmetro e altura

A (Tabela 13) apresenta as equações utilizadas para estimar o volume de árvores de *Eucalyptus urograndis*, juntamente com os valores médios do REMQ e os agrupamentos obtidos pelo teste de Scott-Knott a 1% e 5% de significância. As equações representam diferentes modelos ajustados aos dados experimentais, avaliando a relação entre as variáveis de diâmetro (D), altura (H) e volume (V).

Tabela 13 - Sumário do teste de Scott-Knott de comparação de média do volume *Eucalyptus urograndis* em função do diâmetro e altura

Autores	Equações	Média REMQ	5%	1%
Brody Não Linear	$\hat{V}_i = -0,16231 + (-0,32134 \cdot (-0,56171D_i - 1,67574H_i))$	13,12	a	a
Silva e Bailey Exponencial	$\hat{V}_i = 0,0065211 \cdot e^{[3,9105634 \cdot (3,5742115^D \cdot -2,4935909^H]}$	10,30	b	a
Brody Exponencial	$\hat{V}_i = -0,191010 + 0,094697 \cdot e^{(-1,424233D_i - 1,099860H_i)}$	10,30	b	a
Spurr Logarítmica	$\ln \hat{V}_i = \ln - 2,772149 + 0,940574 \ln (D^2 \cdot H)$	8,93	c	a
Spurr Linear Simples	$\hat{V}_i = 6,7622842 + 0,03640157 \cdot (D^2 \cdot H)$	8,87	d	a
Schumacher-Hall Log-linear	$\ln \hat{V}_i = \ln 2,72499 + 1,51514 \ln D_i + 1,23394 \cdot \ln H_i$	8,83	e	a
Schumacher-Hall Não linear	$\hat{V}_i = 0,067739 \cdot D_i^{1,606098} \cdot H_i^{1,147112}$	8,81	f	a

Em que: Ln = logaritmo neperiano, V_i = Volume das árvores i em m³; D_i = Diâmetro a altura do peito i em cm; H_i = Altura total i em m.

Entre os modelos analisados, o Schumacher-Hall Não Linear apresentou o menor REMQ (8,81), destacando-se como o mais preciso em relação aos dados. Seguido de perto pelo modelo Schumacher-Hall Log-linear (8,83), ambos foram agrupados em categorias distintas pelo teste de Scott-Knott, reforçando a superioridade destes em relação aos demais. Entretanto ao nível de 1% de significância não existe diferenças entre as equações.

Já os modelos Spurr Linear Simples (8,87) e Spurr Logarítmica (8,93) apresentaram desempenhos intermediários, sendo alocados em grupos estatisticamente distintos. Por outro lado, os modelos Silva e Bailey Exponencial e Brody Exponencial compartilharam o mesmo valor de REMQ (10,30), indicando resultados semelhantes.

A (Tabela 14) apresenta os ajustes dos diferentes modelos probabilísticos utilizados para estimar o volume comercial de *Eucalyptus urograndis* em função do diâmetro e da altura. Ela abrange uma variedade de abordagens, desde modelos lineares simples até modelos não lineares e exponenciais.

Tabela 14 - Valores dos parâmetros estimados, os índices de ajustes e os erros padrões das estimativas obtidos a partir da modelagem do crescimento em volume de *Eucalyptus urograndis*

Modelos	Função	Estatísticas									
		REMQ	EMA	IAaj	EPE	MQR	CV%	IF	AIC	AICc	PVEE
Schumacher-Hall	Não Linear	8,806	6,114	9,691	8,863	8,824	8,863	1,804	3195,25	3195,29	4
Schumacher-Hall	Log Linear	8,831	6,074	9,720	8,888	8,849	8,888	1,806	3199,35	3199,39	5
Spurr	Linear Simples	8,867	6,119	9,693	8,919	8,879	8,919	1,809	3203,39	3203,41	8
Spurr	Logarítmica	8,930	6,093	9,675	8,982	8,943	8,982	1,816	3213,82	3213,84	12
Brody	Exponencial	10,604	8,000	9,055	10,371	10,325	10,371	1,951	3425,52	3425,56	16
Brody	Não Linear	13,120	9,356	9,329	13,196	13,138	13,196	2,201	3777,69	3777,71	19
Silva e Bailey	Exponencial	10,304	8,000	9,055	10,371	10,325	10,371	1,951	3425,23	3425,56	16

Entre os modelos apresentado o Schumacher-Hall Não Linear se destacou com o menor REMQ (8,806) e maior Índice de Ajuste corrigido (IAC = 9,691), indicando uma pequena superioridade em precisão. Além disto, apresentou os menores valores de AIC (3195,26), demonstrando bom equilíbrio entre simplicidade e ajuste. Por outro lado, o modelo Brody Não Linear obteve o maior REMQ (13,1201) e o maior VPEE (19), refletindo desempenho inferior. Estes resultados sugerem que o Schumacher-Hall Não Linear é o modelo mais indicado para estimativas na área em estudo.

Resultados semelhantes foram encontrados por Miguel *et al.* (2014), que ajustaram nove modelos volumétricos para *Eucalyptus urograndis* e obtiveram IAC de 0,98, para a equação de Schumacher - Hall. Pereira *et al.* (2016) também testaram o modelo de Schumacher-Hall na forma logarítmica obtiveram os melhores resultados para a estimativa do volume de fuste com casca de *Eucalyptus urograndis*. Martins *et al.* (2021), ao testarem o modelo de Schumacher-Hall em *Eucalyptus urograndis*, obtiveram boas estimativas que cobriram adequadamente o conjunto de dados observados para ambos os volumes avaliados. Além disto, o modelo apresentou uma menor amplitude residual, indicando confiabilidade dos dados nas estimativas volumétricas.

Estes resultados indicam que, embora as diferenças entre os modelos de Schumacher-Hall e Spurr sejam mínimas, o modelo Silva-Bailey Exponencial apresenta um desempenho inferior em termos de IAC com uma diferença mais expressiva em relação ao modelo de referência, indicando que este modelo tem erros de previsão ligeiramente maiores.

Em resumo, os resultados indicam que o modelo de Schumacher-Hall Log-linear foi o mais eficiente para estimativa de volume, seguido de perto pelo modelo Schumacher-Hall Não linear, com ambos os modelos apresentando baixos erros de previsão e altos índices de ajustes. Todos os modelos conseguiram representar mais de 96% da variabilidade dos dados.

Portanto, é essencial avaliar os modelos também com base no valor ponderado das métricas REMQ, AIC e IA_c. Essa abordagem permite ranquear os modelos de forma objetiva, atribuindo menor peso àqueles que apresentam melhor desempenho nas métricas analisadas, como ilustrado na (Figura 18).

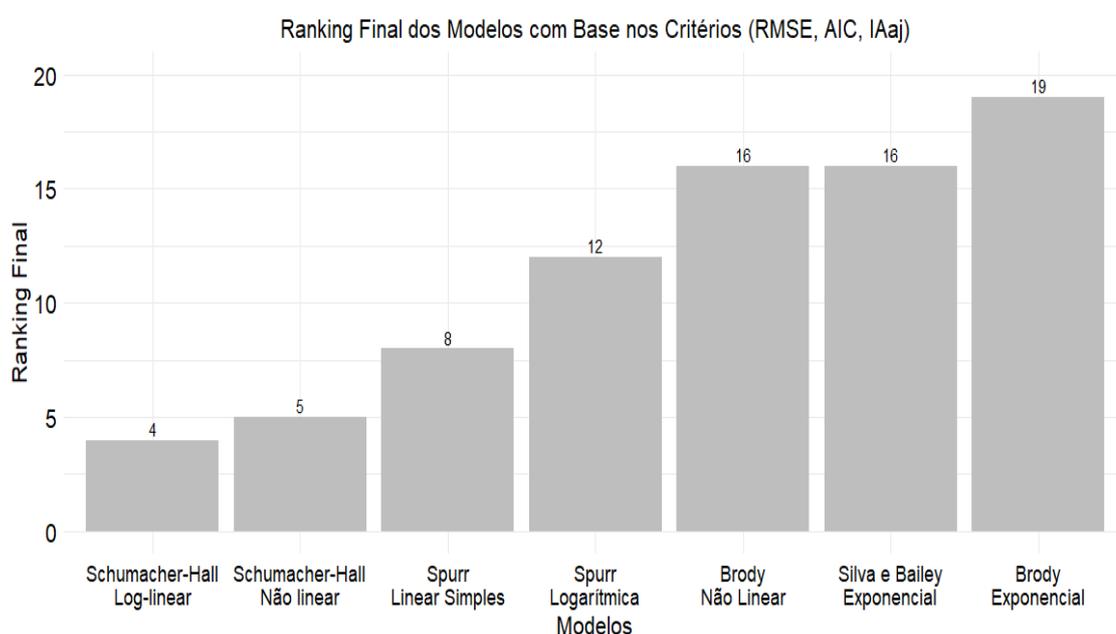


Figura 18 - Ranking final dos modelos com base nos critérios estatísticos (REMQ, AIC e IA_c).

O modelo de Schumacher-Hall Não Linear recebeu o menor VPEE (4), reforçando sua superioridade, enquanto o Brody Não Linear obteve o maior VPEE (19), evidenciando desempenho inferior. A análise gráfica apresentada na (Figura 19) mostra a diferença da dispersão dos volumes observados e estimados pelos modelos de Schumacher-Hall Não Linear e Brody não Linear.

Os modelos Schumacher-Hall Não Linear e Brody Não Linear apresentam diferenças significativas na dispersão dos volumes estimados em relação aos observados.

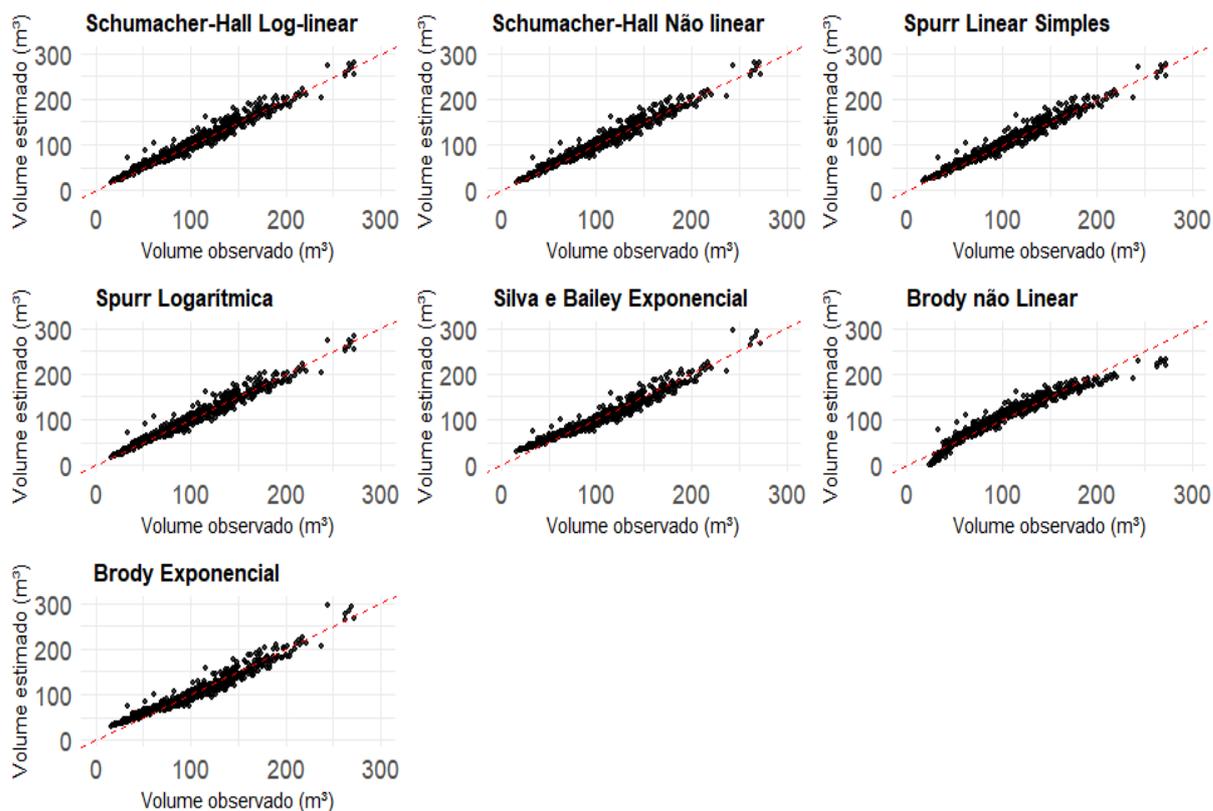


Figura 19 - Dispersão entre os volumes observados e previstos pelos modelos testados na estimativa do volume comercial de *Eucalyptus urograndis* no município de Porto Grande-AP.

Com relação ao desempenho dos modelos, os modelos de Schumacher-Hall (tanto log-linear quanto não linear) e os modelos de Spurr (linear simples e logarítmico) mostram um bom desempenho, com a maioria dos pontos bem alinhados na linha de identidade. Estes modelos demonstram previsões mais precisas para o volume dos clones de *Eucalyptus urograndis*.

O modelo Schumacher-Hall Não Linear demonstra baixa dispersão, com pontos concentrados próximos à linha de identidade, indicando alta precisão e robustez nas estimativas de volume. Em contraste, o modelo Brody Não Linear apresenta maior dispersão, especialmente em volumes mais elevados, com tendência de subestimação em alguns casos e maior variação nos extremos.

Estas diferenças refletem a superioridade do modelo Schumacher-Hall em capturar a relação entre as variáveis predictoras e o volume, enquanto o modelo Brody não linear apresenta limitações em intervalos específicos.

Concomitantemente com a figura anterior à análise da dispersão do volume, a (Figura 20) apresenta os resíduos dos diferentes modelos avaliados.

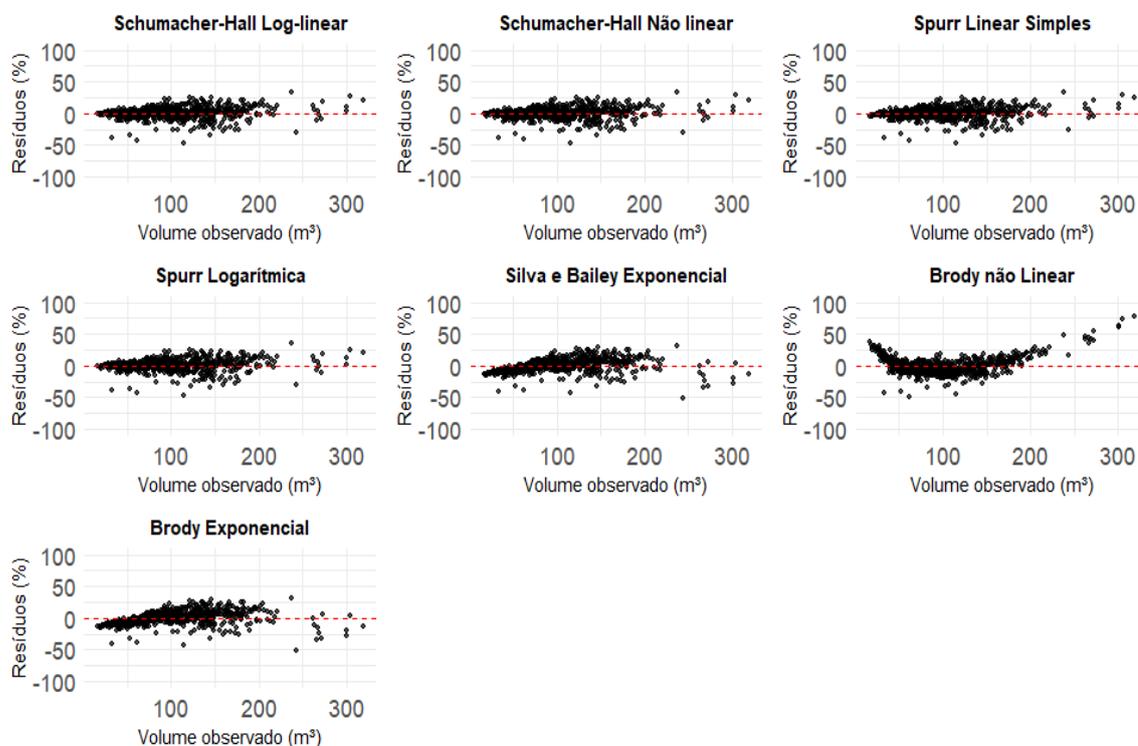


Figura 20- Distribuição residual entre os volumes observados e previstos pelos modelos probabilísticos testados na estimativa do volume comercial de *Eucalyptus urograndis* no município de Porto Grande-AP.

