



ELLIEZER DE ALMEIDA MELO

**MODELAGEM DA RELAÇÃO HIPSOMÉTRICA
E DO CRESCIMENTO E PRODUÇÃO
UTILIZANDO APRENDIZAGEM DE MÁQUINA
E MODELO DE EFEITO MISTO**

**LAVRAS - MG
2019**

ELLIEZER DE ALMEIDA MELO

**MODELAGEM DA RELAÇÃO HIPSOMÉTRICA E DO
CRESCIMENTO E PRODUÇÃO UTILIZANDO APRENDIZAGEM DE
MÁQUINA E MODELO DE EFEITO MISTO**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Florestal, área de concentração em Manejo Florestal, para a obtenção do título de Doutor.

Ph.D. Natalino Calegário
Orientador

**LAVRAS - MG
2019**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Melo, Elliezer de Almeida.

Modelagem da relação hipsométrica e do crescimento e produção
utilizando aprendizagem de máquina e modelo de efeito misto /

Elliezer de Almeida Melo. - 2019.

143 p. : il.

Orientador: Natalino Calegário.

Tese (Doutorado) - Universidade Federal de Lavras, 2019.

Bibliografia.

1. Modelos Mistos. 2. Modelagem da Heterocedasticidade. 3.
Aprendizagem de Máquinas. I. Calegário, Natalino. II. Título.

ELLIEZER DE ALMEIDA MELO

**MODELAGEM DA RELAÇÃO HIPSOMÉTRICA E DO
CRESCIMENTO E PRODUÇÃO UTILIZANDO APRENDIZAGEM DE
MÁQUINA E MODELO DE EFEITO MISTO**

**MODELING OF THE HYPOMETRIC RELATIONSHIP AND
GROWTH AND PRODUCTION USING MACHINE LEARNING AND A
MIXED EFFECT MODEL**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Florestal, área de concentração em Manejo Florestal, para a obtenção do título de Doutor.

APROVADO em 14 de março de 2019.

Prof. Dr. Lucas Rezende Gomide	UFLA
Prof. Dr. Wilian Soares Lacerda	UFLA
Prof. Dr. Bruno Henrique Groenner Barbosa	UFLA
Prof. Dr. Adriano Ribeiro de Mendonça	UFES

Ph.D. Natalino Calegário
Orientador

**LAVRAS - MG
2019**

AGRADECIMENTOS

A Deus, por ter me dado condições de lutar e alcançar os objetivos pretendidos. Pelo amor incondicional, por estar comigo nos momentos de lutas, incertezas e alegrias. Obrigado, por esta conquista.

Aos meus pais, Edson de Melo e Ivalmira Lúcia de Almeida Melo, minha irmã Juliana, meus grandes incentivadores, e por seu amor incondicional. Aos meus familiares e amigos pelo incentivo e, principalmente, pela amizade.

Ao professor, orientador e amigo Natalino Calegario, pela oportunidade, pelos ensinamentos e conselhos, durante o curso de Mestrado e Doutorado.

Ao professor e amigo, Adriano Ribeiro de Mendonça, pelos ensinamentos, conselhos e, principalmente, pela amizade desde a época de UEG.

À UFLA, ao Programa de Pós-Graduação em Engenharia Florestal, aos professores do Departamento de Ciências Florestais, do Departamento de Ciências Exatas e do Departamento de Automática pela contribuição em minha formação profissional.

Aos funcionários, em especial à Chica e Juliano, pela disponibilidade de sempre poder ajudar.

Ao Instituto Federal Goiano/Campus Morrinhos, por tornar possível, por meio do plano de capacitação docente, a concretização do curso de Doutorado.

Aos meus amigos de república, do Programa de Pós-Graduação em Engenharia Florestal e de Lavras. Em especial, Fábio (Zebra), Maurício (Shima), Celso Pedro, João Antônio Lamim, Paulo Henrique (PH), Ivens, Thiago Bastos, Isaac Júnior, Joyce Almeida, Daniel Dantas, Ricardo Tayarol, Rossi Allan, José Edmar Júnior (Juninho), Ernandes e Eduane. Obrigado, pelos ensinamentos e, principalmente, pela amizade.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho, o meu muito OBRIGADO.

RESUMO GERAL

O presente estudo foi desenvolvido a partir de dados de inventário florestal coletados em parcelas temporárias e permanentes. Objetivou-se modelar a relação hipsométrica, o crescimento e a produção das florestas de eucalipto, a partir das técnicas de regressão tradicional, modelagem mista e aprendizagem de máquinas, para a obtenção de modelos capazes de representar a realidade do povoamento florestal. O estudo foi dividido em três capítulos. No primeiro capítulo, realizou-se uma revisão de literatura com o intuito de embasar o desenvolvimento do segundo e terceiro capítulos. O segundo capítulo, constituiu-se na avaliação de quatro modelos não lineares para modelagem da relação hipsométrica. Comparou-se a técnica de regressão tradicional com a técnica de modelagem mista com a adição de covariantes e modelagem da heterocedasticidade, visando obter modelos com melhor acurácia. A utilização dos modelos não lineares mistos possibilitou uma redução do erro padrão em aproximadamente 60%. O modelo de Gompertz, por proporcionar melhor acurácia foi decomposto e incluíram-se as covariáveis clone, área basal, sítio e idade e modelou-se a heterocedasticidade. A modelagem da heterocedasticidade possibilitou uma redução na estimativa do erro padrão residual e percentual de, aproximadamente, 55%, quando comparado ao modelo de efeitos mistos. Para o segundo capítulo, objetivou-se analisar o desempenho de técnicas de aprendizagem de máquinas (Redes Neurais Artificiais - RNA e Máquina de Vetor de Suporte - MVS) na projeção da área basal e volume em comparação ao método tradicional de análise de regressão, utilizando o modelo de Clutter e variações deste. Os dados são provenientes de 2.550 unidades amostrais permanentes de povoamentos clonais do híbrido *Eucalyptus grandis* x *E. urophylla*, compreendendo sete materiais genéticos. Os métodos de aprendizagem de máquinas para projeção da área basal proporcionaram boa capacidade de treinamento e generalização. Para a projeção volumétrica, os métodos apresentaram baixa capacidade de generalização. As estimativas produzidas pelo sistema de Clutter e variações deste foram superiores às técnicas de aprendizagem de máquinas.

Palavras-chave: Modelos Mistos. Modelagem da Heterocedasticidade. Aprendizagem de Máquinas. Redes Neurais Artificiais. Máquina de Vetor de Suporte. Modelo de Clutter.

GENERAL ABSTRACT

We developed this study based on forest inventory data collected from temporary and permanent plots. The objective was to model the hypsometric relationship, the growth, and the production of eucalyptus forests using traditional regression techniques, mixed modeling, and machine learning to obtain models capable of representing the reality of forest stands. We divided the study into three chapters. In the first chapter, we performed a literature review to support the development of the second and third chapters. The second chapter consisted of evaluating four nonlinear models for modeling the hypsometric relationship. We compared the traditional regression technique with the mixed modeling technique with the addition of covariates and heteroscedasticity modeling to obtain more accurate models. The use of nonlinear mixed models allowed us to reduce the standard error by approximately 60%. We decomposed the Gompertz model and included the clone, basal area, site, and age covariates to provide better accuracy, modeling the heteroscedasticity. The heteroscedasticity modeling allowed us to reduce the estimate of the residual standard error and percentage of approximately 55% when compared to the mixed effects model. For the second chapter, we aimed to analyze the performance of machine learning techniques (Artificial Neural Networks - ANN and Support Vector Machine - SVM) in the projection of the basal area and volume, comparing them to the traditional method of regression analysis using the Clutter model and variations thereof. The data derive from 2,550 permanent sample units of clonal stands of the *Eucalyptus grandis* x *E. urophylla* hybrid, comprising seven genetic materials. The machine learning methods to project the basal area provided good training and generalization skills. For the volumetric projection, the methods presented low generalization capacity. The estimates produced by the Clutter system and its variations were superior to the machine learning techniques.

Keywords: Mixed Models. Heteroskedasticity Modeling. Machine Learning. Artificial neural networks. Support Vector Machine. Clutter model.

SUMÁRIO

	PRIMEIRA PARTE	9
1	INTRODUÇÃO GERAL	9
2	REFERENCIAL TEÓRICO	13
2.1	Modelos de crescimento e produção	13
2.2	Dados longitudinais	17
2.3	Modelos mistos	18
2.4	Estruturas das matrizes de variâncias-covariâncias	24
2.5	Redes neurais	27
2.5.1	Funções de ativação	28
2.5.2	Arquitetura, treinamento e generalização de redes neurais artificiais	30
2.5.3	Redes perceptron multicamadas	31
2.5.4	Redes RBF	32
2.5.5	Máquinas de vetor de suporte	33
2.6	Estudos de inteligência artificial aplicados na área florestal	37
	REFERÊNCIAS	41
	SEGUNDA PARTE - ARTIGOS	49
	ARTIGO 1-MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE ÁRVORES CLONAIS DE EUCALIPTO UTILIZANDO MODELOS NÃO LINEARES MISTOS	49
1	INTRODUÇÃO	53
2	MATERIAL E MÉTODOS	55
2.1	Base de dados	55
2.2	Modelos analisados	55
2.3	Modelagem mista	56
2.4	Adição de covariáveis aos modelos	58
2.5	Heterogeneidade da variância	59
2.6	Análise dos modelos	59
3	RESULTADOS E DISCUSSÃO	63
3.1	Análise dos modelos hipsométricos	63
3.2	Análise dos modelos mistos	66
3.3	Inclusão de covariantes no modelo	70
3.4	Modelagem da heterocedasticidade	75
4	CONCLUSÃO	79
	ANEXO A - Ajuste das curvas de regressão para os quatros modelos de efeitos mistos testados.	81
	REFERÊNCIAS	85

ARTIGO 2 - MODELAGEM DO CRESCIMENTO E PRODUÇÃO DE POVOAMENTOS DE EUCALIPTO UTILIZANDO MODELO DE REGRESSÃO E APRENDIZAGEM DE MÁQUINA		87
1	INTRODUÇÃO	91
2	MATERIAL E MÉTODOS	95
2.1	Caracterização da área de estudo e base de dados	95
2.2	Modelagem do crescimento e produção em área basal e volume	96
2.2.1	Preparação dos dados.....	97
2.3	Modelagem pela aprendizagem de máquinas	98
2.3.1	Perceptron de múltiplas camadas (PMC).....	98
2.3.2	Função de base radial (RBF)	101
2.3.3	Máquina de vetor de suporte (MVS)	103
2.4	Modelagem pelo sistema de Equações de Clutter	105
2.5	Análises das estimativas geradas pela aprendizagem de máquinas e pelos modelos de regressão	109
3	RESULTADOS E DISCUSSÃO	111
3.1	Perceptron de múltiplas camadas	111
3.1.1	Modelagem da área basal.....	111
3.1.2	Modelagem do volume	115
3.2	Redes de Função de Base Radial (RBF).....	119
3.2.1	Modelagem da área basal.....	119
3.2.2	Modelagem do Volume.....	121
3.3	Máquina de vetor de suporte (MVS)	124
3.3.1	Modelagem da área basal.....	124
3.3.2	Modelagem do volume	126
3.4	Sistema de Equações de Clutter	130
4	CONCLUSÃO	139
	REFERÊNCIAS	141

PRIMEIRA PARTE

1 INTRODUÇÃO GERAL

Uma das funções do profissional ligado ao manejo de florestas é quantificar e prognosticar, com confiabilidade, o estoque de suas florestas. A quantificação do crescimento e da produção é de fundamental importância no planejamento florestal, pois fornece informações que subsidiam a tomada de decisões relacionadas ao manejo, as quais são baseadas em informações sobre condições de recursos florestais atuais e futuros.

Para obter o volume individual de árvores é necessário o conhecimento da altura dos indivíduos. Como sua coleta é bastante onerosa, e passível de erros de medições, fazem-se necessários o uso de modelos de regressão para estimar a altura em povoamentos florestais. Modelos para estimar essa variável, geralmente são chamados de modelos hipsométricos. Várias são as técnicas adotadas para o ajuste de um modelo hipsométrico. A complexidade do ajuste está correlacionada com a variabilidade dos dados, ou seja, uma base de dados proveniente de vários povoamentos, com particularidades específicas, se torna difícil obter estimativas precisas da variável de interesse.

O uso de modelos não lineares de efeitos mistos juntamente com a adição de covariantes e modelagem da heterogeneidade da variância (heterocedasticidade) são técnicas estatísticas capazes de obter modelos genéricos que modelam a altura de árvores com maior precisão.

Um modelo misto é composto pela parte fixa e aleatória. A presença dos efeitos aleatórios possibilita a estimativa da variância desses efeitos. Nesse caso, juntamente com a variância dos efeitos fixos, é possível modelar a heterogeneidade da variância. Vários são os trabalhos, aos quais são utilizadas essas técnicas para modelagem de dados florestais, dentre eles, podemos citar: a

modelagem da área basal de clones de eucalipto, utilizando como níveis aleatórios as unidades amostrais e as árvores dentro das mesmas (CALEGARIO et al., 2005a); estimação da altura de árvores de *Populus tremuloides* Michx., modelagem da área basal de *Picea mariana* (Mill.) BSP e da altura dominante e codominante de *Pinus contorta* var. *latifolia* Engelm. ex S. Wats, considerando aleatória a unidade amostral (YANG; HUANG, 2011) e modelagem de alturas de eucalipto, utilizando classes de DAP como fatores aleatórios (MENDONÇA; CARVALHO; CALEGARIO, 2015).

Para a quantificação do crescimento e da produção em área basal e volume de povoamentos, os modelos em nível de povoamento são os mais usuais pelas empresas florestais. Dentre eles, os sistemas de Clutter (1963) e Sullivan e Clutter (1972) são os mais difundidos no meio florestal. Tal fato está relacionado à sua facilidade de aplicação, por se tratar de um sistema de Equações consistentes e, por fornecer estimativas relativamente precisas do crescimento e da produção florestal.