Conforme a distribuição residual dos modelos os modelos de Schumacher-Hall (tanto Log-linear quanto não linear), assim como os modelos de Spurr (Linear Simples e Logarítmico), mostram uma distribuição de resíduos que sugere um bom ajuste aos dados. A ausência de padrões ou tendências nos resíduos sugere que estes modelos não apresentam erros que comprometam as estimativas.

Em relação aos modelos com menor desempenho foi o modelo não linear de Brody, pois, o gráfico de resíduos mostra uma indica uma leve tendência de subestimação para volumes observados menores e uma tendência de superestimação para volumes maiores.

Desta forma, modelo de Schumacher-Hall (1933) se destacou na presente pesquisa e na área florestal é dos mais utilizados por suas robustas propriedades estatísticas, fornecendo estimativas precisas e confiáveis de volume individual de árvores, com resultados consistentemente não tendenciosos (Campos; Leite, 2017).

Os histogramas apresentados na (Figura 21) são fundamentais para verificar a adequação dos modelos. Desta forma, a normalidade dos resíduos é uma condição importante para muitos

testes estatísticos e para a validade das inferências feitas a partir dos modelos. Portanto, a simetria e a concentração dos resíduos em torno de zero observadas na maioria dos modelos sugerem que eles são bem ajustados e apropriados para a estimativa do volume comercial de *Eucalyptus urograndis*.

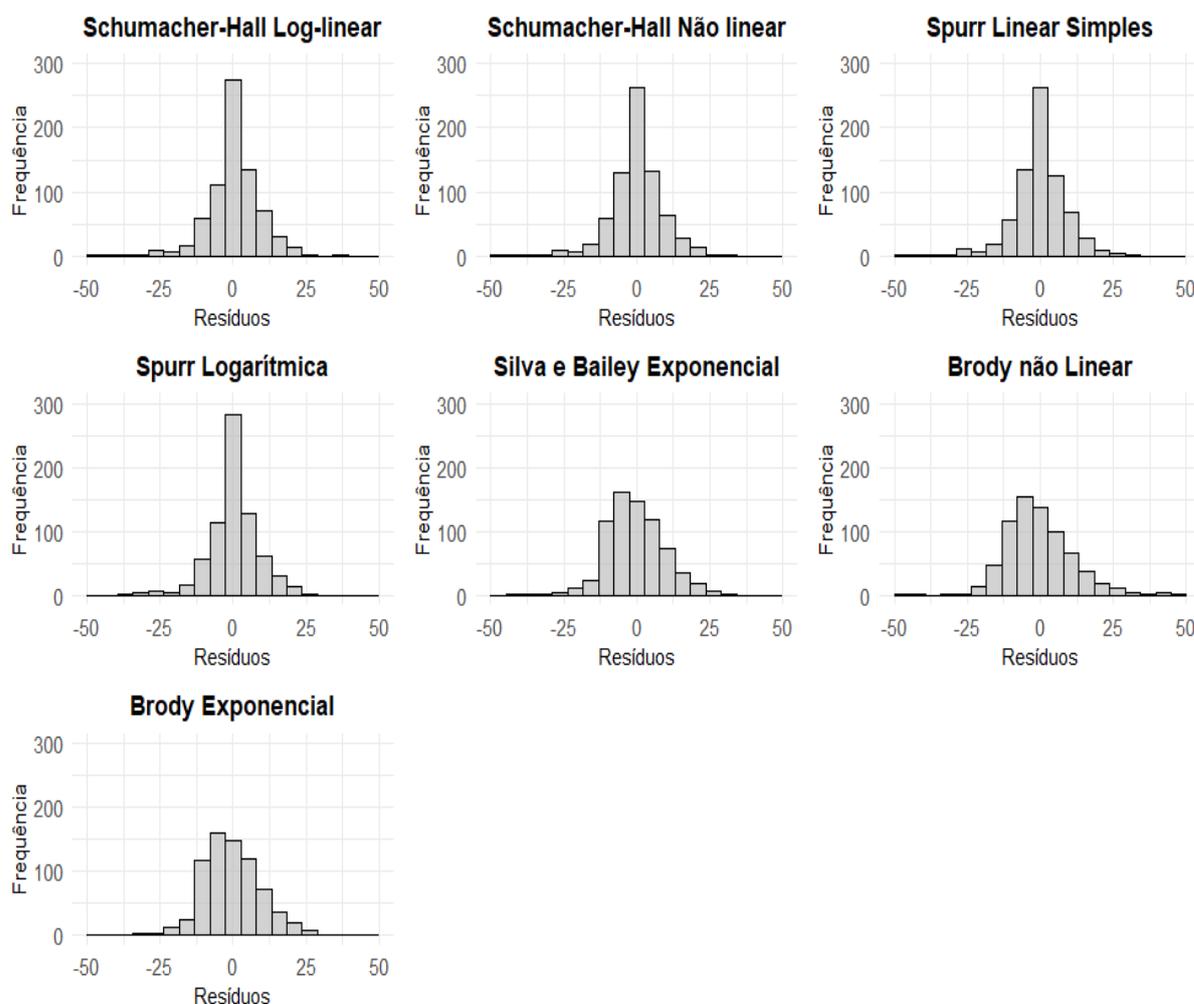


Figura 21 - Histograma de frequência dos resíduos do volume comercial ($\text{m}^3 \text{ha}^{-1}$) de *Eucalyptus urograndis* testados pelos modelos probabilísticos no município de Porto Grande-AP.

A (Figura 22) apresenta a matriz de correlação que mostra a relação entre os valores preditos pelos diferentes modelos probabilísticos utilizados para estimar o volume comercial de clones de *Eucalyptus urograndis*. A matriz de correlação apresentada demonstra o grau de associação entre os diferentes modelos matemáticos utilizados para estimar o volume de árvores.

Os coeficientes de correlação variam de -1 a 1, onde valores próximos de 1 indicam uma forte correlação positiva (resultados semelhantes entre os modelos), enquanto valores próximos de 0 sugerem pouca relação entre as estimativas dos modelos.

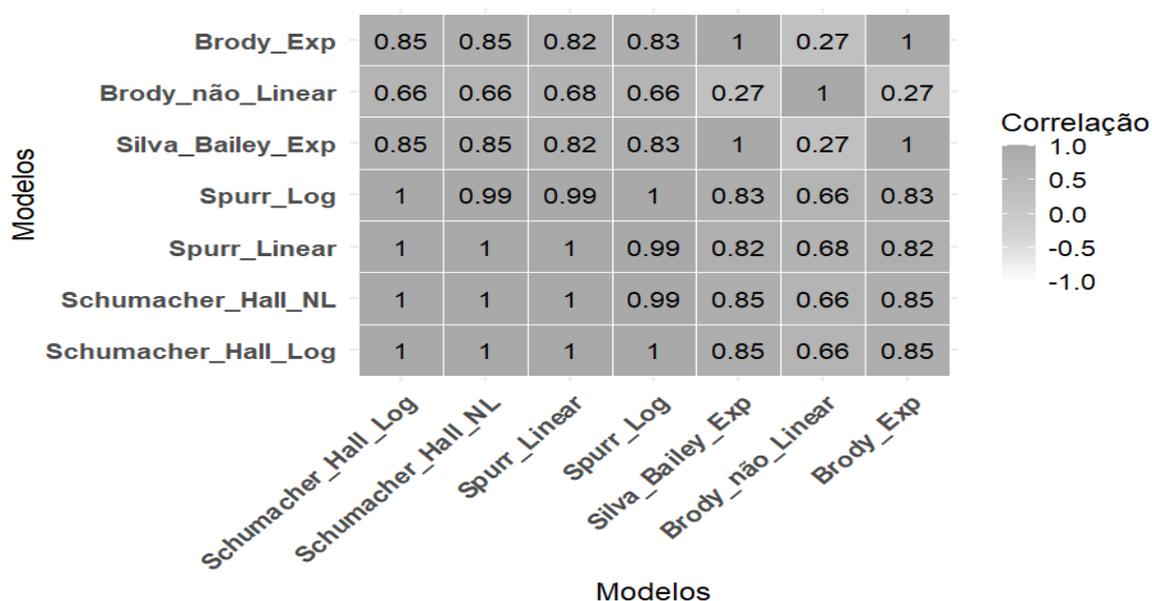


Figura 22- Matriz de correlação utilizada na estimativa dos volumes comerciais de *Eucalyptus urograndis* no município de Porto Grande-AP.

Os resultados indicam que os modelos Schumacher-Hall Log-linear, Schumacher-Hall Não Linear, Spurr Linear e Spurr Logarítmica possuem correlação perfeita (1.00), sugerindo que fornecem previsões altamente similares. Da mesma forma, os modelos Silva e Bailey Exponencial e Brody Exponencial apresentam alta correlação (~0.85), indicando que suas estimativas seguem um padrão semelhante.

Por outro lado, o modelo Brody Não Linear se destaca com as menores correlações (~0.27 a 0.66), sugerindo que suas previsões diferem significativamente das demais. Esse resultado indica que esse modelo pode capturar padrões distintos nos dados ou apresentar um comportamento diferente em relação às variáveis predictoras.

Os resultados deste estudo evidenciaram que o modelo Schumacher-Hall Não Linear apresentou o melhor desempenho na modelagem volumétrica de *Eucalyptus urograndis* em função do diâmetro e altura, sendo a alternativa mais precisa e confiável para estimar o volume comercial da espécie. A utilização do modelo Schumacher-Hall Não Linear permite um melhor dimensionamento da produção ao longo do tempo, auxiliando na previsão de estoques de madeira, maximização da produtividade e rentabilidade das plantações.

Portanto, a implementação do modelo Schumacher-Hall Não Linear pela empresa florestal representa uma estratégia decisiva para elevar a eficiência operacional, maximizando a produtividade dos povoamentos de *Eucalyptus urograndis*.

7.4 Potencial de algoritmos de aprendizado de máquina na modelagem volumétrica de *Eucalyptus urograndis*.

A (Tabela 15) apresenta as especificações dos modelos utilizados para a estimativa do volume comercial de *Eucalyptus urograndis*, incluindo a Rede Neural Artificial (RNA) do tipo Perceptron de Múltiplas Camadas (MLP) e os modelos probabilísticos Clutter e Jones e Schumacher-Hall.

Tabela 15 - Especificações dos modelos utilizados para estimativa do volume comercial de *Eucalyptus urograndis*

Modelos	Arquitetura	Variáveis		Função de ativação
RNA	(10-3-1)	Entrada (Cod. 1 a Cod. 7)	Saída	Tangente Hiperbólica
Clutter e Jones Schumacher-Hall		D I H	Volume	

Em que: Híbridos = Cod1 a Cod.7; D = diâmetro a 1,30 m do solo (cm); I=idade; H = altura total (m); V = volume da árvore com casca (m³).

A modelagem por Redes Neurais Artificiais (RNA), especificamente utilizando a arquitetura de Perceptron Multicamadas (MLP), apresentou resultados promissores na estimativa do volume comercial de clones de *Eucalyptus urograndis*. A (Tabela 16) resume os principais resultados obtidos para o modelo ajustado, destacando a arquitetura da rede, as variáveis de entrada e saída, a função de ativação utilizada e as métricas de desempenho para os conjuntos de treino e validação.

Tabela 16 - Análise das configurações e estatísticas de qualidade do Perceptron Multicamadas (MLP) usando a das funções de ativação tangente hiperbólica aplicado a determinação do volume comercial de *Eucalyptus urograndis* no município de Porto Grande-AP

RNA	Arquitetura	Variáveis		Função de Ativação	Treino			Validação		
		Entrada (Cod.1 a Cod.7).	Saída		REMQ	EMA	IAC	REMQ	EMA	IAC
MLP	10-3-1	D I Ht	Volume	Tangente Hiperbólica	8,1414	5,6172	0,975	8,8053	6,1226	0,971

A arquitetura da RNA consistiu em uma camada de entrada com 10 neurônios, representando as variáveis predictoras (Cod.1 a Cod.7, D, I, Ht), uma camada oculta com 3 neurônios, e uma camada de saída com um único neurônio, representando o volume comercial (Volume) na (Figura 23).

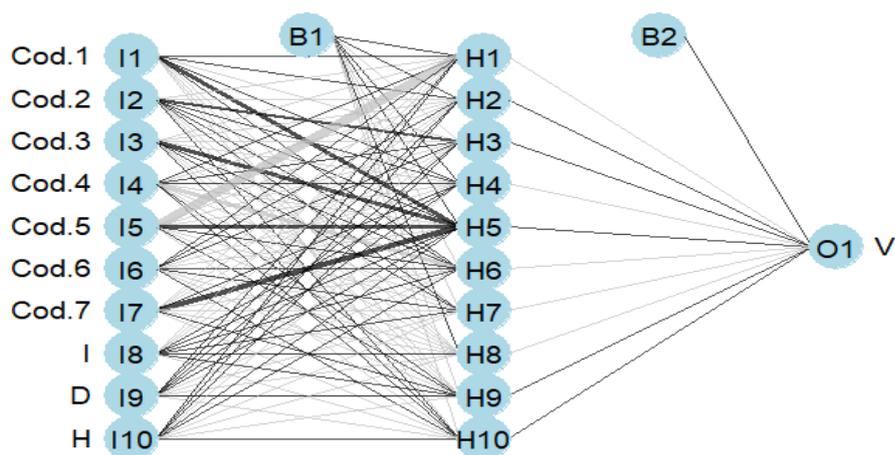


Figura 23- Representação esquemática da arquitetura da RNA. Fonte: Autora, 2024.

Foi empregada a função de ativação Tangente Hiperbólica para capturar as relações não lineares entre as variáveis de entrada e a variável resposta. Os resultados demonstraram desempenho satisfatório tanto no conjunto de treino quanto no de validação. As métricas de erro médio quadrático (REM_Q) e erro médio absoluto (MAE) indicaram bons ajustes do modelo, enquanto o índice de ajuste (I_{Ac}) revelou elevado grau de explicação dos dados pelo modelo.

No conjunto de treino, o REM_Q foi de 8,141 e o EMA de 5,620, com I_{Ac} de 0,973. Já no conjunto de validação, o REM_Q foi de 8,810 e o EMA de 6,120, com I_{Ac} de 0,931. Estes resultados indicam que a arquitetura da RNA proposta foi eficaz em capturar a variabilidade do volume comercial em função das variáveis predictoras.

A (Tabela 17) destaca os parâmetros estimados e as estatísticas de desempenho nos conjuntos de treino e validação. Os modelos Clutter e Jones, Schumacher-Hall e a Rede Neural Artificial (RNA) foram avaliados quanto ao ajuste e capacidade preditiva, utilizando métricas como REM_Q, EMA e Índice de ajuste corrigido (I_{Ac}).

Tabela 17 - Parâmetros ajustados e métricas de desempenho dos modelos testados para a estimativa do volume comercial de *Eucalyptus urograndis* em Porto Grande, AP

	Modelos	Coeficientes				Estatística					
		b ₀	b ₁	b ₂	b ₃	REMQ	EMA	IAc	CV	AIC	IF
Treino	Clutter e Jones	0,847	0,461	0,721	0,080	11,608	8,584	0,967	12,219	2670,593	11,608
	Schumacher-Hall	-0,450	0,574	0,777	0,432	10,196	7,184	0,969	10,732	2529,701	10,196
	RNA					8,180	5,658	0,974	8,610	2292,439	8,180
Validação	Clutter e Jones	0,847	0,461	0,721	0,080	10,237	7,911	0,970	10,512	1091,934	10,237
	Schumacher-Hall	-0,450	0,574	0,777	0,432	9,620	7,385	0,971	9,878	1062,975	9,620
	RNA					9,116	6,207	0,969	9,878	1040,835	9,116

Em geral, os modelos ajustados (regressão e redes) apresentaram bom desempenho na estimativa da capacidade preditiva para a estimativa do volume comercial de clones de *Eucalyptus urograndis*. As métricas de desempenho, como REMQ, EMA e Índice de ajuste corrigido (IAc), revelaram valores satisfatórios tanto nos conjuntos de treino quanto de validação, indicando que os modelos conseguem capturar adequadamente a relação entre as variáveis predictoras e a variável resposta.