Uma alternativa para a modelagem, utilizando modelos de regressão, é a utilização de técnicas de inteligência artificial (IA) como as Redes Neurais Artificiais (RNA) e a Máquina de vetor de suporte (MVS). Essas técnicas são constituídas de uma ou mais unidades de processamento, caracterizadas por neurônios artificiais, que estão conectados em rede, sendo as mesmas representadas por pesos que armazenam as relações entre variáveis e ponderam as entradas de cada unidade (BRAGA; CARVALHO; LUDERMIR, 2011; HAYKIN, 2001; SILVA; SPATTI; FLAUZINO, 2010).

O uso da IA tem despertado interesse na área florestal, por apresentar, em alguns casos, desempenho superior aos modelos de regressão (CORDEIRO et al., 2015; HAYKIN, 2001; LEK et al., 1996), por proporcionar redução dos custos de inventário florestal pela necessidade de uma amostragem reduzida para obter o mesmo nível de precisão (BINOTI; BINOTI; LEITE, 2013; CAMPOS;

LEITE, 2009), pela habilidade de aprender e generalizar, permitindo resolver problemas complexos (HAYKIN, 2001) e pela não necessidade de atender às suposições da análise de regressão (CORDEIRO et al., 2015).

Estudos de crescimento e produção florestal se enquadram na tarefa de aprendizagem denominada aproximação de funções. Essa tarefa consiste em treinar uma rede que aproxime a função desconhecida e que descreva o comportamento dos pares de entrada-saída (HAYKIN, 2001; JAIN; MAO; MOHIUDDIN, 1996). As redes mais utilizadas para a aproximação de funções são as Perceptrons de Múltiplas Camadas (MLP), Rede de Base Radial (RBF) e a MVS. Essas redes possuem uma camada de entrada, a qual recebe as variáveis de entrada e as direcionam à próxima camada e uma camada de saída (BRAGA; CARVALHO; LUDERMIR, 2011; HAYKIN, 2001).

A utilização da IA tem se mostrado uma alternativa promissora em relação às técnicas de regressão para a mensuração dos recursos florestais. Dentre os vários trabalhos que utilizam essa técnica, pode-se citar: estimação do volume de árvores (CORDEIRO et al., 2015; DIAMANTOPOULOU et al., 2009; GORGENS et al., 2009; ÖZÇELİK et al., 2010; SOARES et al., 2011, 2013), crescimento e produção (BINOTI et al., 2015), afilamento (DIAMANTOPOULOU, 2005; LEITE et al., 2011), modelos hipsométricos (BINOTI; BINOTI; LEITE, 2013; DIAMANTOPOULOU, 2012; ÖZÇELİK et al., 2013), modelagem da distribuição diamétrica (BINOTI et al., 2012; LEDUC et al., 2001).

Nesse contexto, objetivou-se, neste estudo:

Obter melhoria na qualidade de ajuste ao modelar o crescimento em altura de árvores de eucalipto, utilizando técnicas de modelagem não linear de efeitos mistos, com adição de covariantes e modelagem da heterocedasticidade;

Construir máquinas de aprendizado que estimem com eficiência o crescimento e a produção de povoamentos de eucalipto a partir de dados de inventário florestal contínuo (IFC) e compará-las ao método tradicional de análise de regressão, utilizando o modelo linear proposto por Clutter (1963) e modificações deste.

2 REFERENCIAL TEÓRICO

2.1 Modelos de crescimento e produção

Desde MacKinney e Chaiken (1939), vários pesquisadores têm usado as técnicas de regressão para construir modelos para estimar o crescimento e a produção florestal. Até 1962, os modelos de crescimento eram desenvolvidos independentemente dos modelos de produção, sem nenhuma tentativa em desenvolver modelos que possuíssem compatibilidade entre valores preditos de crescimento e de produção. A incompatibilidade entre os modelos de crescimento e de produção gerava inconsistências, pois os incrementos estimados pelo modelo de crescimento não apresentavam os mesmos valores de volumes preditos pelo modelo de produção (CLUTTER, 1963).

Clutter (1963) foi o primeiro pesquisador a introduzir a noção de compatibilidade entre modelos de crescimento e de produção, com o uso das técnicas de cálculo diferencial. Segundo esse autor, os referidos modelos são considerados compatíveis quando modelos de produção em área basal e volume podem ser obtidos pelo somatório dos modelos de crescimento preditos, considerando períodos de crescimento definidos ou, mais precisamente, quando o modelo de produção é obtido pela integral do modelo de crescimento e esse é obtido pela derivativa do modelo de produção.

A compatibilidade desenvolvida por Clutter (1963) surgiu do modelo de produção desenvolvido por Schumacher (1939) e aplicado, primeiramente, por MacKinney e Chaiken (1939) em pinus.

$$\ln V_i = \beta_0 + \beta_1 I_i^{-1} + \beta_2 S_i + \beta_3 \ln(G_i) + \varepsilon_i \quad (1)$$

Em que: V_i = volume total da *i-ésima* unidade amostral ($m^3 \text{ ha}^{-1}$); I_i = idade de medição da *i-ésima* unidade amostral (anos); G_i = área basal da *i-ésima* unidade amostral ($m^2 \text{ ha}^{-1}$); S_i = índice de sítio da *i-ésima* unidade amostral; β_s = parâmetros da Equação; \ln = logaritmo neperiano.

Com o intuito de adequar o modelo estatisticamente, são adicionadas iterações ao modelo de produção volumétrica (EQUAÇÃO 1.1). Segundo Clutter (1963), caso as iterações não sejam significativas, isto é, não contribuam para a redução da soma de quadrado do resíduo, deve-se utilizar o modelo sem adição das mesmas.

$$\ln V_i = \beta_0 + \beta_1 I_i^{-1} + \beta_2 S_i + \beta_3 \ln(G_i) + \beta_4 I_i^{-1} S_i + \beta_5 I_i^{-1} \ln(G_i) + \beta_6 S_i \ln(G_i) + \varepsilon_i \quad (1.1)$$

Para obter a função de crescimento em volume, a Equação (1) é derivada em relação à idade, resultando na Equação (2). A derivativa da área basal $\ln(G_i)$ se justifica pela sua relação funcional com a idade.

$$\frac{\partial V}{\partial I} = V_i \left(-\beta_1 I_i^{-2} + \beta_3 G_i^{-1} \frac{\partial G}{\partial I} \right) \quad (2)$$

Clutter (1963) afirma que a Equação (2) não é adequada para aplicação prática, pois envolve a taxa de crescimento em área basal até então não disponível. Para a obtenção da taxa de crescimento volumétrico, Schumacher e Coile (1960) desenvolveram a seguinte função de produção em área basal.

$$\ln(G_i) = \alpha_0 + \alpha_1 I_i^{-1} + \alpha_2 S_i + \alpha_3 I_i^{-1} S_i + \alpha_4 I_i^{-1} \ln(G_{i20}) + \varepsilon_i \quad (3)$$

Em que: α_s = parâmetros da Equação; G_{i20} = área basal da *i*-ésima unidade amostral ($m^2 ha^{-1}$) na idade definida pelos autores como 20 anos.

A função (3), após derivada em função da idade, fornece a taxa de crescimento em área basal, representada pela Equação (4).

$$\frac{\partial G}{\partial I} = I_i^{-1} G_i (\alpha_0 + \alpha_2 S_i - \ln G_i) \quad (4)$$

Assim, é possível encontrar a taxa de crescimento volumétrico adicionando à função (2), a Equação (4), resultando na Equação (5).

$$\frac{\partial V}{\partial I} = V_i \left(-\beta_1 I_i^{-2} + \beta_3 \alpha_0 I_i^{-1} + \beta_3 \alpha_2 I_i^{-1} S_i - \beta_3 \ln(G_i) I_i^{-1} \right) \quad (5)$$

Para obter as funções para projeção da produção em área basal e volume, a Equação (4) é integrada nos intervalos de G_1 a G_2 e I_1 a I_2 , e substituída na Equação (1), resultando nas Equações (6 e 7), respectivamente (CLUTTER, 1963).

$$\ln G_{2i} = \ln G_{1i} \left(\frac{I_{1i}}{I_{2i}} \right) + \alpha_0 \left(1 - \frac{I_{1i}}{I_{2i}} \right) + \alpha_2 \left(1 - \frac{I_{1i}}{I_{2i}} \right) S_i + \varepsilon_i \quad (6)$$

$$\ln V_{2i} = \beta_0 + \beta_1 \left(\frac{1}{I_{2i}} \right) + \beta_2 S_i + \beta_3 \left(\frac{I_{1i}}{I_{2i}} \right) \ln(G_{1i}) + \beta_4 \left(1 - \frac{I_{1i}}{I_{2i}} \right) + \beta_5 \left(1 - \frac{I_{1i}}{I_{2i}} \right) S_i + \varepsilon_i \quad (7)$$

Em que: V_{2i} = volume total na idade (I_2) da *i-ésima* unidade amostral ($\text{m}^3 \text{ha}^{-1}$); I_{1i} e I_{2i} = idades de medição atual e futura da *i-ésima* unidade amostral, respectivamente (anos); $\hat{\beta}_{4*} = \hat{\beta}_{3*}\hat{\alpha}_0$; $\hat{\beta}_{5*} = \hat{\beta}_{3*}\hat{\alpha}_2$; $\varepsilon_i = \varepsilon_1 + \varepsilon_2$ (erro aleatório).

O modelo (7) é utilizado para projetar o volume em função das variáveis idade inicial e futura, índice de sítio, e área basal inicial. Quando a projeção não existir, ou seja, ($I_2 = I_1$), o modelo (7) reduz ao modelo (1). Chamado de sistema de Equações simultâneas, pode ser utilizado tanto para predizer o volume quanto para obter o crescimento ou projeção do volume para idades futuras (SULLIVAN; CLUTTER, 1972).

Apesar de as estimativas dos parâmetros $\hat{\beta}_{4*}$ e $\hat{\beta}_{5*}$ serem obtidas por relações que garantem a consistência do sistema, Sullivan e Clutter (1972) afirmam que estes não são considerados os estimadores mais eficientes estatisticamente. Para melhorar a consistência do ajuste, pesquisadores como Salles et al. (2012), utilizam-se da técnica de estimação dos parâmetros, pelo método de mínimos quadrados em dois estágios.

Embora Clutter (1963) tenha desenvolvido um sistema de Equações consistente, capaz de predizer e projetar o crescimento e a produção em área basal e volume, havia duas irregularidades encontradas no uso desses modelos, mais tarde apresentadas por Sullivan e Clutter (1972).

Segundo Sullivan e Clutter (1972), a primeira irregularidade encontrada está relacionada aos parâmetros dos modelos que são considerados independentes, o que na verdade não são. A outra está relacionada ao descumprimento da suposição de independência da regressão. Nesse caso, para a projeção do crescimento e produção em área basal e volume, são necessárias medições sucessivas dessas variáveis nas unidades amostrais, o que torna inconsistente estatisticamente dizer que essas observações são independentes.

A partir deste trabalho desenvolvido por Clutter (1963), vários outros sistemas para modelar o crescimento e a produção florestal foram desenvolvidos. Merecem destaque os trabalhos de Buckman (1962), Curtis (1967), Moser e Hall (1969) e Sullivan e Clutter (1972).

2.2 Dados longitudinais

Os dados provenientes de inventário florestal contínuo (IFC), que são utilizados para estudos de crescimento e produção florestal são considerados longitudinais pelo fato da base de dados ser formada pela sequência temporal de duas ou mais observações em cada indivíduo ou unidade amostral.

De acordo com McCulloch e Searle (2001), os dados longitudinais geralmente apresentam uma estrutura desbalanceada. Essa estrutura desbalanceada ocorre, em decorrência das medidas, em cada indivíduo, serem obtidas em tempos diferentes. Além disso, o número de observações medidas no tempo nem sempre é igual para todos os indivíduos.

Os dados longitudinais apresentam uma estrutura hierárquica, uma vez que as medidas repetidas são aninhadas dentro do indivíduo. Tal fato permite dizer que as suposições das observações aninhadas no indivíduo são consideradas dependentes com erros correlacionados, enquanto as observações entre indivíduos são independentes. Essa suposição de erros correlacionados exige a modelagem da matriz de variância-covariância dos dados (PINHEIRO; BATES, 2000).

Grande parte dos estudos na área florestal ignora a suposição de dependência das observações aninhadas, tratando-as como independentes. A utilização de modelos lineares ou não lineares sem a modelagem da correlação causa inferências menos confiáveis, comprometendo as estimativas dos parâmetros.

Um modelo estatístico que permite analisar dados longitudinais desbalanceados em estrutura hierárquica, incorporando a dependência e a estrutura de correlação dos erros são os modelos lineares e não lineares de efeitos mistos. São chamados de modelos mistos pela presença de efeitos fixos e aleatórios, os quais estimam tanto os efeitos fixos quanto os componentes de variância (McCULLOCH; SEARLE 2001).

2.3 Modelos mistos

Os modelos mistos são uma importante ferramenta na modelagem do crescimento e da produção florestal, por possibilitarem que se obtenham melhores estimativas, bem como pela capacidade de lidar de maneira adequada com os erros correlacionados advindos de observações mensuradas no tempo (GREGORIE; SCHABENBERGER, 1996; McCULLOCH; SEARLE, 2001; TRINCADO; BURKHART, 2006; YANG; HUANG, 2011; ZHAO; WILSON; BORDER, 2005).

Os modelos mistos para dados longitudinais permitem que os coeficientes da regressão variem entre os indivíduos. Segundo Pinheiro e Bates (2000), esses modelos oferecem uma flexibilidade em modelar a correlação entre e intraindividual, presente nesses tipos de dados. No componente intraindividual, uma mudança longitudinal entre as observações de um indivíduo é descrita pelo modelo de regressão com intercepto e inclinação populacional. Já entre indivíduos ocorre a variação no intercepto e inclinação individual.