Para avaliar a diferença nos erros médios quadráticos (REMQ) entre os grupos analisados foi realizado o teste de Kruskal-Wallis. Este teste é apropriado para verificar diferenças significativas entre grupos independentes quando os dados não atendem aos pressupostos de normalidade e homogeneidade de variâncias. Os resultados indicaram uma diferença estatisticamente significativa entre os modelos analisados (Kruskal-Wallis $\chi^2 = 25,806$; $df = 2$; $p\text{-valor} = 2,47e-06$).

A (Figura 24), apresenta a comparação do REMQ entre os modelos Clutter e Jones, Schumacher-Hall e RNA.

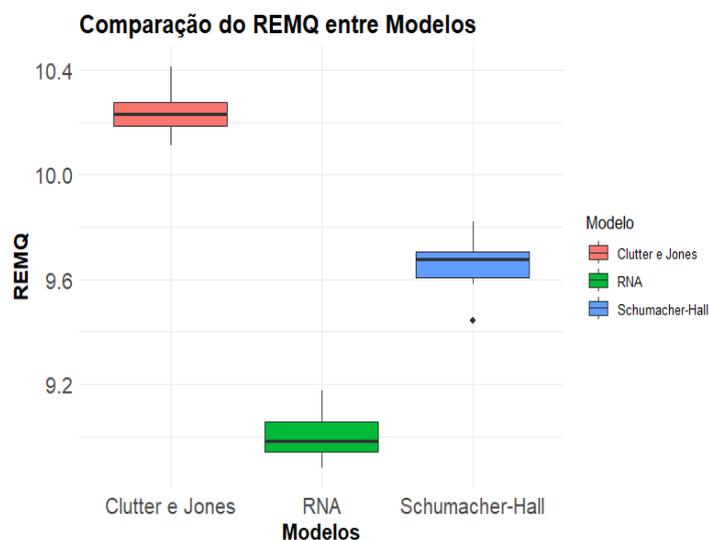


Figura 24- Distribuição do REMQ na modelagem volumétrica de *Eucalyptus urograndis*.

Observa-se que o modelo baseado em RNA apresentou os menores valores de REMQ, sugerindo maior precisão na estimativa do volume comercial de *Eucalyptus urograndis* quando comparado com modelos probabilísticos, utilizando 70% dos dados para treinamento e 20% para validação. O modelo Schumacher-Hall teve um desempenho intermediário, enquanto o modelo Clutter e Jones apresentou os maiores valores de REMQ, indicando menor precisão preditiva.

Vale ressaltar que a escolha do modelo mais adequado para aplicação prática não depende apenas da precisão estatística, mas também de fatores relacionados à conveniência, objetivos específicos e recursos disponíveis da empresa ou instituição que utilizará os resultados.

Se a empresa busca simples interpretação e facilidade de implementação do modelo, pode usar os modelos clássicos por ser modelos de fácil aplicação e manutenção, diferente dos modelos computacionais. Embora mais complexa, a RNA é capaz de capturar relações não lineares e interações mais complexas entre as variáveis, sendo uma escolha interessante para empresas que dispõem de recursos computacionais e equipes capacitadas para trabalhar com modelos mais avançados.

Ao comparar os resultados da tabela de parâmetros ajustados em relação ao REMQ, observa-se que a RNA apresentou os menores valores tanto no conjunto de treino (8,180) quanto na validação (9,116). Em seguida, o modelo Schumacher-Hall obteve valores intermediários (10,196 no treino e 9,620 na validação), enquanto o modelo Clutter e Jones apresentou os maiores REMQ (11,608 no treino e 10,237 na validação). Essa hierarquia está alinhada com a distribuição de resíduos observada na (Figura 25), evidenciando a maior ajuste da RNA em termos de precisão preditiva.

No que se refere ao Índice de Ajuste (IAc), a RNA também demonstrou o melhor desempenho, com os maiores valores na tabela (0,974 no treino e 0,969 na validação), indicando maior aderência entre os valores preditos e reais. Essa tendência é reforçada pelos gráficos de dispersão das predições, nos quais a RNA apresenta os pontos mais próximos à reta de identidade (1:1), evidenciando sua maior capacidade preditiva.

Além disso, os histogramas dos resíduos indicam que a RNA possui resíduos mais concentrados próximos de zero, refletindo menor variabilidade e confirmando os menores valores de Erro Médio Absoluto (EMA) e REMQ reportados na tabela. Essa distribuição reforça a robustez da RNA como o modelo mais eficiente para estimativa do volume comercial de *Eucalyptus urograndis*.

A (Figura 25) mostra a distribuição gráfica da predição, resíduos e histograma dos modelos ajustados.

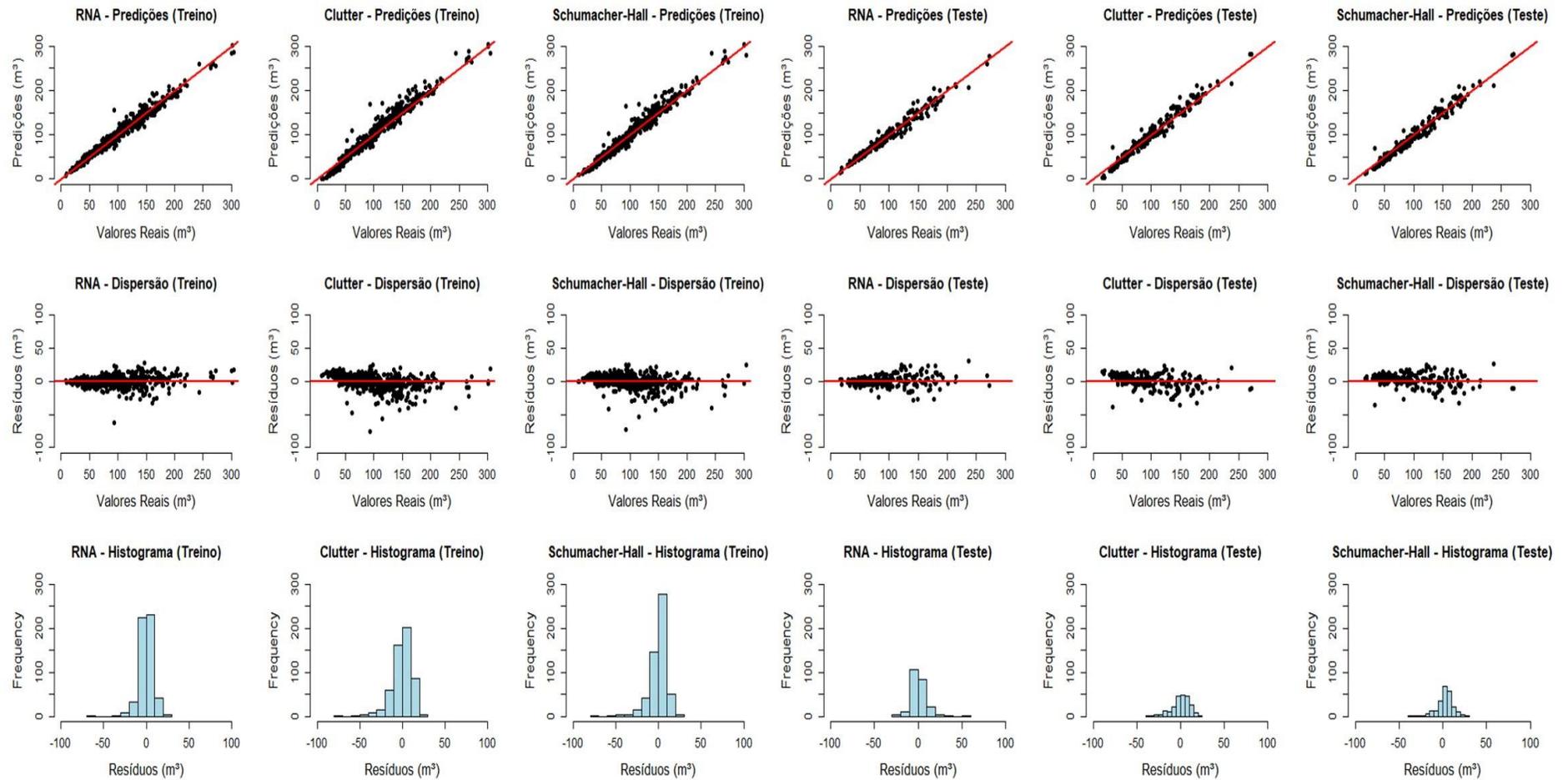


Figura 25- Gráficos de predições, resíduos e histogramas para os modelos Clutter e Schumacher-Hall nos conjuntos de treino e teste, utilizados na estimativa do volume comercial de *Eucalyptus urograndis* em Porto Grande, AP

Além da análise comparativa entre a RNA e os modelos probabilísticos, também foi investigada a importância das variáveis predictoras utilizadas na estimativa do volume comercial de *Eucalyptus urograndis*. A (Figura 26) apresenta a contribuição de cada variável para a predição do modelo, com base no aumento percentual do erro MSE (Mean Squared Error).

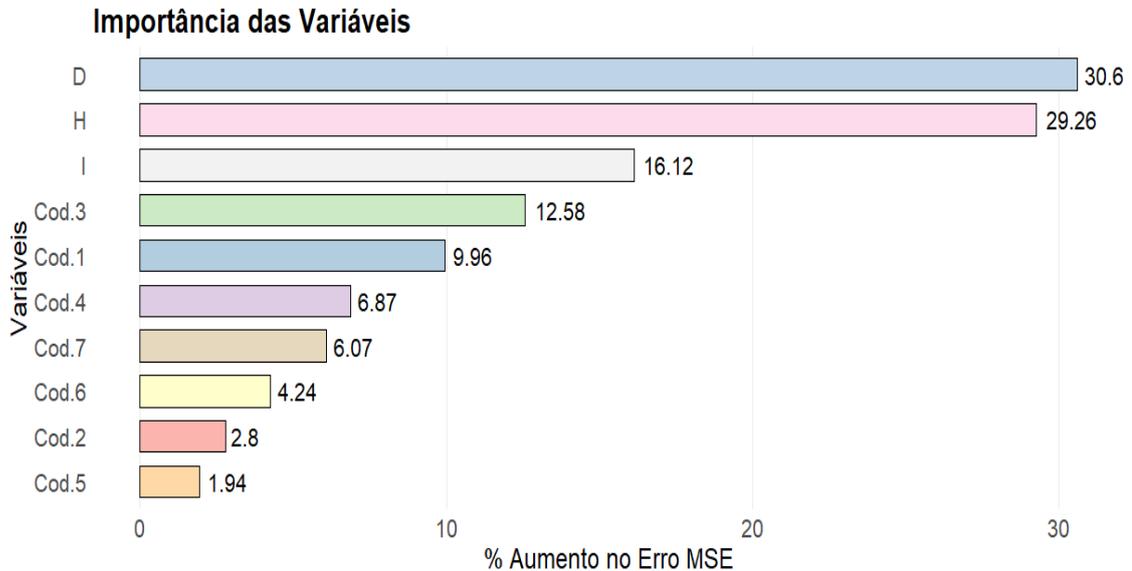


Figura 26 - Importância das variáveis na estimativa do volume comercial de *Eucalyptus urograndis* com base no aumento percentual do erro MSE.

Observa-se que as variáveis diâmetro (D) e altura total (H) foram as mais influentes, contribuindo com 30,6% e 29,26%, respectivamente. A variável idade (I) também apresentou impacto relevante, com 16,12% de contribuição. Entre os códigos dos clones avaliados, Cod.3 e Cod.1 destacaram-se com 12,58% e 9,96%, respectivamente, enquanto os demais híbridos apresentaram menor influência na predição.

Estes resultados reforçam a importância do diâmetro e da altura como principais variáveis explicativas na modelagem do volume comercial. A influência dos diferentes híbridos também são variáveis relevantes no crescimento e na produtividade, destacando a necessidade de abordagens específicas para diferentes materiais genéticos, bem como o valor de importância.

A (Tabela 18) apresenta a importância das variáveis na predição do volume de *Eucalyptus urograndis*, com base em duas métricas principais: %PAEMQ e IPN%PAEMQ, complementadas pela análise de sensibilidade.

Tabela 18 - Importância e sensibilidade das variáveis no modelo Random Forest com base no aumento percentual do erro MSE e pureza dos nós

Variáveis	%PAEMQ	IPN%PAEMQ	Sensibilidade
D	30,6356	801420,131	11,7138
H	29,2250	695981,978	7,7870
I	16,1183	266129,899	0,3101
Cod.3	12,5827	90426,136	0,0000
Cod.1	9,9610	56084,699	0,0000
Cod.4	6,8722	8802,579	0,0000
Cod.7	6,0672	5589,585	0,0000
Cod.6	4,2413	21687,981	0,0000
Cod.2	2,8030	1526,207	0,0000
Cod.5	1,9380	2588,665	0,0000

%PAEMQ (Aumento Percentual no Erro MSE), IPN (Aumento na Pureza dos Nós).

O modelo identifica D e H como variáveis primordiais para a predição do volume comercial, tanto pelo alto %PAEMQ quanto pela alta sensibilidade. Isto é consistente com o conhecimento florestal, pois o volume de uma árvore está diretamente relacionado ao seu diâmetro e altura. Embora a idade tenha alguma influência, seu impacto é muito menor que o de D e H, possivelmente porque o crescimento em volume se estabiliza em árvores mais velhas. Os híbridos contribuíram menos para a predição do volume, indicando que as diferenças genéticas entre os híbridos não são tão críticas quanto as variáveis dendrométricas.

Para aplicações práticas da importância da sensibilidade, é recomendável priorizar o uso de D e H no ajuste de modelos volumétricos, enquanto H e Híbrido devem ser reavaliadas quanto à sua utilidade e forma de inclusão no modelo. Strobl *et al.* (2007) discutem que medidas como %PAEMQ e IPN%PAEMQ podem ser enviesadas devido ao número de categorias ou à escala de medição das variáveis preditoras, levando à seleção incorreta de variáveis e distorcendo os resultados, especialmente quando variáveis irrelevantes possuem muitas categorias. Além disto, Strobl *et al.* (2008) destacam que a correlação entre preditores pode influenciar as medidas de importância das variáveis, resultando em preferências por variáveis correlacionadas nos modelos Random Forest.

De maneira geral, os resultados deste estudo demonstram que o modelo de aprendizado de máquina apresentou desempenho superior na estimativa do volume comercial de *Eucalyptus urograndis*, superando os modelos probabilísticos tradicionais, como Clutter e Jones e Schumacher-Hall, tanto em precisão preditiva quanto em robustez estatística. Com menores valores de REMQ e EMA e um índice de ajuste (IAc) mais elevado, a abordagem baseada em aprendizado de máquina foi mais preciso na modelagem volumétrica.

Além disso, ao analisar a contribuição das variáveis no modelo, as variáveis categóricas demonstraram impacto relevante, ainda que menos expressivo que as variáveis numéricas. Esse resultado reforça a importância de explorar diferentes tipos de variáveis na modelagem volumétrica, possibilitando ajustes mais refinados e projeções mais acuradas, alinhadas às particularidades genéticas e ambientais dos povoamentos analisados

7.5 Impacto de simulações com amostras aleatórias na estimativa do volume de *Eucalyptus urograndis* comparando modelos de aprendizado de máquina e probabilísticos

A (Tabela 19) apresenta as estatísticas descritivas das métricas de desempenho dos modelos testados. São apresentados os valores mínimo, máximo, quartis, mediana, média, desvio padrão e coeficiente de variação para cada métrica de avaliação (RMSE, MAE e IAaj) para a estimativa do volume de *Eucalyptus urograndis*.

Tabela 19 - Estatísticas descritivas das métricas de avaliação dos modelos testados de aprendizado de máquina e probabilísticos com amostras aleatórias na estimativa do volume de comercial de *Eucalyptus urograndis*

Modelos	Simulações		REMQ	EMA	IAaj
RNA – 100%	Mín	1,00	8,123	5,736	0,972
	1st Qu,	5,75	8,202	5,768	0,973
	Mediana	10,50	8,264	5,808	0,973
	Média	10,50	8,275	5,820	0,973
	3rd Qu,	15,25	8,516	5,850	0,974
	Máx	20,00	8,516	5,952	0,974
RNA -70/30	Mín	1,00	7,694	5,497	0,951
	1st Qu,	5,75	8,781	6,163	0,962
	Mediana	10,50	8,986	6,405	0,967
	Média	10,50	9,170	6,423	0,966
	3rd Qu,	15,25	9,652	6,797	0,971
	Máx	20,00	10,325	7,211	0,978
Schumacher-Hall -100%	Mín	1,00	8,831	6,074	0,968
	1st Qu,	5,75	8,831	6,074	0,969
	Mediana	10,50	8,842	6,078	0,970
	Média	10,50	8,881	6,111	0,969
	3rd Qu,	15,25	8,922	6,145	0,970
	Máx	20,00	9,019	6,221	0,970
Schumacher-Hall -70/30	Mín	1,00	7,275	5,211	0,954
	1st Qu,	5,75	8,615	5,996	0,968
	Mediana	10,50	9,013	6,179	0,970
	Média	10,50	8,937	6,166	0,970
	3rd Qu,	15,25	9,422	6,414	0,973
	Máx	20,00	10,200	7,209	0,980

Os modelos treinados com 100% dos dados (RNA e Schumacher-Hall) apresentam menor variação nos resultados, indicando maior estabilidade nas previsões. Em contrapartida, o modelo RNA 70/30 demonstra maior dispersão nos erros (RMSE e MAE), sugerindo que sua performance pode ser mais sensível à amostragem dos dados de treino. O modelo Schumacher-Hall 70/30 apresenta maior consistência em comparação à RNA 70/30, o que sugere que o método Schumacher-Hall pode ser mais robusto quando há menos dados disponíveis para treinamento. Além disso, o IAaj foi elevado em todos os modelos e todas as abordagens testadas foram capazes de capturar bem a relação entre D, H e o volume estimado.

A (Tabela 20) apresentada mostra os parâmetros ajustados (b_0 , b_1 e b_2) para o modelo de Schumacher e Hall em 20 simulações distintas. Este modelo é amplamente utilizado na estimativa do volume comercial de árvores. As simulações foram realizadas para avaliar a estabilidade dos ajustes do modelo ao longo de diferentes particionamentos dos dados. Cada simulação gerou valores ligeiramente diferentes para os parâmetros, refletindo as variações inerentes à amostragem e ao ajuste do modelo.

Tabela 20 - Parâmetros estimados do modelo Schumacher-Hall log-linear em 20 Simulações para a estimativa do volume de *Eucalyptus urograndis*

Autor	Modelo	Simulações	Coeficientes		
			b_0	b_1	b_2
Schumacher e Hall	$V_i = e^{[\beta_0 + \beta_1 \log(D_i) + \beta_2 \log(H_i)]} \varepsilon_i$	1	5,1043	0,0362	-0,1430
		2	4,9418	-0,0082	-0,1121
		3	5,0061	-0,0017	-0,1449
		4	4,6369	0,0257	-0,0419
		5	4,8335	-0,0151	-0,0643
		6	5,1198	-0,0695	-0,1104
		7	4,9412	-0,0155	-0,1072
		8	4,7844	-0,0277	-0,0326
		9	4,7779	-0,0292	-0,0341
		10	5,0377	-0,0359	-0,1178
		11	4,5240	0,0717	-0,0484
		12	4,7443	0,0604	-0,1201
		13	4,9250	-0,0708	-0,0335
		14	4,8034	0,0127	-0,0886
		15	4,7730	-0,0116	-0,0434
		16	4,9787	-0,0148	-0,1216
		17	4,7747	0,0322	-0,0953
		18	5,0575	-0,0706	-0,0884
		19	4,9633	-0,0180	-0,1091
		20	4,9949	-0,0553	-0,0788

b_0 , b_1 e b_2 são os parâmetros

Os valores ajustados desses parâmetros em cada simulação possibilitaram avaliar a consistência e a robustez do modelo ao ser aplicado a diferentes conjuntos de dados.

A (Tabela 21) apresenta os resultados detalhados das 20 simulações realizadas para cada modelo e configuração de treinamento. São exibidas as métricas RMSE, MAE e IAaj para as fases de treino e validação dos modelos RNA e Schumacher-Hall.

Tabela 21 - Estatísticas das 20 simulações independentes para os modelos RNA e Schumacher-Hall nas Configurações de Treino (70% e 100%) e Validação (30%) para a estimativa do volume de *Eucalyptus urograndis*

Modelos	Simulações	Treino (70%)			Validação (30%)			Desvio Padrão Volume (m³)	Media do Volume (m³)	Intervalo de Confiança 95%	Treino (100%)		
		REMQ	EMA	IAaj	REMQ	EMA	IAaj				REMQ	EMA	IAaj
RNA	1	8,414	5,877	0,972	8,917	6,086	0,970	51,885	97,868	6,856	8,281	5,813	0,973
	2	8,545	6,062	0,972	8,434	5,783	0,972	50,679	97,230	6,697	8,389	5,901	0,973
	3	8,418	5,877	0,973	8,788	6,523	0,970	51,232	102,344	6,770	8,294	5,793	0,973
	4	8,210	5,688	0,975	8,760	6,390	0,968	54,814	104,408	7,243	8,286	5,843	0,973
	5	8,119	5,829	0,973	8,940	6,331	0,971	52,391	103,639	6,923	8,249	5,824	0,974
	6	8,578	6,036	0,971	7,694	5,563	0,978	46,408	95,936	6,133	8,271	5,803	0,973
	7	8,041	5,612	0,976	8,879	6,420	0,966	52,434	100,025	6,929	8,220	5,814	0,974
	8	8,199	5,873	0,974	8,716	6,026	0,969	47,449	98,982	6,270	8,263	5,802	0,973
	9	7,842	5,589	0,977	9,707	6,771	0,961	45,894	94,174	6,065	8,213	5,819	0,974
	10	8,384	5,818	0,972	9,027	6,188	0,969	51,136	97,241	6,757	8,280	5,841	0,973
	11	8,178	5,678	0,976	8,945	6,268	0,961	47,985	96,631	6,341	8,937	6,212	0,969
	12	8,266	5,737	0,972	10,171	6,699	0,965	52,354	103,386	6,918	8,238	5,802	0,974
	13	7,950	5,500	0,975	9,531	6,886	0,966	50,597	94,943	6,686	8,135	5,761	0,974
	14	8,537	6,079	0,971	8,058	5,497	0,977	52,152	104,831	6,892	8,242	5,798	0,974
	15	7,586	5,381	0,978	10,325	7,211	0,958	50,647	101,554	6,693	8,234	5,795	0,974
	16	7,786	5,429	0,977	9,603	6,674	0,962	50,027	98,987	6,611	8,177	5,764	0,974
	17	7,936	5,635	0,978	9,892	6,876	0,951	50,609	95,089	6,688	8,208	5,796	0,974
	18	7,918	5,672	0,977	9,633	6,929	0,959	50,827	99,105	6,716	8,205	5,752	0,974
	19	7,654	5,492	0,976	10,246	7,042	0,962	53,175	101,125	7,027	8,305	5,793	0,973
	20	8,249	5,859	0,971	9,127	6,299	0,973	52,453	100,264	6,931	8,241	5,818	0,974
Schumacher - Hall	1	8,203	5,782	0,975	10,200	6,642	0,954	51,885	97,868	6,856	9,073	6,285	0,968
	2	8,629	5,923	0,970	9,280	6,579	0,969	54,489	97,868	7,200	9,053	6,357	0,968
	3	8,902	6,109	0,969	8,705	6,015	0,971	55,530	105,757	7,338	9,366	6,478	0,966
	4	8,919	6,129	0,966	8,647	6,003	0,976	50,985	104,459	6,737	8,959	6,257	0,969
	5	9,013	6,108	0,968	8,323	6,136	0,973	47,977	101,821	6,340	9,129	6,377	0,968
	6	9,108	6,275	0,968	8,195	5,759	0,973	53,457	99,349	7,064	9,104	6,312	0,968
	7	8,563	5,959	0,970	9,412	6,388	0,970	50,121	99,282	6,623	9,247	6,393	0,967
	8	8,619	5,957	0,972	9,418	6,325	0,964	50,970	105,586	6,735	9,079	6,288	0,968
	9	8,575	5,740	0,970	9,437	6,765	0,968	47,592	97,348	6,289	9,049	6,352	0,968
	10	8,658	5,959	0,970	9,245	6,358	0,969	50,156	98,455	6,628	9,073	6,279	0,968
	11	8,812	6,130	0,967	8,836	6,023	0,974	48,800	98,509	6,449	8,992	6,247	0,969
	12	8,471	5,923	0,972	9,615	6,346	0,964	43,689	97,720	5,773	9,125	6,372	0,968
	13	8,710	5,916	0,971	9,198	6,515	0,966	52,062	101,970	6,880	9,109	6,279	0,968
	14	9,414	6,478	0,964	7,275	5,211	0,980	49,651	101,043	6,561	8,901	6,209	0,969
	15	8,531	5,847	0,970	9,546	6,674	0,968	54,135	102,553	7,154	8,913	6,252	0,969
	16	8,682	6,034	0,970	9,203	6,077	0,969	46,988	97,922	6,209	9,138	6,319	0,968
	17	8,496	5,834	0,969	9,549	6,673	0,970	54,169	98,449	7,158	9,124	6,276	0,968
	18	8,684	6,062	0,971	9,564	6,052	0,964	52,334	104,177	6,916	9,115	6,348	0,968
	19	9,364	6,473	0,964	7,491	5,244	0,980	48,314	97,548	6,384	9,083	6,241	0,968
	20	9,023	6,151	0,966	8,382	5,945	0,976	48,058	99,008	6,351	9,241	6,392	0,967

As 20 simulações conduzidas para o modelo RNA apresentaram desempenho consistente, com RMSE variando entre 7,586 e 8,578 no conjunto de treino (70%) e entre 7,694 e 10,325 na validação (30%). O EMA oscilou entre 5,381 e 6,062 no treino e entre 5,497 e 7,211 na validação. O índice de ajustamento (IAaj) indicou elevada acurácia do modelo, com valores acima de 0,96 na maioria das simulações, atingindo até 0,978.

A análise do ajuste global (100% dos dados) revelou um RMSE médio de aproximadamente 8,25, EMA próximo de 5,81 e IAaj em torno de 0,974, indicando um bom ajuste do modelo. A variação no desvio padrão do volume comercial (m³) nas simulações de RNA ficou entre 45,894 e 54,814, enquanto a média do volume comercial oscilou entre 94,174 e 104,831 m³, com um intervalo de confiança de 95% variando de 6,065 a 7,243.

Os resultados das 20 simulações do modelo Schumacher-Hall apresentaram RMSE entre 8,203 e 9,414 no treino (70%) e entre 7,275 e 10,200 na validação (30%). O EMA variou entre 5,740 e 6,478 no treino e entre 5,211 e 7,042 na validação. O IAaj manteve-se entre 0,954 e 0,980, evidenciando a eficiência do modelo na estimativa volumétrica.

No ajuste global, o RMSE médio foi de aproximadamente 9,10, o EMA ficou em torno de 6,31 e o IAaj manteve-se próximo de 0,968. O desvio padrão do volume comercial variou entre 43,689 e 55,530 m³, enquanto a média do volume comercial situou-se entre 97,348 e 105,757 m³, com um intervalo de confiança de 95% variando de 5,773 a 7,338.

Os resultados indicam que o modelo RNA apresentou um desempenho ligeiramente superior ao modelo Schumacher-Hall, especialmente no que se refere ao RMSE e ao EMA. O IAaj do RNA foi consistentemente maior, chegando a 0,978, enquanto o Schumacher-Hall atingiu um máximo de 0,980 em uma simulação isolada. Isso sugere que a RNA pode fornecer previsões mais ajustadas ao volume comercial de clones de *Eucalyptus urograndis*.

Além disso, foi aplicada a validação cruzada para garantir a robustez e a generalização dos modelos. Na validação cruzada os resultados demonstram que os modelos treinados com 100% dos dados apresentaram menor variação nos erros, o que sugere maior estabilidade nas previsões. O modelo RNA 100% manteve valores consistentes de REMQ ($8,275 \pm 0,095$), EMA ($5,820 \pm 0,072$) e IAaj ($0,973 \pm 0,0007$), indicando um bom ajuste ao conjunto total de dados conforme a (Tabela 22) que exhibe as estatísticas das métricas RMSE, MAE e IAaj ao longo dos k-folds utilizados na validação cruzada.

Da mesma forma, o modelo Schumacher-Hall 100% também apresentou baixa variabilidade, com REMQ médio de 8,881, EMA de 6,111 e IAaj de 0,969. Por outro lado, os modelos treinados com 70% dos dados e validados com 30% apresentaram maior variabilidade nos erros. O RNA 70/30 obteve média de REMQ de 9,170, EMA de 6,423 e IAaj de 0,966,

apresentando maior dispersão em relação à sua versão com 100% dos dados. O Schumacher-Hall 70/30, por sua vez, demonstrou desempenho mais estável, com média de REMQ de 8,937, EMA de 6,166 e IAaj de 0,970, sugerindo que este modelo pode ser menos sensível à divisão dos dados de treino e validação.

Tabela 22 - Estatísticas da validação cruzada: RMSE, MAE e IAaj para os modelos testados de aprendizado de máquina e probabilísticos com amostras aleatórias na estimativa do volume de comercial de *Eucalyptus urograndis*

Modelos	REMQ			EMA			IAaj		
	Média	Desv.Padrão	IC_95	Média	Desv.Padrão	IC_95	Média	Desv.Padrão	IC_95
RNA 70/30	9,1696	0,7103	0,3113	6,4230	0,4775	0,2093	0,9664	0,0066	0,0029
Schumacher-Hall 70/30	8,9367	0,6647	0,2913	6,1660	0,4340	0,1902	0,9700	0,0053	0,0023
Schumacher-Hall 100%	8,2754	0,1070	5,8201	5,8201	0,0601	0,0263	0,9733	0,0007	0,0003
RNA 100%	8,8810	0,0252	6,1112	6,1112	0,0439	0,0132	0,9693	0,0004	0,0002

Na validação dos modelos (30% dos dados), os erros do RNA 70/30 apresentaram maior variação, com RMSE entre 7,694 e 10,325, enquanto o Schumacher-Hall 70/30 mostrou menor dispersão, variando entre 7,275 e 10,200. Este resultado indica que o modelo Schumacher-Hall é mais estável e robusto quando treinado com uma quantidade menor de dados.

Os resultados estatísticos com a validação cruzada indicaram que os modelos treinados com 100% dos dados apresentam menor variabilidade nos erros, o que sugere maior estabilidade nos resultados. Segundo Yates *et al.* (2023) a validação cruzada é importante para encontrar um equilíbrio entre desempenho e estabilidade dos modelos.

Após a análise da validação cruzada, foi realizado a análise de variância (ANOVA) para verificar se há diferenças estatisticamente significativas entre os modelos em relação ao (REMQ) conforme (Tabela 23).

Tabela 23 - Análise de variância (ANOVA) para comparação de modelos ajustados de RNA, e Schumacher-Hall na estimativa do volume comercial de *Eucalyptus urograndis*

FV	GL	SQ	QM	F	p>valor
Equações	3	8,691	2,8969	10,72	5,94e-06 ***
Resíduos	76	20,454	0,2703		
Total	79	29,236			

Em que: FV = Fontes de variação; GL = graus de liberdade; SQ = som ade quadrados; QM = quadrados médios; F= valor calculado do teste de F; p= nível de significância.

$$F_{(0.05;3,76GL)} = 2.724944$$

O teste ANOVA revelou uma diferença estatisticamente significativa entre os modelos analisados (F = 10,72; p-valor = 5,94e-06), indicando que pelo menos um dos modelos apresenta desempenho diferentes em relação aos demais. A partir da análise da ANOVA foi

feito o teste de Scott-Knott para identificar diferenças significativas entre os modelos avaliados (Tabela 24).

Observa-se que independente da escolha da significância ($\alpha = 5\%$ ou $\alpha = 1\%$) não houve auteração na formação dos grupos analisados.

Tabela 24 - Resultados do Teste de Scott-Knott para comparação do desempenho dos modelos testados de aprendizado de máquina e probabilísticos com amostras aleatórias na estimativa do volume de comercial de *Eucalyptus urograndis*

Modelo	REMQ	5%	1%
RNA 70/30	9,17	a	a
Schumacher-Hall 70/30	8,98	a	a
Schumacher-Hall 100%	8,83	a	a
RNA 100%	8,28	b	b

Os resultados indicam que o modelo RNA 100% apresentou os menores erros (REMQ = 8,28 e EMA = 5,83), diferenciando-se estatisticamente dos demais modelos. Já o modelo Schumacher-Hall 100% demonstrou um desempenho intermediário, enquanto o RNA 70/30 apresentou os maiores erros, sugerindo maior variabilidade nas previsões.