Para Fausto et al. (2008), o uso desses modelos permite não somente descrever a tendência temporal, levando em conta a correlação que existe entre medidas sucessivas, como também estimar a variação no intercepto e a taxa de mudança ao longo do tempo. Ainda, a utilização de modelos mistos permite

obter a predição de efeitos aleatórios, na presença de efeitos fixos, pelos *Best linear unbiased prediction* - BLUPs (CAMARINHA FILHO, 2002).

A aplicação da teoria de modelos mistos é adequada para dados em que a variabilidade entre os indivíduos é maior do que a variabilidade dentro dos indivíduos, como é o caso das curvas de crescimento. Esse modelo assume que o padrão de crescimento tem a mesma forma funcional para todos os indivíduos, mas que os indivíduos podem apresentar comportamento longitudinal diferentes. Isso faz com que cada indivíduo tenha a sua própria curva de crescimento especificada pelos coeficientes de regressão (FAUSTO et al., 2008).

Matricialmente, o modelo linear misto descrito por McCulloch e Searle (2001) é denotado por (8):

$$y = X\beta + Zu + \varepsilon \quad (8)$$

Em que: y = vetor de dados observados ($n \times 1$); X = matriz de incidência dos efeitos fixos (conhecida) ($n \times p+1$); β = vetor de efeitos fixos desconhecidos ($p+1 \times 1$); Z = matriz de incidência dos efeitos aleatórios (conhecida) ($n \times q$); u = é o vetor de efeitos aleatórios desconhecidos ($q \times 1$) e ε = vetor de erros aleatórios ($n \times 1$).

Modelos não lineares de efeitos mistos são preferíveis para modelar o crescimento e a produção florestal, em razão da relação não linear entre as variáveis dependentes e independentes e por seus parâmetros apresentarem significado biológico. O modelo não linear de efeito misto pode ser representado como em Yang e Huang (2011) pela expressão (9):

$$y_i = f(x_i, \beta, u_i) + \varepsilon_i \quad (9)$$

Em que: y_i = vetor da variável resposta ou variável dependente da *i-ésima* árvore ou unidade amostral (indivíduo); $f(\cdot)$ = função não linear; x_i = matriz de incidência de efeitos fixos; β = vetor de parâmetros de efeitos fixos (comum a todos os indivíduos); u_i = vetor de parâmetros de efeitos aleatórios associado ao *i-ésimo* indivíduo; ε_i = vetor de erros aleatórios para o *i-ésimo* indivíduo.

De acordo com McCulloch e Searle (2001), os vetores de parâmetros u_i e ε_i são independentes com distribuição normal, média zero e matrizes de variâncias-covariâncias \mathbf{D} e \mathbf{R} , respectivamente. Essas matrizes são positivas definidas conhecidas e, portanto, não singulares, dadas por:

$$\text{Var}(\mathbf{u}) = E[\mathbf{u}\mathbf{u}'] = \mathbf{D} \text{ e } \text{Var}(\boldsymbol{\varepsilon}) = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \mathbf{R} \quad (10)$$

Em que: \mathbf{D} é a matriz de variâncias-covariâncias dos efeitos aleatórios e \mathbf{R} é a matriz de variâncias-covariâncias residuais.

Matricialmente, tem-se:

$$\text{Var} \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \sigma^2. \quad (11)$$

Em que: σ^2 é uma constante positiva.

Desse modo, tem-se que:

$$\mathbf{V} = \text{Var}(y) = \text{Var}(X\beta + Z\mathbf{u} + \boldsymbol{\varepsilon}) = Z \text{Var}(\mathbf{u}) Z' + \mathbf{R} = \mathbf{ZDZ}' + \mathbf{R} \quad (12)$$

Assume-se ainda que \mathbf{V} é não singular, e

$$\mathbf{E}[\mathbf{y}] = \mathbf{E}[\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}] = \mathbf{X}\boldsymbol{\beta} \quad (13)$$

Assim, tem-se que:

$$\mathbf{y} \sim \mathbf{N}(\mathbf{X}\boldsymbol{\beta}; \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R}) \quad (14)$$

Conclui-se que os efeitos fixos entram apenas na média, enquanto que a matriz de efeitos aleatórios e a variância entram apenas na variância de \mathbf{y} (McCULLOCH; SEARLE, 2001).

Para McCulloch e Searle (2001), com a presença dos efeitos aleatórios, o interesse do modelador estaria na estimativa da variância desses efeitos. Assim, os dados teriam duas fontes de variação aleatória: a variância ligada aos efeitos aleatórios (covariância) e a variância do erro - matriz de variância-covariância. O conjunto dessas variâncias é denominado de componente de variância (VC). Nesse caso, a soma dos VC é a variância da variável observada.

Considerando um ajuste de regressão aos dados observados, cada medida individual difere da média total dos indivíduos que é definido pelo erro aleatório. Esses erros são independentes e são utilizados para determinação dos componentes da variância. A covariância é determinada pelos erros observados na estimativa da média intraindividual e apresentam a característica de dependência (FAUSTO et al., 2008).

Ao estruturar um modelo misto, a escolha de variáveis com potencial para serem consideradas aleatórias exige uma certa experiência (McCULLOCH; SEARLE, 2001). Para esses autores, a decisão sobre considerar um fator como fixo ou aleatório é dependente da variação dos efeitos (variáveis) em relação aos indivíduos ou unidades amostrais, que são consideradas como amostras aleatórias de uma população. Ou seja, qualquer efeito (variável) que não seja constante para todos os indivíduos ou unidades amostrais podem ser considerados aleatórios.

Na área florestal, diferentes variáveis aleatórias são utilizadas na modelagem do crescimento e produção. Dentre os trabalhos, destacam-se: a modelagem da área basal de clones de eucalipto utilizando como efeitos aleatórios as unidades amostrais e as árvores dentro das mesmas (CALEGARIO et al., 2005a); estimação da altura de árvores de *Populus tremuloides* Michx, modelagem da área basal de *Picea mariana* (Mill.) BSP e da altura dominante e codominante de *Pinus contorta* var. *latifolia* Engelm. ex S. Wats, considerando aleatório a unidade amostral (YANG; HUANG, 2011) e modelagem de alturas de eucalipto utilizando a classe de diâmetro como efeito aleatório (MENDONÇA; CARVALHO; CALEGARIO, 2015).

Yang e Huang (2011), aplicando a teoria dos modelos não lineares mistos para modelagem da relação hipsométrica, área basal e altura média das árvores dominantes e codominantes, afirmam a necessidade de várias combinações de efeitos aleatórios nos parâmetros para se chegar a modelos com boas estimativas. Os critérios de informação de Akaike (AIC) e Bayesiano (BIC) são as estatísticas utilizadas por esses autores e por Calegario et al. (2005a, 2005b) para a escolha das melhores estruturas aleatórias dos modelos.

Calegario et al. (2005a), modelando o crescimento em área basal, baseado em modelos lineares mistos utilizaram dois níveis de efeitos aleatórios: unidades amostrais e árvores dentro das unidades amostrais. De acordo com os resultados, a inclusão dos efeitos aleatórios proporcionaram redução do erro padrão residual, em 15 vezes, quando comparado com o modelo de efeito fixo. Ainda, as estimativas da área basal no segundo nível (árvores) proporcionaram estimativas precisas, com os valores estimados muito próximos dos valores observados.

Segundo Diggle et al. (2002) e Molenberghs e Verbeke (2001), citados por Fausto et al. (2008), efeitos aleatórios no intercepto representam a heterogeneidade entre os indivíduos, decorrente de fatores (genéticos e

ambientais) não medidos. Entretanto, parte da variabilidade dos dados entre e intraindivíduos pode ser explicada pela falta da inclusão no modelo de covariáveis que explicam a variabilidade da variável resposta. Nesse caso, é necessário incluir efeitos aleatórios em outras covariáveis. Para Fausto et al. (2008), as covariáveis definidas como variáveis aleatórias devem ser dependentes do tempo e podem, também, fazer parte da matriz de efeitos fixos.

Diferentes métodos têm sido propostos para estimar os parâmetros dos modelos não lineares mistos. Davidian e Giltinan (1995), Pinheiro e Bates (2004) e Vonesh e Chinchilli (1997) apresentam, de forma detalhada, o desenvolvimento e a aplicação de alguns métodos. Porém, os métodos da máxima verossimilhança (ML) e máxima verossimilhança restrita (RML) são os mais empregados (YANG; HUANG, 2011).

Nos modelos não lineares mistos, a obtenção da convergência do algoritmo exige maior esforço computacional. Geralmente, as integrações numéricas para maximizar a função de probabilidade são difíceis, uma vez que os parâmetros aleatórios entram na função de uma forma não linear. Nesse contexto, torna-se necessário a maximização da função, por meio de alguma aproximação numérica. As aproximações mais utilizadas na área florestal são: a) aproximação adaptativa Gaussiana em quadratura (AGQ) de Pinheiro et al. (1995); b) aproximação de primeira ordem (FO) de Sheiner e Beal (1980) e c) aproximação da expectativa condicional de primeira ordem (FOCE).

Yang e Huang (2011), avaliando os métodos ML e RML para estimativa paramétrica e para a aproximação numérica das técnicas AGQ, FO e FOCE em dados independentes e correlacionados (longitudinais), verificaram que o método FOCE apresentou as melhores estimativas para os dados correlacionados. Já para os dados independentes, o método FO mostrou-se superior. Os métodos ML e RML produziram resultados muito próximos, tanto para os dados independentes como para os dados correlacionados.

Vários pesquisadores têm utilizado esses métodos para modelagem florestal. Merecem destaque os trabalhos de Temesgen, Monleon e Hann (2008), Trincado e Burkhart (2006) e Zheng (2010).

2.4 Estruturas das matrizes de variâncias-covariâncias

Segundo Yang e Huang (2011), em dados longitudinais, é inevitável a presença da autocorrelação residual mesmo com a inclusão de efeitos aleatórios. Portanto, estruturas de covariância são utilizadas para modelar a dependência dos erros (\mathbf{R}_i).

Para Littell (2006) e Yang e Huang (2011), a escolha de uma estrutura de covariância depende do ajuste de possíveis candidatas e para a definição da melhor deve-se utilizar critérios estatísticos como o Critério de Informação da Akaike (AIC) e o Critério de Informação Bayesiano (BIC). Littell (2006) apresenta as matrizes de covariância mais utilizadas na modelagem de dados longitudinais. Considerando dados provenientes de quatro medições têm-se:

a) Componente de variância (VC)

Essa matriz supõe independência e homogeneidade de variâncias em todos os tempos. Possui apenas um parâmetro: a variância (σ^2) e é dada por:

$$\mathbf{R}_i = \begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}.$$

b) Autorregressiva de primeira ordem - AR(1)

A estrutura autorregressiva de primeira ordem tem variâncias homogêneas e correlações que diminuem, exponencialmente, com a distância. Isso quer dizer que as observações feitas em tempos mais próximos têm maior correlação do que aquelas feitas em tempos mais distantes. Essa estrutura tem dois parâmetros desconhecidos: a variância (σ^2) e a correlação (ρ), sendo $|\rho| < 1$. A estrutura é dada por:

$$R_i = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}.$$

c) Autorregressiva heterogênea de primeira ordem - ARH(1)

A matriz ARH(1) apresenta a mesma estrutura de correlação da AR(1), contudo, as variâncias e covariâncias são heterogêneas. O número de parâmetros desconhecidos depende do número de medições repetidas mais o parâmetro de correlação. Sendo o número de tempos medidos igual ao número de variâncias. A estrutura é dada por:

$$R_i = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{pmatrix}.$$

d) Simetria composta (CS)

A matriz CS possui variâncias homogêneas em todos os tempos e as covariâncias são iguais, independentemente do quão próximo foram realizadas

as medidas. Os parâmetros desconhecidos são: a variância (σ^2) e a correlação (ρ). A estrutura é dada por:

$$R_i = \begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}.$$

e) Simetria composta heterogênea (CSH)

A simetria composta heterogênea possui correlações homogêneas e as variâncias são heterogêneas. As covariâncias são dadas pelo produto da raiz quadrada entre as medidas de variância no *i-ésimo* tempo com a correlação. Como a ARM(1), o número de parâmetros é igual ao número de medições, representadas pelas variâncias, mais a correlação. A matriz é dada por:

$$R_i = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{pmatrix}.$$

f) Huynh e Feldt (HF)

A estrutura de Huynh e Feldt possui variâncias e covariâncias heterogêneas. As covariâncias são formadas pela média aritmética entre as medidas de variância no *i-ésimo* tempo, subtraída da diferença entre a média das variâncias e a média das covariâncias (λ). Como nas estruturas ARM(1) e CSH, o número de parâmetros é igual ao número de medições, representadas pelas variâncias, mais a correlação. A estrutura é dada por:

$$R_i = \begin{pmatrix} \sigma_1^2 & \frac{\sigma_1^2 + \sigma_2^2}{2} - \lambda & \frac{\sigma_1^2 + \sigma_3^2}{2} - \lambda & \frac{\sigma_1^2 + \sigma_4^2}{2} - \lambda \\ \frac{\sigma_2^2 + \sigma_1^2}{2} - \lambda & \sigma_2^2 & \frac{\sigma_2^2 + \sigma_3^2}{2} - \lambda & \frac{\sigma_2^2 + \sigma_4^2}{2} - \lambda \\ \frac{\sigma_3^2 + \sigma_1^2}{2} - \lambda & \frac{\sigma_3^2 + \sigma_2^2}{2} - \lambda & \sigma_3^2 & \frac{\sigma_3^2 + \sigma_4^2}{2} - \lambda \\ \frac{\sigma_4^2 + \sigma_1^2}{2} - \lambda & \frac{\sigma_4^2 + \sigma_2^2}{2} - \lambda & \frac{\sigma_4^2 + \sigma_3^2}{2} - \lambda & \sigma_4^2 \end{pmatrix}.$$

Na área florestal, várias estruturas de correlação são utilizadas para a correção da pressuposição da independência dos dados. Dentre os vários trabalhos que modelam a autocorrelação, pode-se citar: modelagem do crescimento em área basal de clones de eucalipto, utilizando-se as estruturas autorregressivas AR(1) e de média móvel (ARMA) de Box, Jenkins e Reinsel (1994) (CALEGARIO et al., 2005a); modelagem da relação hipsométrica de árvores de eucalipto e *Picea abies* (L) Karst., utilizando-se as estruturas AR(1) e ARMA (CALEGARIO et al., 2005b) e poder das covariáveis (varPower) de Pinheiro e Bates (2000) (MEHTÄTALO, 2004; MENDONÇA; CARVALHO; CALEGARIO, 2015); modelagem do crescimento e produção volumétrica de clones de eucalipto, utilizando a estrutura varPower (CARVALHO et al., 2011); modelagem da área basal de povoamentos de *Picea mariana* (Mill.) BSP e da altura média das árvores dominantes e codominantes de povoamentos de *Pinus contorta* var. *latifolia* Engelm. ex S. Wats utilizando a estrutura AR(1) (YANG; HUANG, 2011).

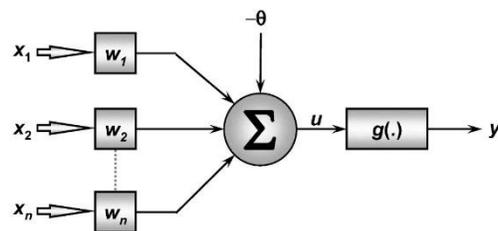
2.5 Redes neurais

Rede neural artificial (RNA) é uma técnica inspirada no funcionamento do cérebro humano, onde neurônios artificiais, conectados em rede, são capazes

de aprender e generalizar (BRAGA; CARVALHO; LUDERMIR, 2011; HAYKIN, 2001; SILVA; SPATTI; FLAUZINO, 2010).

O primeiro modelo de neurônio artificial, proposto por McCulloch e Pitts em 1943, foi baseado no neurônio biológico. Os sinais de entrada (x_1, x_2, \dots, x_n), análogos aos dendritos, são ponderados pelo conjunto de pesos (w_1, w_2, \dots, w_n), de acordo com a relevância de cada uma das entradas. No corpo celular artificial, é produzido um potencial de ativação (u) dado pela diferença entre a soma ponderada das entradas e o limiar de ativação (θ), que, posteriormente, ativa a função de ativação [$g(u)$] e gera a saída (SILVA; SPATTI; FLAUZINO, 2010).

Figura 1 - Estrutura do neurônio artificial.



Fonte: Silva, Spatti e Flauzino (2010).

2.5.1 Funções de ativação

Silva, Spatti e Flauzino (2010) classificam as funções de ativação em dois grupos: parcialmente diferenciáveis (FAPD), quando existem pontos cuja derivada de primeira ordem é inexistente; e totalmente diferenciáveis (FATD), quando é possível obter a derivada em qualquer ponto da função. Assim, as FATDs representadas pelas funções: linear, sigmoideal e gaussiana, são as mais indicadas na construção das RNAs, por possibilitar a minimização do erro durante o processo de treinamento. As principais funções de ativação FAPD e FATD são apresentadas na Tabela 1.

Tabela 1 - Resumo das principais funções de ativação.

Tipo de FA	Notação Matemática	Comentário
Degrau	$g(u) \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases}$	É uma FAPD. Assume valores unitários positivos quando o potencial de ativação do neurônio for maior ou igual a zero; caso contrário, o resultado assume valores nulos.
Degrau bipolar ou sinal	$g(u) \begin{cases} 1, & \text{se } u > 0 \\ 0, & \text{se } u = 0 \\ -1, & \text{se } u < 0 \end{cases}$	É uma FAPD. Assume valores unitários positivos quando o potencial de ativação do neurônio for maior que zero; valor nulo quando o potencial for nulo; e valores unitários negativos quando o potencial for menor do que zero.
Logística	$g(u) = \frac{1}{1 + e^{-\beta u}}$	É uma FATD. O resultado assume valores reais entre zero e um.
Tangente hiperbólica	$g(u) = \frac{1 - e^{-\beta u}}{1 + e^{-\beta u}}$	É uma FATD. O resultado assume valores reais entre -1 e 1.
Gaussiana	$g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}}$	É uma FATD. A saída do neurônio produz resultados iguais para aqueles valores de potencial de ativação que estejam a uma mesma distância de seu centro (média), sendo que a curva é simétrica em relação a este.
Linear	$g(u) = u$	É uma FATD. A função de ativação linear produz resultados de saída idênticos aos valores do potencial de ativação.

Fonte: Silva, Spatti e Flauzino (2010).

Em que: β = constante associada ao nível de inclinação da função logística frente ao seu ponto de inclinação; u = potencial de ativação; c = centro da função gaussiana, associado aos próprios pesos; σ^2 = variância da função gaussiana (raio).

2.5.2 Arquitetura, treinamento e generalização de redes neurais artificiais

Segundo Silva, Spatti e Flauzino (2010), a arquitetura de uma rede neural é definida pela forma em que os neurônios estão arranjados ou dispostos, uns em relação aos outros. De acordo com Lek e Guégan (2000), a arquitetura *feedforward* de camadas múltiplas é a mais empregada na solução de variados problemas, dentre eles aqueles relacionados à aproximação de funções. Entre os principais tipos de redes com essa arquitetura, destacam-se o *Perceptron* multicamadas (*Multilayer Perceptron* - MLP) e as redes de base radial (*Radial Basis Function* - RBF).

São características dessas redes a presença de uma ou mais camadas escondidas (intermediárias) entre as camadas de entrada e saída, em que seus neurônios, na rede MLP, geralmente, possuem funções de ativação de base sigmoideal. Na rede RBF, a função de ativação é gaussiana (BRAGA; CARVALHO; LUDERMIR, 2011; HAYKIN, 2001; SILVA; SPATTI; FLAUZINO, 2010).

O treinamento de uma rede consiste no ajuste de pesos e limiares (bias ou viés) de seus neurônios, tendo-se como objetivo final a generalização de soluções a serem produzidas pelas suas saídas. Esse processo de ajuste envolve a utilização de um algoritmo de aprendizagem capaz de extrair relações entre os valores de entrada-saída. Essa estratégia de ajuste a partir dos pares de amostras (entrada-saída) é chamada de treinamento supervisionado (SILVA; SPATTI; FLAUZINO, 2010).

Por outro lado, a generalização é definida como a capacidade da rede, após treinada, estimar soluções a partir de entradas desconhecidas. Essa capacidade de generalização pode ser influenciada pela definição da topologia de uma RNA. Conforme relatado por Braga, Carvalho e Ludermir (2011), um dimensionamento adequado da topologia representa uma das maiores

dificuldades. Segundo esses autores, o número de camadas, o número de neurônios em cada camada e a função de ativação devem ser definidos de acordo com a análise prévia dos dados e da complexidade do problema. Uma RNA superdimensionada pode resultar em *overfitting* e uma subdimensionada em *underfitting*. Uma forma de evitar o *overfitting* é estimar o erro de generalização durante o processo de treinamento (REED, 1993).

Para a escolha de um modelo, o ideal é que o mesmo apresente um bom treinamento e uma boa capacidade de generalização. Segundo Geman, Bienenstock e Doursat (1992), a compatibilidade entre topologia e o problema deve ser encarada como um dilema entre o viés (rigidez) e a variância (flexibilidade) do modelo, ou seja, que não seja muito rígido a ponto de não modelar fielmente os dados e que não seja excessivamente flexível a ponto de modelar o ruído.

2.5.3 Redes perceptron multicamadas

De acordo com Lek e Guégan (2000), a rede MLP ou PMC é um sistema poderoso por ser capaz de modelar relações complexas entre variáveis, permitindo prever uma saída a partir de variáveis de entrada. Ainda, Braga et al. (2011) mencionam a aplicabilidade dessa estrutura de RNA na resolução de problemas de classificação, regressão e previsão.

A rede MLP pertence à arquitetura *feedforward* de camadas múltiplas, cujo treinamento é efetivado de forma supervisionada. Essa arquitetura garante que o fluxo de informações da rede, independentemente da quantidade de camadas intermediárias, seja sempre realizado em um único sentido: da camada de entrada em direção à camada neural de saída (HAYKIN, 2001; SILVA; SPATTI; FLAUZINO, 2010).

O algoritmo de aprendizado utilizado pela rede MLP é denominado *backpropagation* ou algoritmo de retropropagação do erro, também conhecido como regra delta generalizada, proposto por Rumelhart, Hinton e Williams em 1986. Segundo Haykin (2001) e Silva, Spatti e Flauzino (2010), o treinamento de uma rede é realizado mediante duas fases: propagação adiante (*forward*) e propagação reversa (*backward*).

Na fase *forward*, a partir das entradas da rede associadas aos valores iniciais de pesos e bias obtém-se as saídas. Em seguida, as respostas produzidas pela saída da rede são comparadas com as respectivas respostas desejadas. Caso os desvios (erros) produzidos entre a saída da rede e a saída desejada não forem satisfatórios, aplica-se a fase *backward* (BRAGA; CARVALHO; LUDERMIR, 2011; HAYKIN, 2001; SILVA; SPATTI; FLAUZINO, 2010).

Esses mesmos autores descrevem que na fase *backward* ocorrem as alterações dos valores de pesos e bias de todos os neurônios da rede, por meio do processo de retropropagação do erro para os neurônios das camadas anteriores. Assim, como não se conhecem os valores desejados dos neurônios pertencentes às camadas intermediárias, os ajustes de pesos e bias são obtidos por intermédio de estimativas dos erros produzidos por aqueles neurônios da camada posterior, os quais já foram previamente ajustados.

Essas fases sucessivas são executadas até que se encontrem estimativas de pesos e bias que minimizem a soma do erro quadrático (SEQ ou RMSE) produzido pelas respostas da rede frente àquelas desejadas.

2.5.4 Redes RBF

As redes RBF (*radial basis function*), também denominadas de funções de base radial, são empregadas na solução de diversos problemas que envolvam classificação e aproximação de funções. Estas pertencem à arquitetura

feedforward, cujo treinamento é realizado de forma não supervisionada e supervisionada, respectivamente. Geralmente, as redes RBF são constituídas por apenas duas camadas: a camada intermediária, com neurônios com função de ativação gaussiana, e a camada de saída, com função de ativação linear (BRAGA; CARVALHO; LUDERMIR, 2011; HAYKIN, 2001; SILVA; SPATTI; FLAUZINO, 2010).

Braga, Carvalho e Ludermir (2011) e Silva, Spatti e Flauzino (2010) dividem o treinamento das redes RBF em dois estágios distintos. O primeiro estágio está ligado ao ajuste dos pesos (c) e da variância da função gaussiana (σ^2) dos neurônios da camada intermediária. Nessa etapa, o treinamento utilizado não é supervisionado e está diretamente relacionado com a alocação das funções de base radial. Já no segundo estágio, o treinamento é realizado de forma supervisionada e objetiva estimar os valores de pesos e bias dos neurônios da camada de saída.

Assim, os valores de entrada serão as respostas produzidas pelas funções de ativação gaussiana dos neurônios da camada intermediária e os neurônios da camada de saída (linear) realizará somente uma combinação linear das funções gaussianas, obtendo a saída da rede (HAYKIN, 2001).

O algoritmo de aprendizado é a regra delta generalizada, sendo o processo de treinamento idêntico ao das redes MLP, com exceção do início do treinamento, sendo esse, para as redes RBF, iniciado na camada intermediária em direção à camada de saída após a finalização do primeiro estágio de treinamento (SILVA; SPATTI; FLAUZINO, 2010).

2.5.5 Máquinas de vetor de suporte

De acordo com Vapnik (1995), as máquinas de vetor de suporte (MVS) constituem uma técnica embasada na teoria de aprendizado estatístico (TAE).

Segundo Lorena e Carvalho (2003), a TAE visa estabelecer condições matemáticas que permitam a escolha de uma função com bom desempenho para os conjuntos de treinamento e teste, de forma a maximizar a generalização.

Essas redes são aplicadas em tarefas que envolvem classificação e aproximação de funções como as redes MLP e RBF. São consideradas similares à rede RBF por possuírem alimentação adiante (único sentido) e uma única camada intermediária de unidades não lineares (HAYKIN, 2001).

Segundo Smola et al. (1999), as principais características que têm despertado interesse na utilização das MVSs são: a boa capacidade de generalização, garantindo a geração de classificadores com bons resultados de generalização e, portanto, evitando o *overfitting*; robustez em grandes dimensões, possibilitando trabalhar com base de dados de grande dimensão sem que haja *overfitting*, o que seria impossível nos classificadores gerados por outros métodos inteligentes (MLP e RBF); convexidade da função objetivo, implicando na otimização de uma função quadrática, que possui apenas um mínimo global, diferentemente de outros algoritmos pela presença de mínimos locais; e por possuir teoria bem definida e estabelecida dentro da matemática e estatística.

Diferentemente das redes MLP e RBF, a construção do algoritmo de treinamento por vetor de suporte depende da formação do núcleo do produto interno entre um vetor de suporte, dado por um pequeno subconjunto dos dados de treinamento, e um vetor retirado do espaço de entrada. Dependendo de como esse núcleo de produto interno é gerado, pode-se construir diferentes máquinas de aprendizagem, como: a) Máquinas de aprendizagem polinomial, b) Redes de função de base radial (RBF), e c) Perceptrons de pelo menos uma camada intermediária (HAYKIN, 2001). As funções e algumas características dos tipos de MVS são encontradas na Tabela 2.

Tabela 2 - Resumo dos núcleos de produto interno.