Em relação ao IAaj, o RNA 100% obteve o maior valor (0,97), destacando-se dos demais modelos, enquanto o RNA 70/30 apresentou os menores índices de ajuste. Esses resultados reforçam que o treinamento com 100% dos dados contribui para maior precisão e estabilidade nas estimativas.

A (Figura 27), apresenta a comparação gráfica das métricas de desempenho dos modelos testados, incluindo REMQ (A). EMA (B) e IAaj (C). Estes gráficos reforçam os resultados obtidos nas análises estatísticas e nos testes realizados, evidenciando as diferenças no desempenho dos modelos RNA e Schumacher-Hall sob diferentes configurações de treinamento.

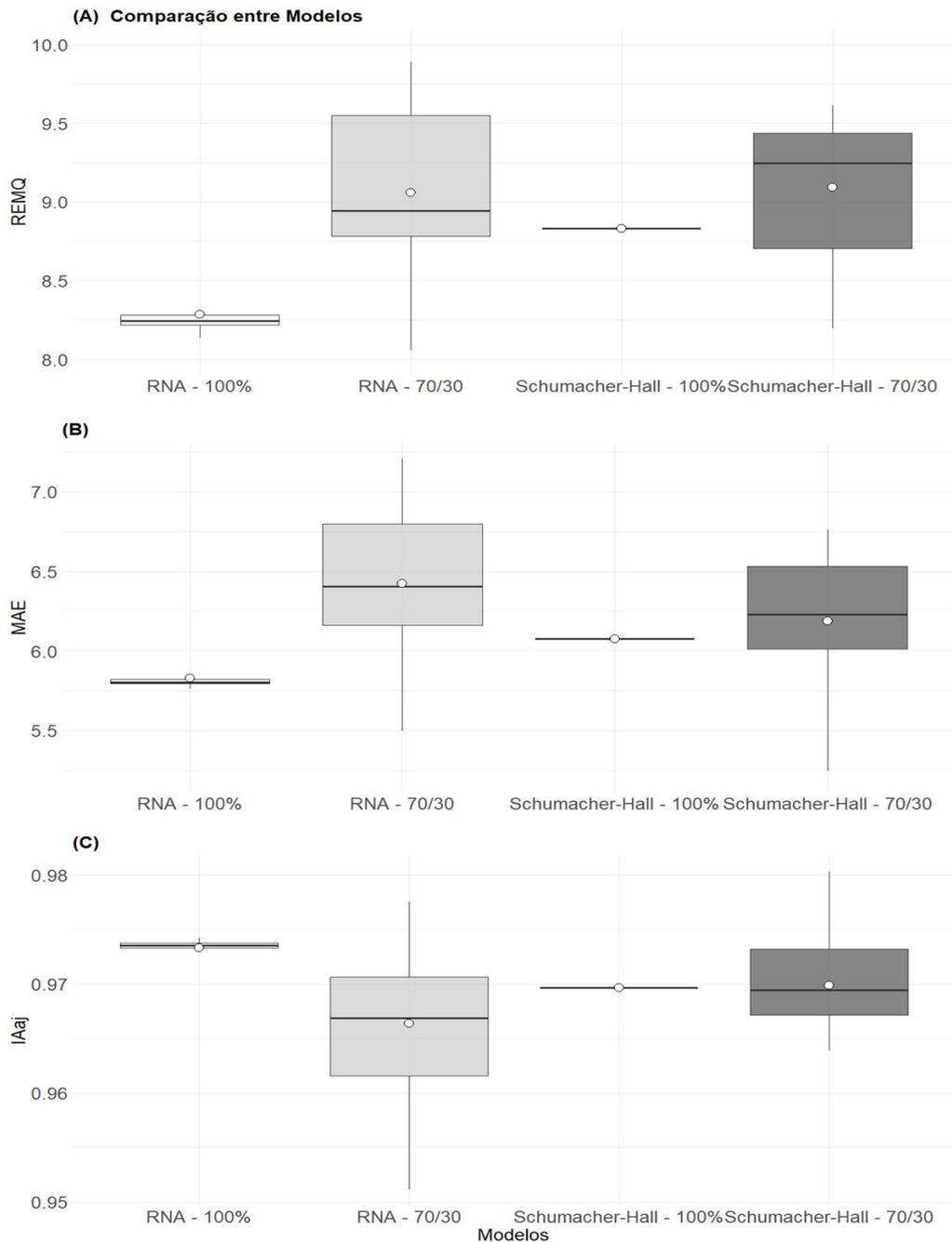


Figura 27 -Comparação dos modelos RNA e Schumacher-Hall para estimativa volumétrica com base em REMQ, MAE e IAaj.

Corroborando com os resultados o fato de os modelos treinados com 100% dos dados apresentarem menor variabilidade nos erros sugere que a amostragem reduzida para validação pode introduzir incertezas na avaliação do desempenho dos modelos. Ao retirar 30% da amostra para validação (aproximadamente 220 árvores), assume-se que essa porção representa adequadamente toda a população de 733 árvores, o que, estatisticamente, é uma suposição

questionável, conforme demonstrado pelo cálculo da probabilidade extremamente baixa de selecionar uma única amostra que represente fielmente a população.

Isto significa que, ao dividir os dados para validação, existe uma grande chance de que a amostra escolhida não seja completamente representativa da população, levando a variações maiores nos resultados. Por outro lado, ao treinar os modelos com 100% dos dados disponíveis, garantimos que todas as informações são utilizadas no ajuste, reduzindo a variabilidade e aumentando a estabilidade dos erros.

Portanto, estes resultados reforçam a importância de considerar o tamanho da amostra e a representatividade estatística na escolha da metodologia de validação, uma vez que a retirada de apenas uma amostra para validação pode levar a inferências menos robustas sobre o desempenho real dos modelos.

7.6 Otimização de modelos preditivos com algoritmo genético para estimativa do volume de *Eucalyptus urograndis*

A (Tabela 25) apresenta os resultados do treinamento e validação de quatro modelos preditivos (RNA, MVS, Random Forest e Schumacher-Hall) aplicados na estimativa do volume comercial de árvores. Os modelos foram ajustados utilizando o Algoritmo Genético (AG), que otimizou os parâmetros de cada método para maximizar o desempenho preditivo.

Para cada modelo, são apresentados os parâmetros ajustados, assim como as métricas de desempenho: REMQ, EMA e IAc nos conjuntos de treino e validação.

Tabela 25 - Desempenho dos modelos preditivos (RNA, MVS, Random Forest e Schumacher-Hall) em treino e validação, com parâmetros ajustados pelo Algoritmo Genético (GA)

Nº	Modelos	Parâmetros			Treino			Validação		
					REMQ	EMA	IAc	REMQ	EMA	IAc
1	RNA	size = 5	decay = 0,01		8,310	5,884	0,972	10,280	6,864	0,965
2	SVM	cost = 10	gamma = 0,1		8,639	5,991	0,970	9,280	6,317	0,971
3	Random Forest	ntree = 200	mtry = 2		4,668	3,239	0,991	9,997	6,238	0,970
4	Schumacher-Hall	b ₀ = 2,6284	b ₁ = 1,5914	b ₂ = 1,1386	8,781	6,113	0,969	0,969	6,257	0,973

Em que: (size): tamanho; (decay): decaimento; (cost): custo; (gamma): gama (parâmetro que define a forma do kernel nas Máquinas de Vetores de Suporte); (ntree): número de árvores; (mtry): número de variáveis e b₀, b₁ e b₃ os coeficientes usados para o modelo de Schumacher-Hall.

Observa-se que a RNA, ajustada com cinco neurônios na camada oculta e uma taxa de decaimento de 0,01, apresentou um desempenho consistente e robusto. O modelo alcançou um IAc superior a 0,96 em ambos os conjuntos, de treino e validação, demonstrando elevada capacidade preditiva e eficiência na modelagem do volume comercial.

A MVS (Máquina de Vetores de Suporte), configurada com um custo de 10 e um parâmetro gamma de 0,1, apresentou resultados satisfatórios, alcançando um IAc de 0,971 no conjunto de validação, indicando boa consistências nos resultados.

Random Forest quando ajustado com 200 árvores e dois preditores por nó, foi o modelo mais eficiente no treino, com menor REMQ (4,668) e maior IAc (0,991).

O modelo de Schumacher-Hall apresentou desempenho consistente, com IAc de 0,973 no conjunto de validação. Estes resultados destacam a eficácia do AG na calibração dos parâmetros e reforçam a flexibilidade dos métodos avaliados na modelagem volumétrica. A (Tabela 26) apresenta a comparação geral dos modelos ajustados utilizando (AG).

Tabela 26 - Comparação dos modelos RNA, SVM, Random Forest e Schumacher-Hall na estimativa do volume comercial de *Eucalyptus urograndis* em Porto Grande-AP, com parâmetros ajustados e métricas de desempenho

Nº	Modelos	Parametros		Estatística			
				REMQ	EMA	IAc	
1	RNA	size = 5	decay = 0,001	11,978	7,091	0,952	
2	RNA	size = 10	decay = 0,001	9,736	6,529	0,968	
3	RNA	size = 15	decay = 0,01	13,229	7,871	0,941	
4	RNA	size = 5	decay = 0,01	10,309	6,782	0,964	
5	RNA	size = 10	decay = 0,01	12,316	7,398	0,949	
6	RNA	size = 15	decay = 0,05	12,648	7,629	0,946	
7	RNA	size = 5	decay = 0,05	9,621	6,416	0,969	
8	RNA	size = 10	decay = 0,05	10,327	6,934	0,964	
9	RNA	size = 15	decay = 0,05	11,546	7,358	0,955	
10	RNA	size = 20	decay = 0,01	10,039	6,941	0,966	
1	SVM	cost = 1	gamma = 0,001	19,478	10,976	0,872	
2	SVM	cost = 10	gamma = 0,001	14,527	9,141	0,929	
3	SVM	cost = 100	gamma = 0,001	10,286	7,012	0,964	
4	SVM	cost = 1	gamma = 0,01	12,270	7,592	0,949	
5	SVM	cost = 10	gamma = 0,01	10,112	6,756	0,966	
6	SVM	cost = 100	gamma = 0,01	9,550	6,640	0,969	
7	SVM	cost = 1	gamma = 0,1	10,494	6,840	0,963	
8	SVM	cost = 10	gamma = 0,1	9,280	6,317	0,971	
9	SVM	cost = 100	gamma = 0,1	8,967	6,204	0,973	
10	SVM	cost = 1000	gamma = 0,01	9,294	6,417	0,971	
1	Random Forest	ntree = 100	mtry = 2	9,556	6,380	0,969	
2	Random Forest	ntree = 200	mtry = 2	9,389	6,174	0,970	
3	Random Forest	ntree = 300	mtry = 2	9,314	6,182	0,971	
4	Random Forest	ntree = 100	mtry = 3	9,460	6,275	0,970	
5	Random Forest	ntree = 200	mtry = 3	9,565	6,348	0,969	
6	Random Forest	ntree = 300	mtry = 3	9,404	6,215	0,970	
7	Random Forest	ntree = 100	mtry = 4	9,363	6,163	0,971	
8	Random Forest	ntree = 200	mtry = 4	9,313	6,123	0,971	
9	Random Forest	ntree = 300	mtry = 4	9,447	6,211	0,970	
10	Random Forest	ntree = 500	mtry = 2	9,398	6,191	0,970	
1	Schumacher-Hall	$b_1 = -2,621364$	$b_2 = 1,592847$	$b_3 = 1,135066$	8,965	6,266	0,973
2	Schumacher-Hall	$b_1 = 2,639286$	$b_2 = 1,606228$	$b_3 = 1,129811$	8,943	6,252	0,973
3	Schumacher-Hall	$b_1 = -2,608619$	$b_2 = 1,592914$	$b_3 = 1,130629$	8,984	6,281	0,973
4	Schumacher-Hall	$b_1 = 2,619934$	$b_2 = 1,617640$	$b_3 = 1,113666$	8,974	6,281	0,973
5	Schumacher-Hall	$b_1 = -2,626090$	$b_2 = 1,611530$	$b_3 = 1,120747$	8,964	6,270	0,973
6	Schumacher-Hall	$b_1 = -2,628457$	$b_2 = 1,605066$	$b_3 = 1,127098$	8,958	6,264	0,973
7	Schumacher-Hall	$b_1 = -2,627852$	$b_2 = 1,615954$	$b_3 = 1,117790$	8,962	6,271	0,973
8	Schumacher-Hall	$b_1 = -2,600066$	$b_2 = 1,571226$	$b_3 = 1,146424$	8,991	6,281	0,973
9	Schumacher-Hall	$b_1 = -2,632018$	$b_2 = 1,615158$	$b_3 = 1,119730$	8,956	6,265	0,973
10	Schumacher-Hall	$b_1 = -2,631047$	$b_2 = 1,592366$	$b_3 = 1,138763$	8,951	6,254	0,973

Redes Neurais Artificiais (RNA)

Para o modelo de Redes Neurais Artificiais (RNA), foram testadas 10 configurações variando o número de neurônios na camada oculta (size) e o valor de regularização (decay).

Configurações com mais neurônios e menor regularização tendem a capturar mais complexidade nos dados, enquanto configurações com mais regularização ajudam a evitar o sobreajuste. Os resultados mostraram que a configuração do modelo RNA de número 7 apresentou tamanho = 5 neurônios na camada oculta, com taxa de aprendizagem igual a 0.05. O valor da regularização apresentou o menor REMQ = 9,3207 e EMA = 6,4159 em comparação com os outros modelos RNA, além de um bom valor de IAc = 0,9690.

Isto indica que este modelo conseguiu prever com maior precisão, minimizando o erro absoluto e o quadrado médio dos erros, além de representar mais de 96% dos dados. Santana *et al.* (2023) avaliaram o desempenho de diferentes técnicas de aprendizado de máquina (ML) para prever o volume de madeira de eucaliptos, utilizando o D e H como variáveis de entrada. Quando utilizaram atributos combinados D e H para estimação do volume obtiveram entre todos os modelos testados $r = 0,95$.

Martins *et al.* (2016), ao investigarem configurações adequadas de RNA para a prognose da produção florestal em plantios de eucaliptos no nível de povoamento, constataram que as redes treinadas com Algoritmos Genéticos, configurado com parâmetros como populações variando de 50 a 1000 (incrementos de 50), taxas de mutação entre 0,1 e 0,5 (incrementos de 0,1), e taxas de cruzamento de 0,05 a 0,70 (incrementos de 0,05). A combinação destes parâmetros resultou em 114 configurações para o treinamento das RNA. Para os seus resultados o uso do algoritmo não foi satisfatório. Em alguns casos, as configurações não convergiram corretamente ou apresentaram resultados tendenciosos, comprometendo a precisão das previsões.

Os resultados deste estudo, juntamente com os obtidos por Binoti *et al.* (2014), que também aplicaram estes algoritmos para estimar volumes de povoamentos florestais, sugerem que estes métodos podem apresentar limitações em determinadas estimativas florestais, como na previsão de crescimento e na avaliação volumétrica de povoamentos. Isto indica que, para algumas aplicações específicas, estes algoritmos podem não ser a abordagem mais adequada.

Máquina de Vetores de Suporte (SVM)

Para o SVM, foram testadas diferentes combinações de custo (cost) e parâmetro de suavização (gamma). No caso da Máquina de Vetores de Suporte (SVM), os hiperparâmetros ajustados foram o custo que controla a penalidade para erros, e o gamma, que controla a flexibilidade do modelo. Observou-se que um valor intermediário para ambos os parâmetros proporcionou melhores resultados, como evidenciado pela equação número 9 do modelo SVM (custo = 100, gamma = 0,1), apresentou um REMQ de 8,9668 e um IAc de 0,97. As melhores

configurações de SVM, foram as equações 9, sugerindo que o custo mais alto e gamma intermediário proporcionam um bom ajuste ao modelo. Configurações com custo muito baixo como pode ser visto na equação 1 do modelo SVM (custo = 1), resultou em maior REMQ, sugerindo que o modelo estava subajustado para estes parâmetros.

O uso de Aprendizagem de Máquina se destaca por apresentar resultados precisos e acurados na predição de variáveis biométricas florestais (Binoti *et al.*, 2014; Martins *et al.*, 2016; Gao *et al.*, 2018; Pandit *et al.*, 2018; Cunha Neto *et al.*, 2019). Com base nos resultados apresentados, observa-se que os modelos de aprendizado de máquina, em particular o SVM com custo = 100 e gamma = 0.01, demonstraram melhor desempenho geral em termos de REMQ e IAc.