Tipo de MVS	Núcleo de produto interno	Comentário
Máquina de aprendizagem polinomial	$(X^T X_i)^p$	A potência p deve ser especificada pelo usuário
Rede de função de base radial	$\exp\left(-\frac{1}{2\sigma^2} \ X-X_i\ ^2\right)$	A amplitude σ^2 é especificada pelo usuário
Perceptron de duas camadas	$\tanh(\beta_0 X^T X_i + \beta_1)$	O teorema de Mercer é satisfeito apenas para alguns valores de β_0 e β_1

Fonte: Haykin (2001).

Em que: X_i = vetor de suporte; X = vetor retirado do espaço de entrada; \tanh = tangente hiperbólica.

Segundo Haykin (2001), independentemente da máquina de aprendizagem é utilizado o algoritmo de treinamento por vetor de suporte para implementar o processo de aprendizagem. Contudo, existem variações do algoritmo de treinamento de uma MVS quando este é utilizado para classificação ou regressão não linear. Diferentemente das redes MLP e RBF que é utilizado um algoritmo específico (retropropagação). O algoritmo de retropropagação minimiza uma função de erro quadrático, independentemente da tarefa de aprendizagem.

Quando a tarefa for classificação, o algoritmo minimiza o número de amostras de treinamento que se encontram dentro da margem de separação entre exemplos de uma classe e de outra. Esse critério é mais apropriado do que o critério do erro médio quadrático utilizado pelo algoritmo de retropropagação (HAYKIN, 2001).

Por outro lado, quando se trata de uma tarefa de regressão não linear, o algoritmo minimiza uma função de perda (erro) insensível, que é uma extensão do critério do erro médio absoluto da teoria *minimax* proposta por Huber em 1964 e 1981. Isso se dá pela necessidade de um estimador robusto e que seja

insensível a pequenas variações no modelo. Assim, um estimador de mínimos quadrados torna-se inadequado pela sua sensibilidade à presença de pontos extremos (*outliers*), necessitando, portanto, de um algoritmo robusto.

O número de neurônios da camada intermediária e seus respectivos valores de peso e bias são obtidos, automaticamente, a partir de conjuntos de dados de treinamento (vetores de suporte). Assim, na MVS do tipo RBF o número de funções de base radial, seus centros e bias são obtidos automaticamente pelo número de vetores de suporte e seus valores, respectivamente. Já na MVS do tipo MLP, o número de neurônios na camada intermediária, seus vetores de peso e bias são obtidos pelo número de vetores de suporte e seus valores (HAYKIN, 2001).

Haykin (2001) afirma que nas redes PMC e RBF a complexidade do modelo é controlada mantendo-se pequeno o número de neurônios na camada escondida. Essas redes diferem-se das redes MVS, nas quais a complexidade não é controlada pela dimensionalidade, ou seja, pelo número de núcleos de produto interno (neurônios), e sim pela imposição de restrições, que resulta na extração de vetores de suporte, garantindo uma boa generalização da rede.

A alta dimensionalidade (número) de vetores de suporte extraídos dos dados de treinamento está ligada à complexidade e inexistência de conhecimento prévio do problema, bem como a falta de controle do número de dados selecionados pelo algoritmo de aprendizagem para compor os vetores de suporte. Tal problema sofre da “maldição” da dimensionalidade. Esse problema computacional é evitado usando a ideia da formação do núcleo de produto interno, definido de acordo com o teorema de Mercer (HAYKIN, 2001).

O fato da alta dimensionalidade de vetores de suporte justifica um desempenho mais lento das redes MVS em relação às redes PMC e RBF, treinadas com o algoritmo de retropropagação, para um desempenho de generalização similar.

2.6 Estudos de inteligência artificial aplicados na área florestal

Em razão das características inerentes às técnicas de inteligência artificial (IA) discutidas durante o desenvolvimento do tópico de RNAs, seu uso tem sido motivo de estudos e pesquisas e, em alguns casos, sua aplicação é preferível à regressão convencional por produzir estimativas mais precisas.

Vários trabalhos na área florestal são desenvolvidos com base em técnicas de IA para modelagem de variáveis do povoamento. Dentre os trabalhos que modelam essas variáveis, pode-se citar: estimação de volume de árvores de *Pinus brutia* Ten., *Pinus sylvestris* L., *Abies cilicica* Carr., *Cedrus libani* A. Rich., *Pinus pinaster*, *Acacia mangium*, eucalipto e *Tectona grandis* L.f. utilizando as redes PMC (DIAMANTOPOULOU, 2005; GORGENS et al., 2009, 2014; ÖZÇELIK et al., 2010), RBF (CASTELLANOS; BLANCO; PALENCIA, 2007) e MVS (CORDEIRO et al., 2015); modelagem da altura de árvores de Crimean juniper e eucalipto utilizando a rede PMC (BINOTI; BINOTI; LEITE, 2013; ÖZÇELIK et al., 2013); modelagem do afilamento de fustes de eucalipto e de vegetação nativa (NUNES; GORGENS, 2016; SCHIKOWSKI; CORTE; SANQUETTA, 2015); projeção da área basal e volume em nível de povoamento de clones de eucalipto, utilizando as redes perceptron, PMC e RBF (BINOTI et al., 2015).

Diamantopoulou (2005), utilizando a rede PMC para estimar o volume de árvores de *Pinus brutia* Ten., constatou que a RNA apresentou o melhor desempenho em comparação aos modelos de regressão não lineares testados. Segundo o autor, a superioridade da rede está relacionada à capacidade desta em superar os problemas ligados aos dados florestais, tais como: relações não lineares, distribuições não gaussianas, *outliers* e ruídos.

Resultados semelhantes foram encontrados por Özçelik et al. (2010), ao utilizar as arquiteturas *backpropagation* (BPANN) e correlação em cascata

(CCANN), para estimar o volume individual de árvores de *Pinus sylvestris* L., *Pinus brutia* Ten., *Abies cilicica* Carr. e *Cedrus libani* A. Rich., em comparação aos métodos do centroide, de funções de afilamento e tabelas de volume padrão.

Gorgens et al. (2014), ao testar a influência da arquitetura da rede PMC para estimar o volume de árvores de eucalipto e teca de cinco regiões diferentes, concluíram que para obter melhor precisão das estimativas, é necessária uma quantidade de dez neurônios na primeira camada, recomendando o uso de mais de uma camada intermediária. Gorgens et al. (2009) utilizou esses mesmos dados para comparar os resultados de volume obtidos pela rede PMC e pelo método de regressão utilizando o modelo de Schumacher e Hall. De acordo com esses autores, a presença de duas camadas intermediárias e função de ativação exponencial proporcionou as melhores estimativas. Ainda, afirmam que a modelagem em rede permite a redução da quantidade de Equações estimadas, uma vez que uma única rede consegue abordar diversos locais e espécies.

Özçelik et al. (2013) modelou a altura de árvores de Crimean juniper utilizando técnicas de RNAs, modelos não lineares de efeitos mistos e modelos generalizados. Para esses autores, a incorporação da variabilidade do povoamento nos modelos de redes e modelos mistos proporcionou uma redução de 20% da raiz quadrada do erro médio (RMSE), quando comparados ao modelo de regressão não linear simples. Uma redução de 13% foi observada quando utilizou o modelo generalizado. Além disso, atribuíram às redes os melhores resultados, pelo fato de terem melhor capacidade de generalização em comparação com outras técnicas testadas.

Binoti, Binoti e Leite (2013) modelaram a altura de árvores de eucalipto utilizando as redes PMC visando à melhoria da precisão das estimativas, redução dos custos de inventário florestal e diminuição da necessidade de aferição da altura em campo. Neste trabalho, os autores testaram dois cenários para validação da metodologia proposta. No primeiro cenário, os autores trabalharam

com dados de parcelas de apenas um material genético (clone). O objetivo era analisar o desempenho/generalização da rede quando apresentado a ela dados de parcelas de outro clone. Para o segundo cenário, utilizaram-se duas abordagens: a) treinamento de uma rede para cada clone e b) treinamento de uma única rede com dados de todos os clones, considerados como variável categórica. Os autores concluíram que os modelos foram eficientes e cumpriram com os objetivos da pesquisa.

Schikowski, Corte e Sanquetta (2015) aplicaram a rede PMC para estimar o diâmetro relativo, volume total e comercial de eucalipto e compararam os resultados obtidos com os resultados decorrentes das funções de afilamento de Hradetzky e Garay. Os autores concluíram que o desempenho das RNAs foi muito semelhante ao das funções de afilamento para a estimação do diâmetro relativo. Já para a estimação do volume total e comercial as RNAs se mostraram mais precisas e com menor dispersão dos resíduos.

As redes RBF são uma boa alternativa às redes PMC. Segundo Haykin (2001) e Silva, Spatti e Flauzino (2010), as redes RBF diferem das PMC, por utilizar a função de ativação gaussiana na camada intermediária, a qual está ligada a formação de clusters. Assim, os centros dos clusters são os centros das funções gaussianas e a distância entre os clusters fornece a largura ou variância da função gaussiana. Para Castellanos, Blanco e Palencia (2007), as vantagens das RBFs estão relacionadas à formação de clusters. Esses autores afirmam que a formação de agrupamentos com características comuns, pelos neurônios gaussianos reduz a quantidade de amostras para o treinamento da rede, garantindo um treinamento mais rápido.

Castellanos, Blanco e Palencia (2007), comparando as precisões fornecidas pela estatística erro quadrático médio (RMSE), quando aplicadas às técnicas: rede RBF, Huber e regressão múltipla para predição volumétrica de

árvores de *Pinus pinaster* e *Pinus sylvestris*, constataram a obtenção de um menor RMSE quando utilizada a rede RBF.

Cordeiro et al. (2015), analisando a estimação do volume de *Acacia mangium* aplicando RNA do tipo *perceptrons* de múltiplas camadas; MVS com função de erro tipo II (nu-SVM) e função de Kernel do tipo RBF e regressão (modelo de Schumacher e Hall), constataram uma maior precisão das estimativas quando utilizadas a RNA e a MVS.

Binoti et al. (2015), objetivando treinar, aplicar e avaliar a eficiência de RNAs na prognose da produção da área basal e volume de povoamentos de clones de eucalipto, concluíram que as redes apresentaram resultados satisfatórios, comprovando que a técnica possui potencial para ser aplicada em dados florestais. Neste trabalho, a rede do tipo PMC foi a mais adequada para projeção da área basal, apresentando as melhores estatísticas de RMSE (4,78%) e correlação (0,98). Já, para projeção volumétrica, a rede do tipo RBF foi superior, com valores de RMSE e correlação de 10,89% e 0,96, respectivamente.

REFERÊNCIAS

- BINOTI, D. H. B. et al. Modelagem da distribuição diamétrica em povoamentos de eucalipto submetidos a desbaste utilizando autômatos celulares. **Revista Árvore**, Viçosa, v. 35, n. 5, p. 125-136, set./out. 2012.
- BINOTI, M. L. M. da S. et al. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, Lavras, v. 21, n. 1, p. 97-105, jan./mar. 2015.
- BINOTI, M. L. M. da S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, Viçosa, v. 37, n. 4, p. 639-645, jul./ago. 2013.
- BOX, G.; JENKINS, G.; REINSEL, G. **Time series analysis: forecasting and control**. New Jersey: Prentice Hall, 1994. 598 p.
- BRAGA, A. P.; CARVALHO, A. P. de L. F. de; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2011. 226 p.
- BUCKMAN, R. E. **Growth and yield of red pine in Minnesota**. Washington: U.S. Department of Agriculture, 1962. (Technical Bulletin, 1272). 50 p.
- CALEGARIO, N. et al. Estimativa do crescimento de povoamentos de *Eucalyptus* baseada em modelos lineares em multiníveis de efeito misto. **Revista Árvore**, Viçosa, v. 29, n. 2, p. 251-261, mar./abr. 2005a.
- _____. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não lineares generalizados. **Scientia Forestalis**, Piracicaba, n. 69, p. 38-50, dez. 2005b.
- CAMARINHA FILHO, J. A. **Modelos lineares mistos: estruturas de matrizes de variâncias e covariâncias e seleção de modelos**. 2002. 85 p. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2002.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3. ed. Viçosa: Ed. UFV, 2009. 548 p.
- CARVALHO, S. de P. C. e et al. Modelos não lineares generalizados aplicados na predição da área basal de volume de *Eucalyptus* clonal. **Cerne**, Lavras, v. 17, n. 4, p. 541-548, out./dez. 2011.

CASTELLANOS, A.; BLANCO, A. M.; PALENCIA, V. Applications of radial basis neural networks for area forest. **International Journal Information Theories and Applications**, Piscataway, v. 14, p. 218-222, 2007.

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

CLUTTER, J. L. et al. **Timber management: a quantitative approach**. New York: Wiley & Sons, 1983. 333 p.

CORDEIRO, M. A. et al. Estimativa do volume de *Acacia mangium* utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. **Pesquisa Florestal Brasileira**, Colombo, v. 35, n. 83, p. 255-261, set. 2015.

CURTIS, R. O. A method of estimation of gross yield of Douglas-fir. **Forest Science**, Bethesda, v. 13, p. 24, Mar. 1967. Supplement 1.

DAVIDIAN, M.; GILTINAN, D. M. **Nonlinear models for repeated measurement data**. New York: Chapman & Hall, 1995. 360 p.

DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and Electronics in Agriculture**, New York, v. 48, n. 3, p. 235-244, Sept. 2005.

_____. Assessing a reliable modeling approach of features of trees through neural network models for sustainable forests. **Sustainable Computing: informatics and systems**, Washington, v. 2, n. 4, p. 190-197, Dec. 2012.

DIAMANTOPOULOU, M. J. et al. Artificial neural network modeling for reforestation design through the dominant trees bole-volume estimation. **Natural Resource Modeling**, Tempe, v. 22, n. 4, p. 511-543, Nov. 2009.

DIGGLE, P. J. et al. **Analysis of longitudinal data**. 2nd ed. Oxford: Oxford University Press, 2001. 379 p.

FAUSTO, M. A. et al. O modelo de regressão linear misto para dados longitudinais: uma aplicação na análise de dados antropométricos desbalanceados. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 24, n. 3, p. 513-524, mar. 2008.