Random Forest

No modelo de Random Forest, o número de árvores (ntree) e o número de variáveis selecionadas para cada divisão (mtry) foram ajustados. A equação número 8 do modelo Random Forest (ntree = 200, mtry = 4) apresentou o melhor desempenho, com um REMQ de 9,3128, EMA = 6,123 e um IAc de 0,97. Esta configuração mostrou que um número moderado de árvores combinado com mais variáveis selecionadas por nó proporciona um ajuste mais preciso. Configurações com muitas árvores e menos variáveis, como nas equações 1 e 2, não apresentaram melhorias significativas no desempenho do modelo.

Estudos recentes, como o de Hu *et al.* (2023), também demonstram o potencial de otimizações baseadas em algoritmos genéticos para melhorar modelos florestais. Utilizando o algoritmo genético adaptativo de floresta aleatória (RF-AGA), assim como, aprimoraram a estimativa da biomassa acima do solo (AGB) em florestas por meio da seleção de características baseadas em mecanismos de dispersão e análise de sensibilidade. Embora o REMQ de 10,42 obtidos por eles seja superior aos resultados deste estudo, o uso de algoritmos genéticos se destaca como uma abordagem promissora para melhorar a precisão em estimativas florestais, especialmente em ambientes complexos.

Modelo de Schumacher-Hall

Os resultados indicaram que pequenas variações nos coeficientes b_1 , b_2 e b_3 tiveram impactos limitados nas métricas de desempenho do modelo Schumacher-Hall. No geral, o modelo apresentou uma REMQ variando de 8,95 a 8,98, e um IAc de 0,97 entre as 10 equações testadas, indicando um ajuste consistente e robusto. Das 10 equações a que melhores se ajustou foi a equação modelo 2. Schumacher-Hall, que apresentou REMQ de 8,9433 e EMA de 6,251.

Os coeficientes $b_1 = -2,6392$, $b_2 = 1,6062$ e $b_3 = 1,1298$, ajustam a relação entre o diâmetro (D), a altura total (H) e o volume estimado (\hat{V}). O parâmetro b_1 é um intercepto que ajusta a escala da relação, enquanto b_2 e b_3 controlam a influência relativa de D e H, respectivamente. Os valores de b_2 variaram entre 1,571 e 1,617, indicando uma influência relativamente estável do diâmetro no volume. O coeficiente b_3 , que ajusta a relação com a altura, variou entre 1,113 e 1,146, o que também sugere uma variação limitada da contribuição da altura no ajuste. Todas as 10 equações apresentaram um IAc acima de 0,96 o que indica que os 10 modelos avaliados capturaram mais de 96% da variabilidade nos dados, demonstrando que os coeficientes otimizados são adequados para a modelagem volumétrica nestes dados específicos.

Resultados semelhantes foram observados no estudo de Azevedo *et al.* (2020), que exploraram diferentes abordagens para modelar volumes múltiplos do fuste de árvores de eucalipto, incluindo o uso das equações de Schumacher e Hall. Neste estudo, os modelos também se mostraram eficazes, apresentando uma boa distribuição dos resíduos, com um viés leve apenas para volumes menores que 0,15 m³. As estatísticas recalculadas para ajuste (REMQ = 15,29% e EMA = 11,10%) e validação (REMQ = 15,17% e EMA = 11,70%) reforçam a eficiência da estratégia, com os resíduos distribuídos de forma adequada em torno de zero, validando a robustez do modelo. Silva *et al.* (2009) avaliaram o ajuste do modelo de Schumacher-Hall por diferentes algoritmos e a aplicação de RNA para estimar o volume de árvores de eucalipto.

Para selecionar as equações mais adequadas a este estudo, foi realizada uma análise de variância tendo como variável resposta a REMQ, cujos resultados estão apresentados na (Tabela 27)

Tabela 27 - Análise de variância (ANOVA) para comparação de modelos ajustados de RNA, SVM, Random Forest e Schumacher-Hall na estimativa do volume comercial de *Eucalyptus urograndis*

FV	GL	SQ	QM	F	p>valor
Equações	3	9,27	3,088	2,005	0,131
Resíduos	36	55,45	1,540		
Total	39	64,72			

Em que: FV = Fontes de variação; GL = graus de liberdade; SQ = som ade quadrados; QM = quadrados médios; F= valor calculado do teste de F; p= nível de significância.

$$F_{(0,05;3,36GL)} = 3,46$$

O valor do F calculado foi 2,005, enquanto o valor-p obtido foi 0,131, acima do nível de significância usual de 0,05. Assim, com base na comparação do F calculado com o F crítico

$(0.05; 3,36 \text{ GL}) = 3,46$, não há evidências suficientes para rejeitar a hipótese nula de que as médias dos modelos são iguais.

Esses resultados indicam que a maior parte da variação observada nos dados se deve aos resíduos, sugerindo que as diferenças entre os modelos, apesar de existirem, não são estatisticamente significativas ao nível de 5%.

O Teste de Scott-Knott foi aplicado para complementar a análise de variância (ANOVA), buscando identificar grupos homogêneos de modelos com base nas suas médias de REMQ. No presente estudo, o teste foi realizado para comparar os modelos RNA, SVM, Random Forest e Schumacher-Hall, a fim de identificar quais deles apresentaram desempenhos estatisticamente semelhantes no ajuste do volume comercial de *Eucalyptus urograndis*.

A (Tabela 28) apresenta o sumário do teste de Scott-Knott, cada modelo ajustado (equação) utilizado e suas respectivas médias de REMQ. Os resultados indicam que todos os modelos foram agrupados no mesmo cluster, evidenciando que, embora existam diferenças nas médias, estas não foram suficientemente significativas para formar clusters distintos.

Tabela 28 - Sumário do teste de Scott-Knott de comparação de média dos modelos RNA, SVM, Random Forest e Schumacher-Hall na estimativa do volume comercial de *Eucalyptus urograndis*

Modelos	Média REMQ	5%
SVM	10,28	a
Random Forest	9,280	a
RNA	9,421	a
Schumacher Hall	8,955	a

Conforme as análises estatísticas todos os modelos foram na (Figura 28), apresenta a dispersão entre os volumes observados e previstos para os modelos mais ajustados, incluindo RNA, SVM, Random Forest e Schumacher-Hall.

Em todos os casos, observa-se uma forte correlação entre os volumes observados e previstos, com os pontos próximos à linha de tendência, corroborando a precisão dos modelos analisados.

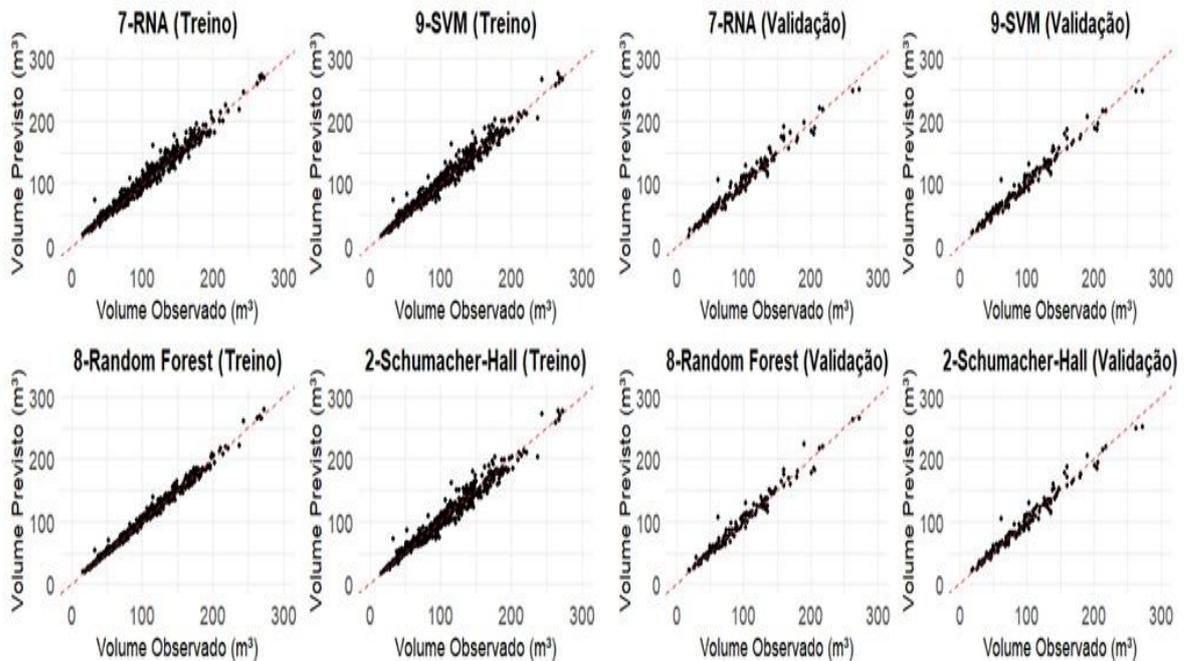


Figura 28- Comparação entre volumes observados e previstos pelos modelos RNA, SVM, Random Forest e Schumacher-Hall para estimativa de volume comercial de *Eucalyptus urograndis*.

O modelo 7 de RNA os pontos estão concentrados próximos à linha de identidade, indicando que o modelo RNA com a configuração $size = 5$ e $decay = 0.05$ foi capaz de prever os volumes com razoável precisão. No entanto, existem alguns desvios visíveis, particularmente para volumes maiores, indicando que o modelo pode ter alguma dificuldade em capturar volumes muito elevados. Na Máquina de Vetores de Suporte modelo 9 da SVM apresenta um ajuste muito bom, com a maioria dos pontos próximos à linha de identidade. Os desvios são mínimos, o que indica que este modelo foi eficiente em prever tanto volumes baixos quanto elevados, com um bom desempenho geral. O modelo 8 Random Forest, também apresenta uma boa proximidade dos pontos em relação à linha de identidade, com previsões bastante precisas para a maioria dos volumes.

O modelo 2 de Schumacher-Hall se destaca com pontos muito bem distribuídos em torno da linha de identidade. Isto indica que ele foi capaz de prever o volume de forma muito precisa, com poucos desvios visíveis. Este resultado corrobora com o fato de Schumacher - Hall ter o menor REMQ e maior IAc entre os modelos testados.

No gráfico de distribuição dos resíduos (em %) e Histograma dos resíduos na (Figura 29) todos os modelos apresentaram uma boa distribuição dos resíduos, com a maioria dos valores próximos à linha de referência (0%), indicando um ajuste adequado dos modelos.

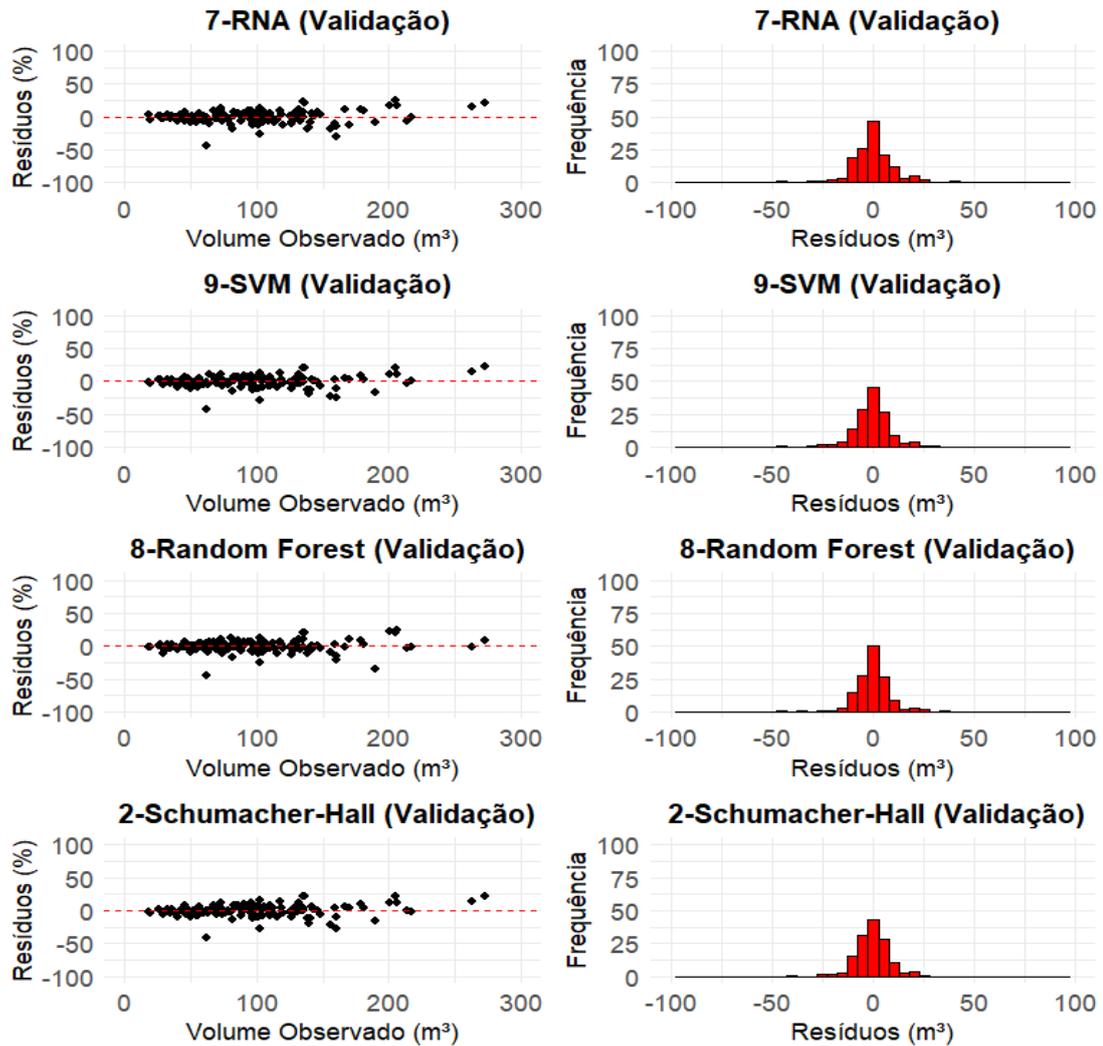


Figura 29 - Distribuição gráfica dos resíduos e histogramas dos modelos RNA, SVM, Random Forest e Schumacher-Hall na validação para estimativa do volume comercial de *Eucalyptus urograndis*.

A aplicação do AG nos modelos RNA, SVM, Random Forest e Schumacher-Hall possibilitou o ajuste otimizado de seus parâmetros, resultando em previsões mais estáveis e com menor variabilidade nos erros. Embora diferenças numéricas tenham sido observadas entre os modelos testados, os resultados estatísticos indicaram que não houve diferenças significativas ao nível de 5% entre os modelos após a otimização.

Para futuras pesquisas, os resultados deste estudo reforçam a necessidade de explorar novas abordagens híbridas, combinando modelos estatísticos clássicos com técnicas avançadas de aprendizado de máquina e otimização evolutiva.

8 CONCLUSÃO

A escolha da melhor equação na modelagem florestal muitas vezes se baseia em diferenças mínimas entre os modelos, tornando o processo subjetivo. Nesta pesquisa, os modelos testados demonstraram grande potencial para a modelagem volumétrica de *Eucalyptus urograndis*, oferecendo uma contribuição relevante a empresa AMCEL Amapá Celulose, fornecendo desempenhos robustos tanto para abordagens probabilísticas quanto para aprendizado de máquina.

O modelo Schumacher-Hall se destacou pela alta precisão e consistência nas estimativas baseadas em diâmetro e altura, enquanto o modelo Clutter-Jones foi eficiente na projeção do volume em função da idade, tornando-se uma excelente alternativa para estudos de crescimento e prognose. Ambos são amplamente utilizados na modelagem florestal e oferecem soluções confiáveis para planejamento produtivo.

Os modelos de aprendizado de máquina (RNA, SVM e Random Forest) também apresentaram bons ajustes preditivos, especialmente quando seus hiperparâmetros foram otimizados. No entanto, não superaram de forma absoluta os modelos probabilísticos, reforçando que a escolha do modelo deve considerar a complexidade da aplicação e os recursos disponíveis. Além disso, os modelos treinados com 100% dos dados mostraram métricas superiores aos treinados com 70/30, indicando maior estabilidade e confiabilidade nas previsões.

A otimização dos parâmetros, via Algoritmo Genético (AG), aprimorou o desempenho dos modelos, mas não houve diferenças estatísticas significativas entre RNA, SVM, Random Forest e Schumacher-Hall, conforme indicado pelo teste de Scott-Knott para REMQ.