GEMAN, S.; BIENENSTOCK, E.; DOURSAT, R. Neural networks and the bias/variance dilemma. **Neural Computation**, Massachusetts, v. 4, n. 1, p. 1-58, Jan. 1992.

GORGENS, E. B. et al. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, Viçosa, v. 33, n. 6, p. 1141-1147, nov./dez. 2009.

_____. Influência da arquitetura na estimativa de volume de árvores individuais por meio de redes neurais artificiais. **Revista Árvore**, Viçosa, v. 38, n. 2, p. 289-295, mar./abr. 2014.

GREGOIRE, T. G.; SCHABENBERGER, O. Nonlinear mixed-effects modeling of cumulative bole volume with spatially correlated within-tree data. **Journal of Agricultural, Biological, and Environmental Statistics**, Alexandria, v. 1, n. 1, p. 107-119, Mar. 1996.

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

HUBER, P. J. Robust estimation of a location parameter. **Annals of Mathematical Statistics**, Ann Arbor, v. 35, p. 73-101, 1964.

_____. **Robust statistics**. New York: Wiley & Sons, 1981. 380 p.

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

LEDUC, D. J. et al. **Predicting diameter distributions of longleaf pine plantations: a comparison between artificial neural networks and other accepted methodologies**. Asheville, NC: U.S. Department of Agriculture, 2001. 24 p.

LEITE, H. G. et al. Estimation of inside-bark diameter and heartwood diameter for *Tectona grandis* Linn. trees using artificial neural networks. **European Journal of Forest Research**, Heidelberg, v. 130, n. 2, p. 263-269, Mar. 2011.

LEK, S. et al. Application of neural networks to modelling nonlinear relationships in ecology. **Ecological Modelling**, Amsterdam, v. 90, n. 1, p. 39-52, Sept. 1996.

LEK, S.; GUÉGAN, J. F. **Artificial neuronal networks: application to ecology and evolution**. Berlin: Springer, 2000. 288 p.

LITTELL, R. C. et al. **SAS for mixed models**. 2nd ed. Cary: SAS Institute, 2006. 814 p.

LORENA, A. C.; CARVALHO, A. C. P. L. F. de. **Introdução às máquinas de vetores suporte (*Supporte Vector Machines*)**. São: Carlo Instituto de Ciências Matemáticas e de Computação, 2003. 66 p.

MACKINNEY, A. L.; CHAIKEN, L. E. **Volume, yield, and growth of loblolly pine in the Mid-Atlantic Coastal Region**. Washington: USDA Forest Service, 1939. 30 p.

MCCULLOCH, C. E.; SEARLE, S. R. **Generalized, linear, and mixed models**. New York: Wiley Series in Probability and Statistics, 2001. v. 1, 325 p.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, New York, v. 5, n. 4, p. 115-133, Dec. 1943.

MEHTÄTALO, L. A longitudinal height-diameter model for Norway spruce in Finland. **Canadian Journal of Forest Research**, Ottawa, v. 34, n. 1, p. 131-140, 2004.

MEHTÄTALO, L.; MALTAMO, M.; KANGAS, A. The use of quantile trees in the prediction of the diameter distribution of a stand. **Silva Fennica**, Helsinki, v. 40, n. 3, p. 501-516, Jan. 2006.

MENDONÇA, A. R. de et al. Modelos hipsométricos e de crescimento em altura das árvores dominantes e codominantes para *Pinus caribaea* var. *hondurensis*. **Scientia Forestalis**, Piracicaba, v. 39, n. 90, p. 151-160, jun. 2011.

MENDONÇA, A. R. de; CARVALHO, S. de P. C. e; CALEGARIO, N. Modelos hipsométricos generalizados mistos na predição da altura de *Eucalyptus* sp. **Cerne**, Lavras, v. 21, n. 1, p. 107-114, jan./mar. 2015.

MOLENBERGHS, G.; VERBEKE, G. A review on linear mixed models for longitudinal data, possibly subject to dropout. **Statistical Modelling**, London, v. 1, n. 4, p. 235-269, Dec. 2001.

MOSER, J. W.; HALL, O. F. Deriving growth and yield functions for uneven-aged forest stands. **Forest Science**, Bethesda, v. 15, n. 2, p. 183-188, June 1969.

NUNES, M. H.; GORGENS, E. B. Artificial intelligence procedures for tree taper estimation within a complex vegetation mosaic in Brazil. **PLoS One**, San Francisco, v. 11, n. 5, p. 16, May 2016.

ÖZÇELİK, R. et al. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, Amsterdam, v. 306, p. 52-60, Oct. 2013.

_____. Estimating tree bole volume using artificial neural network models for four species in Turkey. **Journal of Environmental Management**, London, v. 91, n. 3, p. 742-753, Jan./Feb. 2010.

PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer, 2000. 528 p.

_____. **Mixed-effects models in S and S-PLUS**. New York: Springer, 2004. 530 p.

PINHEIRO, J. C.; BATES, D. M.; PINHEIRO, J. C. Approximations to the log-likelihood function in the nonlinear mixed effects model. **Journal of Computational and Graphical Statistics**, Alexandria, v. 4, n. 1, p. 12-35, Mar. 1995.

REED, R. Pruning algorithms – a survey. **IEEE Transactions on Neural Networks**, New York, v. 4, n. 5, p. 740-747, Sept. 1993.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations by error propagation. In: McCLELLAND, J. L.; RUMELHART, D. E.; HINTON, G. E. **Parallel distributed processing**. Cambridge: MIT Press, 1986. v. 1, chap. 8, p. 318-362.

SALLES, T. T. et al. Modelo de Clutter na modelagem de crescimento e produção de eucalipto em sistemas de integração lavoura-pecuária-floresta. **Pesquisa Agropecuária Brasileira**, Brasília, v. 47, n. 2, p. 253-260, fev. 2012.

SCHIKOWSKI, A. B.; CORTE, A. P. D.; SANQUETTA, C. R. Estudo da forma do fuste utilizando redes neurais artificiais e funções de afilamento. **Pesquisa Florestal Brasileira**, Colombo, v. 35, n. 82, p. 119-127, abr./jun. 2015.

SCHUMACHER, F. X. A new growth curva and its application to timber-yield studies. **Journal of Forest**, Bethesda, v. 37, n. 4, p. 819-820, 1939.

SCHUMACHER, F. X.; COILE, T. S. **Growth and yield of natural stands of the southern pines**. Durham: T.S. Coile, 1960. 115 p.

SHEINER, L.; BEAL, S. L. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis–Mentem model: routine clinical pharmacokinetic data. **Journal of Pharmacokinetics and Biopharmaceutics**, New York, v. 8, n. 6, p. 553-571, Dec. 1980.

SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais: para engenharia e ciências aplicadas**. São Paulo: Arliber, 2010. 399 p.

SMOLA, A. J. et al. Introduction to large margin classifiers. In: _____. (Ed.). **Advances in large margin classifiers**. Cambridge: MIT Press, 1999. p. 1-28.

SOARES, F. A. A. M. N. et al. Recursive diameter prediction and volume calculation of eucalyptus trees using Multilayer Perceptron Networks. **Computers and Electronics in Agriculture**, New York, v. 78, n. 1, p. 19-27, Aug. 2011.

_____. Recursive diameter prediction for calculating merchantable volume of eucalyptus clones using Multilayer Perceptron. **Neural Computing & Applications**, London, v. 22, n. 7/8, p. 1407-1418, June 2013.

SULLIVAN, A. D.; CLUTTER, J. L. A simultaneous growth and yield model for loblolly pine. **Forest Science**, Bethesda, v. 18, n. 1, p. 76-86, Mar. 1972.

TEMESGEN, H.; MONLEON, V. J.; HANN, D. W. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. **Canadian Journal of Forest Research**, Ottawa, v. 38, n. 3, p. 553-565, 2008.

TRINCADO, G.; BURKHART, H. E. A generalized approach for modeling and localizing stem profile curves. **Forest Science**, Bethesda, v. 52, n. 6, p. 670-682, Dec. 2006.

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

VONESH, E. F.; CHINCHILLI, V. M. **Linear and nonlinear models for the analysis of repeated measurements**. New York: Marcel Dekker, 1997. 560 p.

YANG, Y. et al. A multilevel individual tree basal area increment model for aspen in boreal mixedwood stands. **Canadian Journal of Forest Research**, Ottawa, v. 39, n. 11, p. 2203-2214, 2009.

YANG, Y.; HUANG, S. Comparison of different methods for fitting nonlinear mixed forest models and for making predictions. **Canadian Journal of Forest Research**, Ottawa, v. 41, n. 8, p. 1671-1686, Aug. 2011.

ZHAO, D.; WILSON, M.; BORDER, B. E. Modeling response curves and testing treatment effects in repeated measures experiments: a multilevel nonlinear mixed-effects model approach. **Canadian Journal of Forest Research**, Ottawa, v. 35, p. 122-132, 2005.

ZHENG, M. Fitting linear and nonlinear growth curve models using PROC NLMIXED. In: SAS GLOBAL FORUM, 2010, Seattle. **Proceedings...** Washington: SAS, 2010. p. 260-2010.

SEGUNDA PARTE - ARTIGOS

**ARTIGO 1 - MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE
ÁRVORES CLONAIIS DE EUCALIPTO UTILIZANDO MODELOS NÃO
LINEARES MISTOS**

MODELING OF EUCALYPTUS CLONES TREE HEIGHT GROWTH USING
NON-LINEAR MODELS OF MIXED EFFECTS

**Artigo redigido conforme a NBR 6022 (ABNT, 2003) e formatado de acordo
com o Manual da UFPA de apresentação de teses e dissertações.**

RESUMO

Para obter o estoque de madeira de um povoamento florestal, é necessário o conhecimento da altura de árvores, pois essa variável é correlacionada com o volume. Com isso, objetivou-se, neste trabalho, analisar a estimação da altura de árvores de eucalipto, utilizando modelos não lineares de efeitos mistos com a adição de covariantes e modelagem da heterocedasticidade. Foram utilizados dados de 90 unidades amostrais, compreendendo parcelas temporárias e permanentes medidas em 2010 e 2012, distribuídas em diferentes idades e locais. Foram testados os modelos: Logístico, Chapman-Richards, Weibull e Gompertz. Para aplicação dos modelos mistos, a unidade amostral foi considerada como efeito aleatório. A utilização dos modelos não lineares mistos possibilitou uma redução do erro padrão em, aproximadamente, 60%. Todos os modelos testados apresentaram comportamentos e estatísticas adequadas para a predição de alturas de árvores individuais. Contudo, para fins de adição de covariantes e de modelagem da heterogeneidade da variância foi utilizado apenas o modelo de Gompertz. Procedeu-se com a decomposição dos parâmetros do modelo de Gompertz para a inclusão das covariantes: clone, área basal, sítio e idade. A adição de variáveis ao modelo misto não possibilitou uma melhoria de acurácia da estimativa, obtendo o mesmo valor de erro padrão residual - 1,22m. A aplicação da estrutura de variância “*varPower*” para modelagem da heterocedasticidade possibilitou uma redução na estimativa do erro padrão residual e percentual, quando comparado ao modelo de efeitos mistos com inclusão de covariantes de, aproximadamente, 55% (de 1,22m para 0,55m). Para o modelo de Gompertz misto heterocedástico, foi verificada uma melhor distribuição de resíduos, ao comparar com o gráfico de resíduos do modelo de Gompertz original. Além disso, o problema de heterogeneidade da variância verificado no modelo original foi solucionado pela modelagem da estrutura de variância.

Palavras-chave: Regressão. Efeitos aleatórios. Covariável. Heterocedasticidade.

**ARTICLE 1 - MODELING OF THE HYSOMETRIC RELATIONSHIP
OF CLONAL EUCALYPTUS TREES USING NONLINEAR MIXED
MODELS**

ABSTRACT

It is necessary to know the height of the trees to obtain the wood stock from a forest stand since this variable is correlated with the volume. Therefore, the objective of this study was to analyze the estimation of the height of eucalyptus trees using mixed effect nonlinear models with the addition of covariates and heteroscedasticity modeling. We used the data from 90 sample units, comprising temporary and permanent plots measured in 2010 and 2012, and distributed at different ages and sites. We tested the Logistic, Chapman-Richards, Weibull, and Gompertz models. We considered one sampling unit as a random effect for the application of the mixed models. The use of the nonlinear mixed models allowed us to reduce the standard error by approximately 60%. All the tested models presented adequate behavior and statistics to predict individual tree heights. However, we used only the Gompertz model to add covariates and model the heterogeneity variance. We proceeded with the decomposition of the Gompertz model parameters to include the clone, basal area, site, and age covariates. The addition of variables to the mixed model hindered the improvement in the estimation accuracy, obtaining the same residual standard error value - 1.22m. The application of the *varPower* variance structure to model the heteroscedasticity allowed us to reduce the estimation of the residual standard error and percentage when compared to the mixed effect model with the inclusion of covariates, of approximately 55% (from 1.22m to 0.55m). We verified a better residue distribution for the mixed heteroscedastic Gompertz model when comparing it to the residue graph of the original Gompertz model. Furthermore, the variance heterogeneity problem verified in the original model was solved by modeling the variance structure.

Keywords: Regression. Random effects. Covariable. Heteroscedasticity.

1 INTRODUÇÃO

A medição de alturas de árvores para estimar a produção é um procedimento oneroso, podendo ocorrer erros e, por isso, se faz necessária a mensuração de parte das alturas nas unidades amostrais. Como o objetivo do manejador é conhecer a produção volumétrica e esta é correlacionada com a altura, modelos de regressão são utilizados para obter sua estimativa. Nesse contexto, é justificável a aplicação de conhecimentos estatísticos visando a uma melhor acurácia das estimativas de alturas.

Conforme relatado por Calegario et al. (2005b), existem dificuldades na construção de modelos genéricos que conseguem captar toda a variabilidade do povoamento, em razão da base de dados ser proveniente de vários projetos, com diferenças de material genético, qualidade de local, densidade e idade. Para solucionar esse problema, alguns pesquisadores utilizam a técnica de adição de variáveis que são específicas de cada povoamento. Trabalhos que abordam a decomposição dos parâmetros para a inclusão de covariantes que explicam a variabilidade da variável resposta são frequentes, podendo citar: modelagem da relação hipsométrica de *Toona ciliata* (ALVES et al., 2017), *Pinus caribaea* var. *hondurensis* (MENDONÇA et al., 2011), *Picea abies* (L) Karst. (MEHTÄTALO, 2004), *Populus tremuloides* Michx. (YANG; HUANG, 2011) e eucalipto (CALEGARIO et al., 2005b; MENDONÇA; CARVALHO; CALEGARIO, 2015) e predição da área basal e volume (CARVALHO et al., 2011).

Essa variabilidade do povoamento causa problemas na análise dos dados, em razão da violação da pressuposição de homogeneidade da variância, provocando distorções nas estimativas paramétricas. Uma alternativa para a modelagem da heterocedasticidade é a utilização de modelos lineares ou não

lineares generalizados mistos (McCULLOCH; SEARLE, 2001; PINHEIRO; BATES, 2000).

Os modelos mistos são uma importante ferramenta na modelagem de dados florestais por obterem melhores estimativas e por sua capacidade de lidar de maneira adequada com os erros heterocedásticos, a partir da modelagem de estruturas de variâncias (GREGORIE; SCHABENBERGER, 1996; McCULLOCH; SEARLE, 2001; TRINCADO; BURKHART, 2006; YANG; HUANG, 2011; ZHAO; WILSON; BORDER, 2005).

Na área florestal, várias estruturas de variância foram utilizadas, destacando-se: modelagem da altura de árvores e área basal de povoamentos de eucalipto, utilizando a função de variância exponencial (CALEGARIO et al., 2005a, 2005b) e modelagem da altura de *Picea abies* (L) Karst. e eucalipto, utilizando a função varPower de Pinheiro e Bates (2000) (MEHTÄTALO, 2004; MENDONÇA; CARVALHO; CALEGARIO, 2015).

Nesse contexto, conduziu-se o presente estudo, com o objetivo de analisar a modelagem da relação hipsométrica de árvores de eucalipto, empregando modelos não lineares mistos, visando a uma melhoria da qualidade de ajuste tanto pela adição de covariantes quanto pela modelagem da heterogeneidade de variância.

2 MATERIAL E MÉTODOS

2.1 Base de dados

Os dados para a realização deste estudo são provenientes de inventário florestal contínuo (IFC) de plantios clonais de eucalipto, localizados na região sudeste do estado de Goiás, Brasil. O clima da região enquadra-se no tipo Aw que, de acordo com a classificação climática de Köppen, apresenta dois períodos distintos: um chuvoso e outro seco. A temperatura média anual é de 22°C, apresentando precipitação média anual entre 1200 e 1800 mm.

A área total é de 1.000 hectares. O método de amostragem foi a casual simples (ACS), com parcelas de área variável de, aproximadamente, 450m². No total, foram amostradas 90 unidades amostrais, compreendendo parcelas temporárias e permanentes medidas em 2010 e 2012, distribuídas em diferentes idades e locais. A análise descritiva das variáveis do povoamento é apresentada na Tabela 1.

Tabela 3 - Estatísticas descritivas relacionadas às variáveis do povoamento.

Variável do Povoamento	Mínimo	Máximo	Média	CV (%)
I (anos)	1,70	6,90	-	25,63
N (N ha ⁻¹)	891	1.612	1.189	7,15
S (m)	27,00	39,00	33,86	6,46
\overline{Dq} (cm)	10,16	18,53	15,50	8,95
G (m ² ha ⁻¹)	10,18	34,74	22,58	17,37

Fonte: Do autor (2018).

Em que: I = idade (anos); N = número de árvores por hectare; S = índice de sítio; \overline{Dq} = diâmetro quadrático; G = área basal; CV = coeficiente de variação.

2.2 Modelos analisados

Para modelagem da relação hipsométrica foram testados os modelos não lineares Logístico, Chapman-Richards, Weibull e Gompertz (Tabela 2).

Tabela 2 - Modelos não lineares para estimar a altura total de árvores de eucalipto.

Nome	Modelo
Logístico	$H_i = \frac{\beta_0}{1 + \exp\left[\frac{(\beta_1 - DAP_i)}{\beta_2}\right]} + \varepsilon_i$
Chapman-Richards	$H_i = \beta_0 \left[1 - \exp(-\beta_1 DAP_i)\right]^{\beta_2} + \varepsilon_i$
Weibull	$H_i = \beta_0 \left[1 - \exp(-\beta_1 DAP_i^{\beta_2})\right] + \varepsilon_i$
Gompertz	$H_i = \beta_0 \exp\left[-\beta_1 \exp(-\beta_2 DAP_i)\right] + \varepsilon_i$

Fonte: Do autor (2018).

Em que: H_i = altura total da *i-ésima* árvore (m); DAP_i = diâmetro a aproximadamente 1,30 m de altura (cm) da *i-ésima* árvore presente em cada unidade amostral; exp = base do logaritmo neperiano; β_s = parâmetros a serem estimados; ε_i = erro aleatório.

2.3 Modelagem mista

O modelo misto não linear é expresso pela seguinte forma:

$$y_i = f(x_i, \beta, u_i) + \varepsilon_i \quad (1)$$

Em que: y_i = vetor da variável altura do *i-ésimo* grupo de árvores (m) ($n \times 1$); f (.) = função não linear; x_i = matriz da variável DAP (cm) do *i-ésimo* grupo de árvores com dimensão ($n \times p+1$), onde p é o número de parâmetros de efeitos fixos; β = vetor de parâmetros de efeitos fixos ($p+1 \times 1$); u_i = vetor de parâmetros aleatórios associado ao *i-ésimo* grupo de árvores ($q \times 1$); ε_i = vetor de erros aleatórios ($n \times 1$).

Os parâmetros aleatórios u_i e ε_i não são correlacionados e apresentam distribuição normal com média zero e matrizes de variâncias-covariâncias \mathbf{D} e \mathbf{R}_i , respectivamente.

$$\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{D}) \text{ e } \boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \mathbf{R}_i) \quad (2)$$

Em que: $\mathbf{D}_{(q \times q)}$ é a matriz de variâncias-covariâncias dos efeitos aleatórios e $\mathbf{R}_{(n \times n)}$ é a matriz de variâncias-covariâncias residuais.

Para estimação paramétrica dos modelos de altura, será utilizado o método da máxima verossimilhança (MV). Segundo Pinheiro e Bates (2000) e Yang e Huang (2011), as estimativas obtidas pela MV são desejáveis estatisticamente, em razão da sua capacidade de produzir estimativas assintoticamente eficientes e seguir distribuições normais assintóticas. Além disso, seu uso possibilita a comparação dos valores da verossimilhança entre modelos, pelos critérios de informação de Akaike (CIA) e Bayesiano (CIB), por estes serem baseados na estatística da MV.

As estimativas são obtidas pela maximização da função de densidade probabilidade conjunta, após os efeitos aleatórios na verossimilhança condicional serem integrados, conforme expressão (3).

$$p(y | \beta, D) = \prod_{i=1}^m \int p(y_i | u_i, \beta, D) p(u_i, D) du_i \quad (3)$$

Em que: $p(y_i | u_i, \beta, D)$ é a densidade condicional de y_i dado u_i , β e D ; $p(u_i, D)$ é a densidade marginal de u_i .

Os parâmetros aleatórios dos modelos mistos serão preditos, usando o melhor preditor linear não viesado - BLUP (McCULLOCH; SEARLE, 2001), que foi aplicado, por exemplo, por Lappi (1991, 1997) e Mehtätalo (2004). Essa técnica está implementada computacionalmente na biblioteca nlme do *software* R.

Nessa implementação para predição dos parâmetros aleatórios, estimam-se as matrizes D e R . A estimativa de D é construída a partir dos desvios padrões estimados e correlações dos parâmetros aleatórios. As variâncias dos resíduos são calculadas, usando algum modelo de variância, nesse caso será utilizada a função de variância “varPower” (PINHEIRO; BATES, 2000). Essas variâncias são colocadas na diagonal da matriz \hat{R} . Como os resíduos são não correlacionados, os elementos não diagonais de \hat{R} são zeros. O melhor preditor linear não viesado dos parâmetros aleatórios é calculado pela expressão (4).

$$\hat{u} = (Z' R^{-1} Z + D^{-1})^{-1} Z' R^{-1} (y - \mu) \quad (4)$$

Quando os parâmetros aleatórios são calculados, as estimativas das matrizes de variância-covariância \hat{R} e \hat{D} são substituídas pelas matrizes R e D e a predição da parte fixa é substituída pelo vetor μ (LAPPI, 1997; MEHTÄTALO, 2004).

Como a relação H-DAP dos dados é não longitudinal e, portanto, não são correlacionados, unidades amostrais foram consideradas como efeitos aleatórios.

2.4 Adição de covariáveis aos modelos

A adição de covariáveis aos modelos de altura baseia-se na decomposição dos parâmetros dos modelos e inclusão de variáveis que explicam a variabilidade da variável resposta. Isso se dá pela ineficiência de apenas uma variável, DAP, explicar a variação da altura total de árvores individuais. Assim, aos modelos serão associadas variáveis relativas à idade, ao clone, à G e ao sítio. Dentre os modelos testados, essa técnica será aplicada apenas ao que produzir as melhores estatísticas de acurácia.

2.5 Heterogeneidade da variância

Para a modelagem da heterocedasticidade, foi utilizada a função de variância denominada “varPower”. Segundo Pinheiro e Bates (2000), o modelo de variância é expresso pela Equação (5).

$$\text{Var}(\varepsilon_{ij}) = \sigma^2 |v_{ij}|^{2\delta} \quad (5)$$

O modelo de variância apresentado em (5) corresponde a função de variância expressa pela expressão (6).

$$g(v_{ij}, \delta) = |v_{ij}|^{\delta} \quad (6)$$

* v_{ij} deve ser diferente de 0.

A função de variância é uma potência do valor absoluto da covariável (v_{ij}). O coeficiente da função de variância é representado pelo parâmetro δ , o qual é considerado irrestrito, ou seja, pode assumir qualquer valor real e, portanto, modela casos onde a variância aumenta ou diminui com o valor absoluto da variância da covariável.

2.6 Análise dos modelos

Para a avaliação dos modelos de regressão será utilizada a análise gráfica dos resíduos, a fim de se verificar as suposições básicas de análise de regressão, juntamente com algumas medidas de acurácia, que servirá de subsídio na comparação dos modelos. Os modelos ajustados terão sua acurácia avaliada,

por meio do erro padrão residual (S_{yx}) e percentual [$(S_{yx} (\%))$], critério de informação de Akaike (CIA) (SAKAMOTO; ISHIGURO; KITAGAWA, 1986), Bayesiano (CIB) (SCHWARZ, 1978) e, pelo teste da razão da máxima verossimilhança (TRMV).

O erro padrão residual e percentual é dado pelas Equações (7) e (8), respectivamente.

$$S_{yx} = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n - p}} \quad (7) \quad S_{yx} (\%) = \frac{S_{yx}}{\bar{Y}} 100 \quad (8)$$

Em que: S_{yx} = erro padrão residual (m); $S_{yx} (\%)$ = erro padrão relativo (%); Y = valor observado da variável dependente; \hat{Y} = valor estimado da variável dependente pelo modelo; n = número de observações; p = número de parâmetros.

Os critérios de informação de Akaike e Bayesiano são representados pelas Equações (9) e (10), respectivamente.

$$CIA = -2 \ln(mv) + 2p \quad (9) \quad CIB = -2 \ln(mv) + p \ln(n) \quad (10)$$

Em que: mv = valor da máxima verossimilhança; \ln = logaritmo neperiano; p = número de parâmetros do modelo; n = número de observações.

Em relação aos critérios, nota-se que, quanto menores os valores de CIA e CIB melhor o modelo ajustado, pois penalizam modelos com número excessivo de parâmetros, selecionando modelos mais parcimoniosos.

O teste TRMV é uma ferramenta de comparação de modelos aninhados dado pela diferença entre os valores máximos da função de verossimilhança dos modelos (EQUAÇÃO 11). Esse valor é testado pela estatística χ^2 . Se o valor de TRMV for significativo ao nível α de probabilidade, o modelo com maior valor de máxima verossimilhança apresenta maior precisão.

$$\text{TRMV} = 2[\ln(mv_2) - \ln(mv_1)] \quad (11)$$

Em que: mv = valor da máxima verossimilhança; \ln = logaritmo neperiano.

3 RESULTADOS E DISCUSSÃO

3.1 Análise dos modelos hipsométricos

Na Tabela 3, são apresentadas as estatísticas de ajuste dos modelos de altura e seus respectivos parâmetros estimados.

Tabela 3 - Estimativa dos parâmetros e estatísticas de ajuste dos modelos de relação hipsométrica.

Logístico ($S_{yx} = 3,07m - 11,94\%$)				
Parâmetro	Estimativa	Erro Padrão	t-valor	p-valor
β_0	33,84713	0,436	77,691	< 0,0001
β_1	10,00832	0,142	70,643	< 0,0001
β_2	5,11851	0,203	25,227	< 0,0001
Chapman-Richards ($S_{yx} = 3,11 - 12,07\%$)				
β_0	37,329612	0,961	38,850	< 0,0001
β_1	0,104574	0,009	12,020	< 0,0001
β_2	1,782399	0,153	11,680	< 0,0001
Weibull ($S_{yx} = 3,10 - 12,04\%$)				
β_0	34,928925	0,768458	45,453	< 0,0001
β_1	0,022251	0,002682	8,295	< 0,0001
β_2	1,479157	0,058414	25,322	< 0,0001
Gompertz ($S_{yx} = 3,09 - 12,01\%$)				
β_0	35,807051	0,628124	57,010	< 0,0001
β_1	2,834652	0,153553	18,460	< 0,0001
β_2	0,134730	0,006614	20,370	< 0,0001

Fonte: Do autor (2018).