Assim, todos os modelos avaliados são viáveis na estimativa do volume comercial, cabendo à empresa florestal optar entre simplicidade e confiabilidade (Schumacher-Hall e Clutter-Jones) ou técnicas mais flexíveis e adaptáveis (RNA, SVM e Random Forest), considerando os recursos operacionais, a necessidade de automação e a facilidade de implementação no manejo florestal.

9 REFERÊNCIAS BIBLIOGRÁFICAS

- ABREU, E.C.R. 2000. **Modelagem para prognose precoce do volume por classe diamétrica para *Eucalyptus grandis***. Lavras, 70 p. Dissertação (Mestrado). Universidade Federal de Lavras.
- ABREU, J. C. de. 2019. **Alternativas para estimar o volume de árvores individuais em formações florestais no estado de Minas Gerais**. Programa de Pós-graduação em Engenharia Florestal. Tese. Universidade Federal de Viçosa. 82f. 2019.
- ABREU, J. C. DE, SOARES, C. P. B., LEITE, H. G., BINOTI, D. H. B., & SILVA, G. F. Alternatives to estimate the volume of individual trees in forest formations in the state of Minas Gerais-Brazil. **Cerne**, v.26 n.3, p.393-402. 2020. <https://doi.org/10.1590/01047760202026032728>.
- AKAIKE, H. A new look at the statistical model identification. **Automatic Control, IEEE Transactions on**, v. 19, n. 6, p. 716-723, 1974.
- ALCÂNTARA, A. E. M. de. 2015. **Redes Neurais artificiais para a progênese do crescimento e da produção de povoamentos de eucaliptos em Minas Gerais**. Programa de Pós-graduação em Ciência Florestal. Tese. Universidade Federal de Viçosa. 43f.
- ANDRADE, V. C. L.; GAMA, R. da C.; TERRA, D. L. C. V. Revisão sobre modelos volumétricos empregados em povoamentos florestais brasileiros. **Advances in Forestry Science**. Cuiabá, v.6, n.1, p.561-566, 2019. DOI: <http://dx.doi.org/10.34062/afs.v6i1.7313>
- AZEVEDO, G. B.; TOMIAZZI, H. V.; AZEVEDO, G. T. DOS; TEODORO, L. P. R.; TEODORO, P. E.; DE SOUZA, M. T. P. et al. Multi-volume modeling of *Eucalyptus* trees using regression and artificiais neurais networks. ***PLoS ONE***, v. 15, n. 9, e0238703, 2020. <https://doi.org/10.1371/journal.pone.0238703>.
- BARROS, K. N. N. O. **Abordagem clássica e bayesiana em modelos simétricos transformados aplicados à estimativa de crescimento em altura de *Eucalyptus urophylla* no Polo Gesseiro do Araripe – PE**. 2010. Dissertação (Mestrado em Biometria) – Universidade Federal Rural de Pernambuco, Recife, 2010.
- BERTALANFFY, L. V. “A QUANTITATIVE THEORY OF ORGANIC GROWTH (INQUIRIES ON GROWTH LAWS. II).” *Human Biology* 10, no. 2 (1938): 181–213. <http://www.jstor.org/stable/41447359>.
- BERTALANFFY, L. V. “Quantitative Laws in Metabolism and Growth.” *The Quarterly Review of Biology* 32, no. 3 (1957): 217–31. <http://www.jstor.org/stable/2815257..>
- BHERING, L. L., CRUZ, C. D.; PEIXOTO, L.de A.; ROSADO, A. M.; LAVIOLA, B. G.; NASCIMENTO, M. Application of neurais networks to predict volume in eucalyptus. **Crop Breeding and Applied Biotechnology**, v.15 p.125-131, 2015. <http://dx.doi.org/10.1590/1984-70332015v15n3a23>

BINOTI, M. L. M. da S. **Redes neurais artificiais para progênese da produção de povoamentos não desbastados de eucalipto**. Dissertação. Programa de Pós-graduação em Ciências Florestais. Universidade Federal de Viçosa. 54f. 2010.

BINOTI, D. H. B., BINOTI, M. L. M. S., LEITE, H. G. Configuração de Redes Neurais Artificiais para estimação do volume de árvores. **Ciência da Madeira**, v.5 n.1, p.58-67. 2014. <https://doi.org/10.12953/21776830.v05n01a06>.

BOZDOGAN, H. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. **Psychometrika**, 52: p. 345–370, 1987.

BONETE, I. P., ARCE, J. E., FILHO, A. F., RETSLAFF, F. A. S., & LANSSANOVA, L. R. Artificial neural networks and mixed-effects modeling to describe the stem profile of *Pinus taeda* L. **Floresta**, 50(1), 1123-1132. 2019. <http://dx.doi.org/10.5380/ufv.v50i1.61764>.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento. **Florestas do Brasil em resumo**: 2019. Brasília DF, 207p. Disponível em: <https://www.florestal.gov.br/documentos/publicacoes/4261-florestas-do-brasil-em-resumo-digital/file> Acesso em: 24 de maio de 2021.

BREIMAN, L. Random Forests. **Machine Learning**, [s.l.], v. 45, n. 1, p.05-32, out. 2001. <http://dx.doi.org/10.1023/A:1010933404324>

BRITO, C. C. R.; SILVA, J. A. A.; FERREIRA, R. L. C.; SANTOS, E. de S.; FERRAZ, I. Modelos de crescimento resultantes da combinação e variações dos modelos de Chapman-Richards e Silva-Bailey em *Leucaena leucocephala* (Lam.) de Wit. **Ciência Florestal**, v. 17, n. 2, p. 175-185, 2007. DOI: <https://doi.org/10.5902/198050981949>

BRODY, S. **Bioenergetics and growth**. New York: Reinhold Publishing Corp, 1945. 1023 p.

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

CARNEIRO, M. G. **Redes complexas para classificação de dados via conformidade de padrão, caracterização de importância e otimização estrutural**. 2017. Tese de Doutorado. Universidade de São Paulo. <https://doi.org/10.11606/T.55.2017.tde-01022017-100223>

CARVALHO, S. P. C.; CALEGARIO, N.; SOLVA, F. F.; BORGES, L. A. C.; MENDONÇA, A. R.; LIMA, M. P. Modelos não lineares generalizados aplicados na predição de área basal e Eucalyptus clonal. **Cerne**, 17: p. 541-548, 2011. <https://doi.org/10.1590/S0104-77602011000400013>

CASTRO, R.V.O.; SOARES, C. P. B.; MARTINS, F. B.; LEITE, H. G. Crescimento e produtividade de plantios comerciais de eucalipto estimados por duas categorias de modelos. **Pesquisa Agropecuária Brasileira**, v.48, n.3, p.287-295, 2013. <https://doi.org/10.1590/S0100-204X2013000300007>

CHAPMAN, D.E. Statistical problems in dynamics of exploited fishing populations, **In: Berkeley Symposium on mathematical statistic and probability**, p. 153-158, Proceedings 1961.

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

CLUTTER, JEROME L., EARLE P. JONES. **Prediction of Growth after Thinning in Old-Field Slash Pine Plantations**. Asheville, N.C.: **Southeastern Forest Experiment Station**. United States Department of Agriculture, Southeastern Forest Experiment Station (Asheville, N.C.), and United States Forest Service. 1980.

CLUTTER, J. L.; FORTSON, J. C.; PIENAAR, L. V.; BRISTER, G. H.; BAILEY, R. L. (1983). *Timber Management: A Quantitative Approach*. Wiley. p.333.

CORDEIRO, M. A.; ACER, J. E.; RETSALFF, G. BONETE, I. P.; SILVA, A. V. dos S.; ABREU, J. C. de.; BINOTI, D. H. B. Estimativas volumétricas em povoamentos de eucalipto utilizando máquinas de vetores de suporte e redes neurais artificiais. **Madera y Bosques** v. 28, n. 1, Primavera, 2022. <https://doi.org/10.21829/myb.2022.2812252>.

CORDEIRO, M. A.; PEREIRA, N. N. J.; BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G. Estimativa do volume de Acácia mangium utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. **Pesquisa florestal brasileira**, Colombo, v. 35, n. 83, p. 255-261, jul./set. 2015. DOI: <https://doi.org/10.4336/2015.pfb.35.83.596>

COSENZA, D. N.; SOARES, A. A. V.; DE ALCÂNTARA, A. E. M.; DA SILVA, A. A. L.; REDE, R.; SOARES, V. P.; LEITE, H. G. Classificação de locais para povoamentos de eucalipto utilizando rede neural artificial baseada em características ambientais e de manejo. **Cerne**, v.23, n.3, p.310-320, 2017. url = <https://www.redalyc.org/articulo.oa?id=74452671004>.

COSTA FILHO, S. V. S. et al. Configuração de algoritmos de aprendizado de máquina na modelagem florestal: um estudo de caso na modelagem da relação hipsométrica. **Ciência Florestal**. Santa Maria, v. 29, n. 4, p. 1501-1515, 2019. <https://doi.org/10.5902/1980509828392>.

CUNHA NETO, E. M. et al. Estimativa do afilamento do fuste de indivíduos de eucalipto por meio de técnicas de inteligência artificial. ***BIOFIX Scientific Journal***, v.4, n.2, p.166-171, 2019. <http://dx.doi.org/10.5380/biofix.v4i2.65831>.

CUNHA NETO, E. M. da.; MOURA, M. M.; ARAUJO, E. C. G.; SANTANA, G. M.; DALLA CORTE, A. P.; SANQUETA, C. R. Aprendizado de máquina e regressão linear na estimativa do volume de Eucalyptus na Amazônia Oriental. **BIOFIX Scientific Journal** v. 6 n. 1 p. 08-14 2021. <http://dx.doi.org/10.5380/biofix.v6i1.75531>.

FALCÃO, A.O.; BORGES, J.G. Designing an evolution program for solving integer forest management scheduling models: an application in Portugal. **Forest science**, v.47, n.2, p.158-168, 2001. DOI:[10.1093/forestscience/47.2.158](https://doi.org/10.1093/forestscience/47.2.158)

FIORENTIN, L. D.; MIRANDA, R. O. V.; FIGUEIREDO FILHO, A.; TÊO, S. J. DA COSTA, R. H. Modelos não lineares aplicados para classificação de sítios florestais de *Pinus Elliottii* E *pinus taeda*. **Rev. Bras. Biom.**, Lavras, v.35, n.3, p.474-496, 2017.

FRITZSONS, E.; PARRON, L. M. Plantações florestais comerciais e a água. cap. 3 In: OLIVEIRA, Y. M. M. de; OLIVEIRA, E. B. de (Ed.). *Plantações florestais: geração de benefícios com baixo impacto ambiental*. Brasília, DF: Embrapa, 2017.

FRÜHAUF A. C.; SILVA E. M.; GRANATO-SOUZA D.; SILVA E. M.; MUNIZ J. A.; FERNANDES T. J. Description of height growth of hybrid eucalyptus clones in semi-arid region using non-linear models. **Brazilian Journal of Biometrics**. Lavras, v.40, n.2, p.138-151, 2022. DOI: <https://doi.org/10.28951/bjb.v40i2.543>

FURNIVAL, G. M. An index for comparing equation used in construction of volume table, *Forest Science*, v. 7, n. 4, p. 337-341, 1961.

GAO, Y.; LU, D.; LI, G.; WANG, G.; CHEN, Q.; LIU, L.; LI, D. Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. **Remote Sens**. 2018, 10, 627. <https://doi.org/10.3390/rs10040627>.

GARCIA NETO, P. J.; MARTÍNEZ, J; ARAÚJO, M; ORDÓÑEZ, C. Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*. **Applied Mathematical Modelling**. v.36. p.6137–6145.2012. DOI:[10.1016/j.apm.2012.02.016](https://doi.org/10.1016/j.apm.2012.02.016)

GENUER, R. POGGI, J. TULEAU-MALOT, C. VSURF: An R Package for Variable Selection Using Random Forests. **The R Journal**. v. 7, n2, 2015. DOI:[10.32614/RJ-2015-018](https://doi.org/10.32614/RJ-2015-018)

GENUER, R.; POGGI, J.; TULEAU-MALOT, C Variable selection using Random Forests. **Pattern Recognition Letters**. v.31. p.2225-2236. 2010. DOI:[10.1016/j.patrec.2010.03.014](https://doi.org/10.1016/j.patrec.2010.03.014)

GOMIDE LR, ARCE JE, SILVA ACL (2009) Uso do algoritmo genético no planejamento florestal considerando seus operadores de seleção. **Cerne**, 15(4): 460-467. <http://www.redalyc.org/articulo.oa?id=74413024011>

GONÇALVES, J. L. M.; ALVARES, C. A.; HIGA, A. R.; SILVA, L. D.; ALFENAS, A. C.; STAHL, J.; et al. Integrating genetic and silvicultural strategies to minimize abiotic and biotic constraints in Brazilian eucalypt plantations. **Forest Ecology and Management**, v. 301, p. 6-27, jan. 2013. <https://doi.org/10.1016/j.foreco.2012.12.030>

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. DOI:[10.1016/j.compag.2015.07.004](https://doi.org/10.1016/j.compag.2015.07.004)

GOUVEIA, J. F.; SILVA, J. A. A.; FERREIRA, R. L. C.; GADELHA, F. H. L.; LIMA FILHO, L. M. A. Modelos volumétricos mistos em clones de *Eucalyptus* no Polo Gesseiro do Araripe, Pernambuco. **Floresta**, v. 45, p. 587-598, 2015. DOI: <http://dx.doi.org/10.5380/ufpa.v45i3.36844>

HAYKIN, S. **Redes Neurais: princípios e prática**. Porto Alegre, 2001. 2ª Ed. Editora: Bookman. 900p.

HODSON, T. O.: Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not, *Geosci. Model Dev.*, 15, 5481–5487, <https://doi.org/10.5194/gmd-15-5481-2022>.

HOLLAND, J. H. *Adaptation in Natural and Artificial Systems*. 2. ed. [S.l.]: The MIT Press, 1975.

HU, Y.; NIE, Y.; LIU, Z.; WU, G.; FAN, W. Improving the potential of coniferous forest aboveground biomass estimation by integrating C- and L-Band SAR data with feature selection and non-parametric model. **Remote Sens**. 2023, 15, 4194. <https://doi.org/10.3390/rs15174194>

HUSCH, B.; MILLER, C. I.; BEERS, T. W. **Forest Mensuration: Krieger Publishing Company**. Malabar. Florida, 1982. p402.

IBÁ – INDÚSTRIA BRASILEIRA DE ÁRVORES. **Relatório Anual 2022**. Disponível em: <https://twosides.org.br/wp-content/uploads/sites/15/2023/02/relatorio-anual-iba2022-compactado.pdf> > Acesso em: 20 de set. 2023.

IBÁ- INDÚSTRIA BRASILEIRA DE ÁRVORES **Relatório Anual 2023**. Disponível em: <https://www.iba.org/datafiles/publicacoes/relatorios/relatorio-anual-iba2023-r.pdf>. Acesso em: 07 jun. 2021.

KARIME, H.; KANI, I. M. Finding the worst imperfection pattern in shallow lattice domes using genetic algorithms. **Journal Of Building Engineering**, Netherlands, v. 23, p.107-113, 2019. <https://doi.org/10.1016/j.jobe.2019.01.018>

KARTHIK, C.; VALARMATHI, K.; RAJALAKSHMI, M. Support vector regression and model reference adaptive control for optimum control of nonlinear drying process. If you had to choose a kaolin supplier based on one quality, which would you choose? **TAPPI Journal** ·v. 6, p. 111-126, 2016. DOI:[10.32964/TJ15.2.111](https://doi.org/10.32964/TJ15.2.111)

KIVINEN, J. A.; SMOLA, J.; WILLIAMSON, R. C. Online learning with kernels. ***IEEE Transactions on Signal Processing***, v. 52, n. 8, p. 2165-2176, 2004. <https://doi.org/10.1109/TSP.2004.830991>.

KUHN, M. Building Predictive Models in R Using the caret Package. **Journal of Statistical Software**, v. 28, n. 5, p. 1-26, 2008. <https://doi.org/10.18637/jss.v028.i05>.

LEAL, A.F.; MIGUEL, E. P.; MATRICARDI, E. A. T.; PEREIRA, R.S. REDES NEURAIAS ARTIFICIAIS NA ESTIMATIVA DE VOLUME EM UM PLANTIO DE EUCALIPTO EM FUNÇÃO DE FOTOGRAFIAS HEMISFÉRICAS E NÚMERO DE ÁRVORES. **Revista Brasileira Bioma.**, São Paulo, v.33, n.2, p.233-249, 2015.

LEAL, F. A.; LEAL, G. da S. A.; SILVA, T. C. Redes neurais artificiais e modelos alométricos aplicados para estimativa de volume e altura em *Eucalyptus urophylla* S. T. Blake. **Advances in Forestry. Science** Cuiabá, v. 7, n. 3, p. 1181-1188, 2020 DOI: <http://dx.doi.org/10.34062/afs.v7i3.10720>

LEITE, H. G.; SILVA, M. L. M. da; BINOTI, D. H. B.; FARDIN, L.; TAKIZAWA, F. H. Estimation of inside-bark diameter and heartwood diameter for *Tectona grandis* Linn. trees using artificial neurais networks. **European Journal of Forest Research**, v. 130, n. 2, p. 263–269, 2011. <https://doi.org/10.1007/s10342-010-0427-7>

LIAW, A.; WIENER, M. Classification and regression by randomForest. **R News**, v. 2, n. 3, p. 18-22, 2002.

LIMA FILHO, L. M. A.; SILVA, J. A. A.; CORDEIRO, G. M.; FERREIRA, R. L. C. Modelagem do crescimento de clones de *Eucalyptus* usando o modelo de Chapman-Richards com diferentes distribuições simétricas dos erros. **Ciência Florestal**, v. 22, p. 777-785, 2012. <https://doi.org/10.5902/198050987558>.

MAESTRI, R.; SANQUETTA, C. R.; SCOLFORO, J. R.; do Amaral MACHADO, S.; Dalla CORTE, A. P. Modelagem do crescimento florestal considerando variáveis do ambiente: revisão **Scientia Agraria**, vol. 14, n. 3, pp. 103-110, 2013. DOI: [10.5380/rsa.v14i3.40908](https://doi.org/10.5380/rsa.v14i3.40908)

MARTINS, E. dos R.; BINOTI, M. L. M. da S.; LEITE, H.G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de redes neurais artificiais para prognose da produção de povoamentos clonais de eucalipto. **Agrária - Revista Brasileira de Ciências Agrárias**. v.10, n.4, p.532-537, 2015. DOI: <https://doi.org/10.5039/agraria.v11i2a5373>

MARTINS, E. DOS R., BINOTI, M. L. M. DA S., LEITE, H. G., BINOTI, D. H. B., & DUTRA, G. C. Configuração de redes neurais artificiais para estimação do afilamento do fuste de árvores de eucalipto. **Revista Brasileira de Ciências Agrárias**, 11(1), 33-38, 2016. DOI: <https://doi.org/10.5039/agraria.v11i1a5354>

MARTINS, M. T., COSTA, E. A., MARANGON, G. P. Estimativas volumétricas em plantios de eucalipto para duas mesorregiões do Rio Grande do Sul. **BIOFIX Scientific Journal** v. 6 n. 1 p. 28-36 2021. <http://dx.doi.org/10.5380/biofix.v6i1.75708>.

MELO, R. R.; MIGUEL, E. P. Use of artificiais neurais networks in predicting particleboard quality parameters. **Revista Árvore**, Viçosa-MG, v.40, n.5, p. 949-958, 2016. <https://doi.org/10.1590/0100-67622016000500019>.

MENDONÇA, N. de P. CARVALHO, M. C.; GOMIDE, L. R. FILHO, A. C. F.; FERREIRA, M. A. Previsão de diâmetros ao longo do fuste de eucalipto via redes neurais artificiais. **ENCICLOPÉDIA BIOSFERA**, Centro Científico Conhecer - Goiânia, v.11 n.22; p. 2015. DOI: http://dx.doi.org/10.18677/Enciclopedia_Biosfera_2015_212.

MEYER, D.; DIMITRIADOU, E.; HORNIK, K.; WEINGESSEL, A.; LEISCH, F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien; 2019. **R package version** 1.7-2. Available from: <https://CRAN.R-project.org/package=e1071>.

MIGUEL, R. P.; LEAL, F. A.; ONO, H. A.; LEAL, U. A. S. Modelagem na predição do volume individual em plantio de *Eucalyptus Urograndis*. **Rev. Bras. Biom.**, São Paulo, v.32, n.4, p.584-598, 2014.

MITCHELL, M. An introduction to genetic algorithms. Cambridge: Mit Press. 1997. 207 p.

MOJENA, P. A. FERNANDO, S. C. SILVA, D. S. JESUS . F. W. Curvas de índices de sítio para povoamentos de *Eucalyptus urograndis* Site index curves for *Eucalyptus urograndis* stands. **Scientific Electronic Achilles**. v. 12 p. 2, 2019. <http://www.seasinop.com.br/revista/index.php?journal=SEA&pa>

MONTEIRO T. C.; JUNIOR, C. A. A.; SANTOS, J. H.; SILVA, T. C.; MASCIMENTO, T. M.; JUNIOR, J. L. F. C.; MATOS, J. L. M.; KLITZKE, R. J.; ROCHA, M. P. Artificial intelligence to growth stresses predicting in eucalyptus clones using dendrometric variables and wood density. **Maderas. Ciencia y tecnología** v.25, n.30, 1-12. 2023. <http://dx.doi.org/10.4067/s0718-221x2023000100430>.

MOREIRA, J. M. M. A. P.; SIMIONI, F. J.; OLIVEIRA, E. B. Importância e desempenho das florestas plantadas no contexto do agronegócio brasileiro. **Floresta**, Curitiba, PR, v. 47, n. 1, p. 85 - 94, jan. / mar. 2017. DOI: <http://dx.doi.org/10.5380/ufv.v47i1.47687>

NASCIMENTO, A. F.; DIAS, A. N.; FILHO, A. F.; ACER, J. E.; MIRANDA, G. DE M. Uso da Meta-Heurística otimização por exame de partículas no planejamento Florestal. **Scientia Florestales**, Piracicaba, v. 40, n. 96, p. 557-565. 2012

PANDIT, S.; TSUYUKI, S.; DUBE, T. Estimating Above-Ground Biomass in Sub-Tropical Buffer Zone Community Forests, Nepal, Using Sentinel 2 Data. **Remote Sens**. 2018, 10, 601. <https://doi.org/10.3390/rs10040601>.

PENIDO, T. M. A., LAFETÁ, B. O., NOGUEIRA, G. S., ALVES, P. H., GORGENS, E. B., & OLIVEIRA, M. L. R. (2020). Modelos de crescimento e produção para a estimativa volumétrica em povoamentos comerciais de eucalipto. **Scientia Forestalis**, 48(128), e3340. <https://doi.org/10.18671/scifor.v48n128.06>

PERCONTINI, A.; GOMES-SILVA1, F.S.; RAMOS, M.W.A.; VENANCIO, R.; CORDEIRO, G.M. A distribuição Gama Weibull Poisson aplicada a dados de sobrevivência. **Tendências em Matemática Aplicada e Computacional**, 15, N. 2 (2014), 165-176. <https://doi.org/10.5540/tema.2014.015.02.0165>

PEREIRA, J. C.; DIAS, P. A. S.; MERGULHÃO, R. C.; THIERSCH, C. R.; FARIA, L. C. Modelo de crescimento e produção de Clutter adicionado de uma variável latente para predição do volume em um plantio de *Eucalyptus urograndis* com variáveis correlacionadas espacialmente. **Scientia Forestalis**, v. 44, n. 110, p. 393-403, 2016. <http://dx.doi.org/10.18671/scifor.v44n110.12>

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

PRODAN, M. Forest biometrics. Oxford : Pergamon Press, 1968. 447p.

R DEVELOPMENT CORE TEAM. R: **A language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2014. Disponível em: <https://www.R-project.org/>. Acesso em: 03 jan. 2024.

RAHMAN, I.; MOHAMAD-SALEH, J. Hybrid bio-Inspired computational intelligence techniques for solving power system optimization problems: A comprehensive survey. **Applied Soft Computing**, **Netherlands**, v. 69, p.72-130, 2018.

REGAZZI, A. J. Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear. **Revista Ceres**, Viçosa, v.50, n.287, p.9-26, 2003.

REIS, L.P.; SOUZA, A.L.DE.; REIS, P.CM DOS MAZZEI L, BIONOTI, DHB, LEITE HG. Prognose da distribuição diamétrica na Amazônia utilizando redes neurais artificiais e autômatos celulares. **Floresta**, Curitiba, PR, 2018; 48(1): 93-102. <http://dx.doi.org/10.5380/rf.v48i1.527484>

RETSLAFF, FAS, FIGUEIREDO FILHO A, DIAS AN, BERNETT LG, FIGURA MA. Curvas de sítio e relações hipsométricas para *Eucalyptus grandis* na região dos Campos Gerais, Paraná. **Cerne**, v. 21, n. 2.p 219-225. 2015.

RIBEIRO, M. D. S. B. JORGE, L. A. B.; MISCHAN, M. M.; SANTOS, A. L. ; BALLARIN, A. W. Avaliação da produção de biomassa do fuste de um clone híbrido de eucalipto sob diferentes espaçamentos. **Ciência Florestal**, Santa Maria, v. 27, n. 1, p. 31-45, 2017. <https://doi.org/10.5902/1980509826445>.

RICHARDS, F. J. A flexible growth function of empirical use. **Journal of Experimental Botany**, Oxford, v. 10, p. 290-300, 1959.

RIPLEY, B. D. (1996). Pattern recognition and neural networks. Usado para treinar redes neurais artificiais com o pacote nnet. **Cambridge University Press**, 1996, ISBN 0-521-46086-7.

RODRIGUES, F.L. LEITE, H. G.; SANTOS, H. do N.; SOUZA, A. L.; SILVA, G. F. Metaheurística algoritmo genético para solução de problemas de planejamento florestal com restrições de integridade. **Revista Árvore**, v. 22, n. 2, p. 233–245, 2004. <https://doi.org/10.1590/S0100-67622004000200010>.

SALES, F. C. V.; SILVA, J. A. A.; FERREIRA, R. L. C.; GADELHA, F. H. L. Ajustes de modelos volumétricos para o clone *Eucalyptus grandis* x *E. urophylla* cultivados no Agreste de Pernambuco. **Floresta**, v. 45, p. 663-669, 2015. DOI: <http://dx.doi.org/10.5380/rf.v45i4.37594>

SANTANA, D. C.; SANTOS, R. G. D.; DA SILVA, P. H. N.; PISTORI, H.; TEODORO, L. P. R.; POERSCH, N. L.; DE AZEVEDO, G. B.; DE OLIVEIRA SOUSA AZEVEDO, G. T.; DA SILVA JUNIOR, C. A.; TEODORO, P. E. Machine Learning Methods for Woody Volume Prediction in *Eucalyptus*. **Sustainability**, v.15, p.10968, 2023. <https://doi.org/10.3390/su151410968>.

SANTOS, C. J. THIERSCH, C. R.; MOREIRA, M. F. B. Regressão não linear quantílica para classificação de sítio em povoamentos de *Eucalyptus* sp. *Scientia Forestalis*, Piracicaba, v. 46, n. 120, p. 551-561, dez. 2018. DOI:[10.18671/scifor.v46n120.04](https://doi.org/10.18671/scifor.v46n120.04)

SCHLAEGEL, R. (1981). Testing, Reporting, and Using Regression Models for Tree Biomass Estimation. In Proceedings of the First Forest Biomass Workshop, Forest Biomass Studies, General Technical Report.

SCHNEIDER, P.R.; SCHNEIDER, P.S.P.; SOUZA, C.A.M. **Análise de regressão aplicada à Engenharia Florestal**. 2a ed., Santa Maria, FACOS, 2009. 294p.

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

SCOLFORO, J. R. S. **Biometria florestal: modelos de crescimento e produção florestal**. Lavras: UFLA/FAEPE, 2006. 393 p. (Textos Acadêmicos).

SCOTT, A. J.; KNOTT, M. A cluster analysis method for grouping mean in the analysis of variance. **Biometrics**, Alexandria, v.30, p. 507-5042, 1974.

SILVA, J. A. A. **Dynamics of stand structure in fertilized slash pine plantations**. Athens: University of Georgia, 1986. 133 p. Thesis (Doctor of Philosophy) – University of Georgia, GA, USA.

SILVA, J. A. A.; BAILEY, R. L. Uso de transformações normalizadoras no ajuste de modelos volumétricos. **Revista Árvore**, Viçosa, MG, v. 15, n.12 p. 199-206, 1991.

SILVA, J. A. A; SILVA, I. P. **Estatística experimental aplicada à ciência florestal**. Universidade Federal Rural de Pernambuco 269f. 1995.

SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Ajuste do modelo de Schumacher e Hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucalipto. **Revista Árvore**, 33(6), 1133-1139, 2009. <https://doi.org/10.1590/S0100-67622009000600015>.

SILVA, J.A.A. Conceitos e princípios básicos de modelagem matemática em ciências florestais. **Anais da Academia Pernambucana de Ciência Agronômica**, Recife, vols. 11/12, p.195-215, 2015.

SILVA, J. A. A.; ROCHA, K. D.; FERREIRA, R. L. C.; TAVARES, J. A. Modelagem do crescimento volumétrico de clones de eucalipto (*Eucalyptus urophylla*) no Polo Gesseiro do Araripe-PE. **Anais da Academia Pernambucana de Ciência Agronômica**, Recife, v. 13/14, p.173-190, 2017.

SILVA, A. V. dos S.; SOUZA, R. G. T.; LIARTE, G. V. C. C.; LIMA, R. B.; ABREU, J. C. Classificação de sítio em plantio florestal de eucalipto no estado do Amapá. **Revista Arquivos Científicos (IMMES)**, v. 3, n. 1, p. 106-110, 2020. DOI: <https://doi.org/10.5935/2595-4407/rac.immes.v3n1p106-110>

SILVA, J. W. L. **Modelos estatísticos e técnicas de inteligência artificial para estimativa do volume de clones de *Eucalyptus* spp. com adição de variáveis climáticas** – 2020. 153 f. Tese. Pós-Graduação em Biometria e Estatística Aplicada, Universidade Federal Rural de Pernambuco. Recife, 2020.

SOUZA. S. R. R.; SILVA, J.A.A.; FERREIRA, T. AL. E.; GUERRA, O. G.M. REDES NEURAIAS PARA ESTIMATIVA VOLUMÉTRICA DE CLONES DE EUCALYPTUS SPP. NO PÓLO GESSEIRO DO ARARIPE. *Brazilian Journal of Biometrics*, 36(3), 715–729. 2018. <https://doi.org/10.28951/rbb.v36i3.286>

SKOVSGAARD, J.P.; VANCLAY, J.K. Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. **Forestry**, v. 81, n. 1, p. 13-31, 2008. <https://doi.org/10.1093/forestry/cpm041>

SPURR, S. H. **Forest inventory**. New York: Ronald Press, 1952. 476 p

STROBL, C.; BOULESTEIX, A. L.; ZEILEIS, A.; HOTHORN, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. **BMC Bioinformatics**, 8(1), 25.

STROBL, C.; BOULESTEIX, A. L.; KNEIB, T.; AUGUSTIN, T.; ZEILEIS, A. (2008). Conditional variable importance for random forests. **BMC Bioinformatics**, 9, 307.

TONINI, H.; ARCO-VERDE, M. F.; SCHWENGBER, D.; MOURÃO JUNIOR, M. Avaliação de espécies florestais em área de mata no estado de Roraima. **Cerne**, v. 12, n. 1, 2006. Disponível em: <http://www.redalyc.org/articulo.oa?id=74412102>. ISSN 0104-7760.

TURNBULL, K. J. **Population dynamics in mixed forest stands: a system of mathematical models of mixed stand growth and structure**. Washington, 1963. 186f. Dissertation (Magister Science) - University of Washington, Washington, 1963.

VALENÇA, M. **Fundamentos das redes neurais: Exemplos em Java**. 2ª. Edição, Revisado e ampliado. Olinda-PE: Livro Rápido, 2015. 386p.

VAPNIK, V. N. **The nature of Statistical learning theory**. Springer-Verlag, New York, 1995. 188p.

YATES, LUKE A., ZACH AANDAHL, SHANE A. RICHARDS, AND BARRY W. “ Cross Validation for Model Selection: A Review with Examples from Ecology.” *Ecological Monographs* 93(1): e1557. 2023. <https://doi.org/10.1002/ecm.1557>

YU, F.; XU, X. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. **Applied Energy, Amsterdam**, v. 134, p.102-113, 2014.