Na Tabela 3, mostra-se que todos os parâmetros dos modelos testados foram significativos ($p < 0,05$). Considerando apenas esse critério, todos os

quatro modelos poderiam ser utilizados para se estimar a altura de árvores individuais de eucalipto. Em relação ao erro padrão de cada parâmetro, apesar de o modelo logístico não apresentar os menores valores em todos os seus parâmetros quando comparados aos demais, apresentou os maiores valores de t , motivo por que pode ser considerado como levemente superior. Fato este é confirmado quando se leva em consideração os critérios: erro padrão residual e percentual, CIA, CIB e $\ln mv$. Na Tabela 3, observa-se que o modelo logístico apresentou o menor valor de erro padrão (3,07m), seguido dos modelos de Gompertz (3,09m), Weibull (3,10m) e Chapman e Richards (3,11m). Na Tabela 4, nota-se que o modelo logístico apresentou os menores valores de CIA e CIB e o maior valor da máxima verossimilhança em relação aos outros modelos. Contudo, uma leve superioridade é atribuída ao modelo logístico, porém, nessa primeira análise, não há restrições quanto à utilização dos modelos de Chapman e Richards, Weibull e Gompertz.

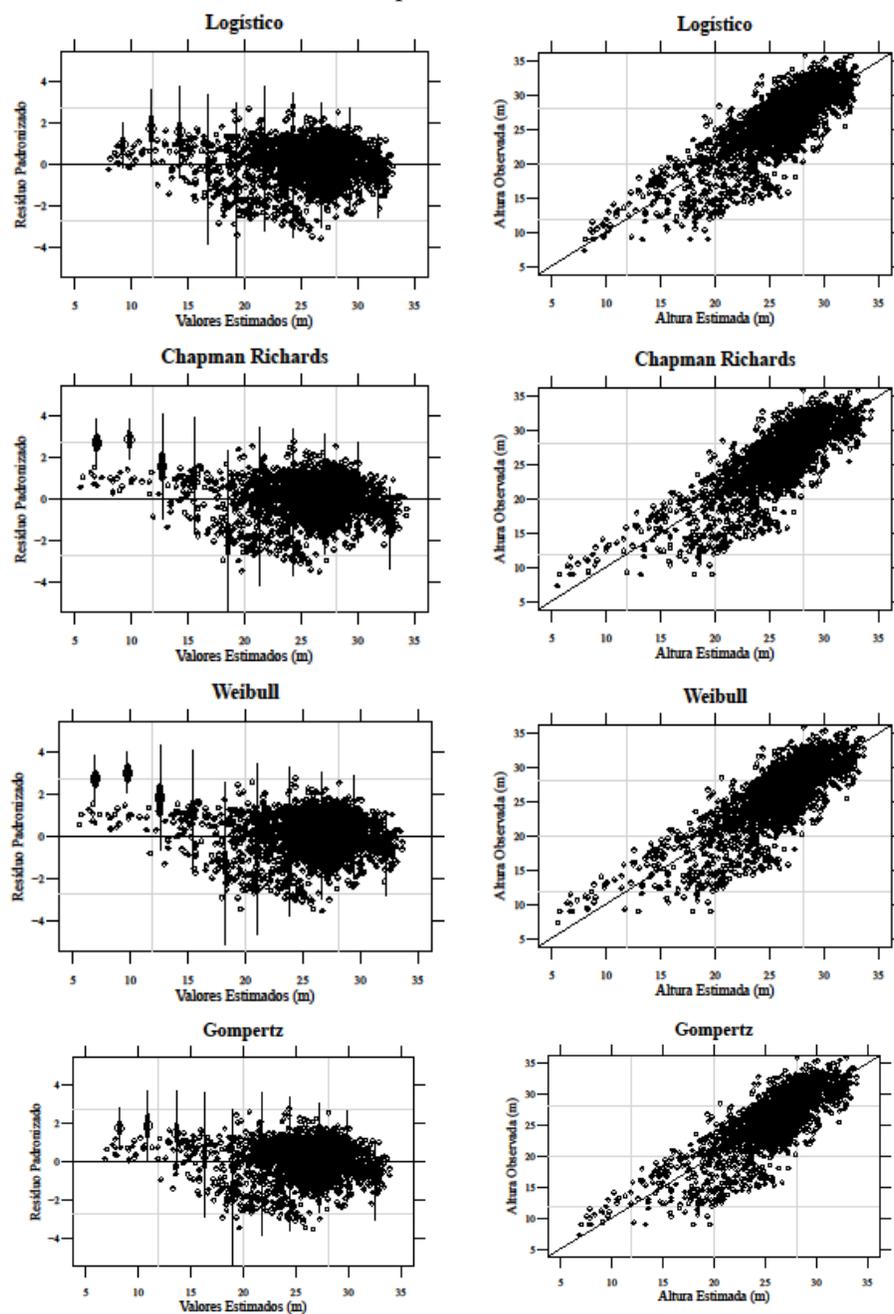
Tabela 4 - Critério de informação de Akaike (CIA), critério de informação Bayesiano (CIB), e logaritmo neperiano da máxima verossimilhança ($\ln mv$), para os modelos hipsométricos ajustados.

Modelo	CIA	CIB	$\ln mv$
Logístico	10.907,63	10.930,31	-5.449,813
Chapman e Richards	10.953,94	10.976,63	-5.472,971
Weibull	10.942,47	10.965,16	-5.467,237
Gompertz	10.932,19	10.954,80	-5.462,095

Fonte: Do autor (2018).

Na Figura 1, são apresentados os gráficos dos resíduos padronizados e a relação entre a altura total estimada e a altura total observada para os quatro modelos testados, como forma de verificar se as suposições da análise de regressão são atendidas.

Figura 1 - Gráficos de resíduos padronizados e da relação entre os valores observados e estimados para os modelos testados.



Fonte: Do autor (2018).

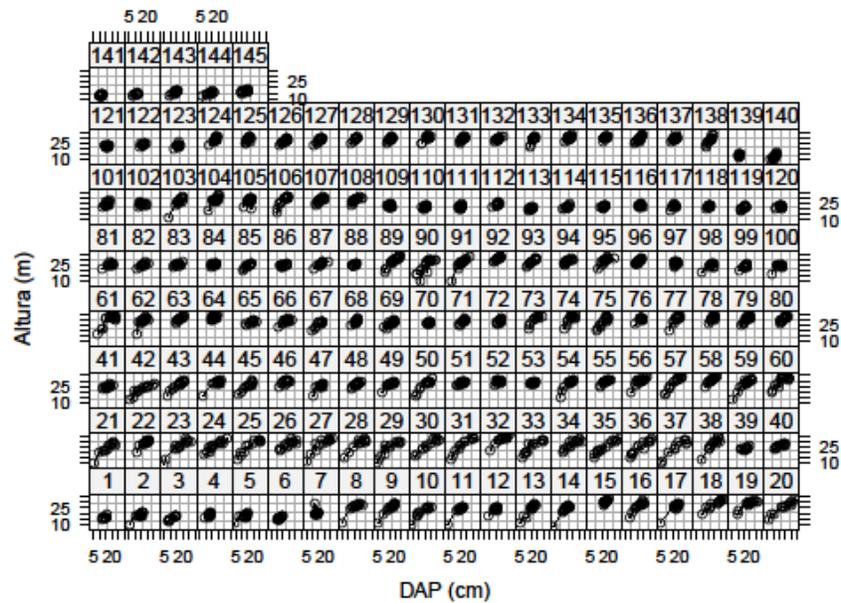
Ao se verificar os gráficos de resíduos, nota-se uma semelhança entre a qualidade de ajuste dos modelos. Uma tendência de superestimação é verificada nas menores árvores. Antes de se penalizar o uso de todos os modelos testados pela disposição dos intervalos de confiança para a média (linhas verticais grossas), a superestimação é em decorrência do pequeno número de indivíduos amostrados com altura menor do que 15m, e não dos modelos ajustados.

Além disso, é verificado um problema de heterogeneidade da variância, uma vez que se observa um comportamento em forma de funil dos resíduos. Tal problema pode ser verificado também pelos desvios-padrão específicos das classes (linhas verticais finas). A presença da heterocedasticidade é comum em base de dados florestais, em razão da variação na dispersão de uma variável em relação a outra.

3.2 Análise dos modelos mistos

Para aplicação da teoria de efeitos mistos, a unidade amostral foi considerada como efeito aleatório, conforme Figura 2.

Figura 2 - Relação das alturas e diâmetros separados por unidades amostrais.



Fonte: Do autor (2018).

Os modelos mistos utilizando a função “nlme” foram ajustados. Consideraram-se, como efeito aleatório, os parâmetros β_0 e β_2 para os modelos Logístico e Weibull, β_0 e β_1 para o modelo de Gompertz e o parâmetro β_0 para o modelo de Chapman e Richards. Para todos os modelos testados, a unidade amostral foi considerada aleatória. Na Tabela 5, são apresentadas as estimativas paramétricas e a correlação entre os parâmetros aleatórios.

Tabela 5 - Estimativa dos parâmetros e da correlação para os modelos hipsométricos testados utilizando os modelos não lineares mistos.

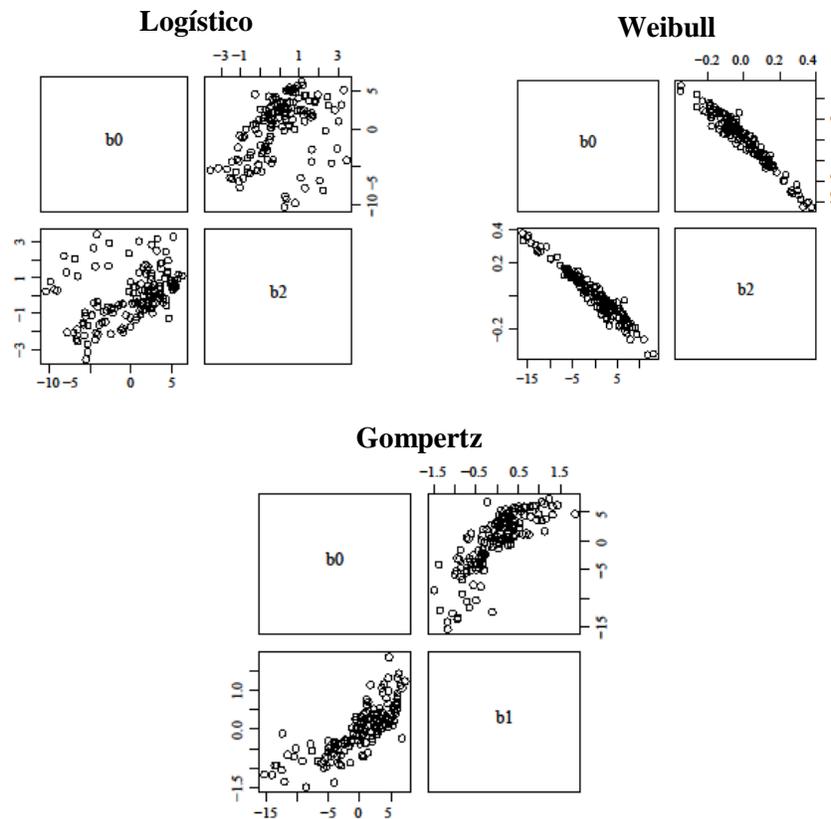
Logístico 3 Parâmetros ($S_{yx} = 1,27m - 4,94\%$)							
Parâmetro	Estimativa	Erro Padrão	t-valor	p-valor	Correlação Aleatório		
					β_0	β_2	
β_0	29,435696	0,373	78,871	< 0,0001	X	X	
β_1	8,272134	0,081	101,861	< 0,0001	X	X	
β_2	3,954391	0,169	23,350	< 0,0001	0,46	X	
Chapman-Richards ($S_{yx} = 1,35m - 5,27\%$)						β_0	
β_0	32,35901	0,481	67,277	< 0,0001	X		
β_1	0,12040	0,006	21,235	< 0,0001	X		
β_2	1,48310	0,078	18,958	< 0,0001	X		
Weibull ($S_{yx} = 1,23m - 4,79\%$)						β_0	β_2
β_0	29,401853	0,545	53,910	< 0,0001	X	X	
β_1	0,031264	0,002	16,301	< 0,0001	X	X	
β_2	1,559266	0,030	51,865	< 0,0001	-0,95	X	
Gompertz ($S_{yx} = 1,22m - 4,73\%$)						β_0	β_1
β_0	30,687925	0,460	66,739	< 0,0001	X	X	
β_1	2,221878	0,100	22,182	< 0,0001	0,76	X	
β_2	0,157322	0,004	42,138	< 0,0001	X	X	

Fonte: Do autor (2018).

Na Tabela 5, mostra-se que todos os modelos testados tiveram significância em seus parâmetros ($p < 0,05$). A utilização da teoria dos modelos não lineares de efeitos mistos com a correta associação dos parâmetros aos efeitos aleatórios possibilitou uma redução do erro padrão em aproximadamente 60%. Esse fato ocorre em decorrência do ajuste “nlme” captar a variabilidade entre as unidades amostrais e, com isso, produzir estimativas de erro padrão menores, e pelos baixos valores de correlação para os modelos Logístico (46%) e de Gompertz (76%) observados na Tabela 5 e na análise gráfica da Figura 3.

A superioridade dos modelos Logístico e de Gompertz pode ser comprovada pelos baixos valores de correlação entre os parâmetros (aleatórios), indicando que tais modelos são parcimoniosos, ou seja, não estão sobreparametrizados e, por produzir boas estimativas de erro padrão - 1,27 e 1,22m, respectivamente.

Figura 3 - Gráfico de correlação para as estimativas paramétricas de efeitos aleatórios.



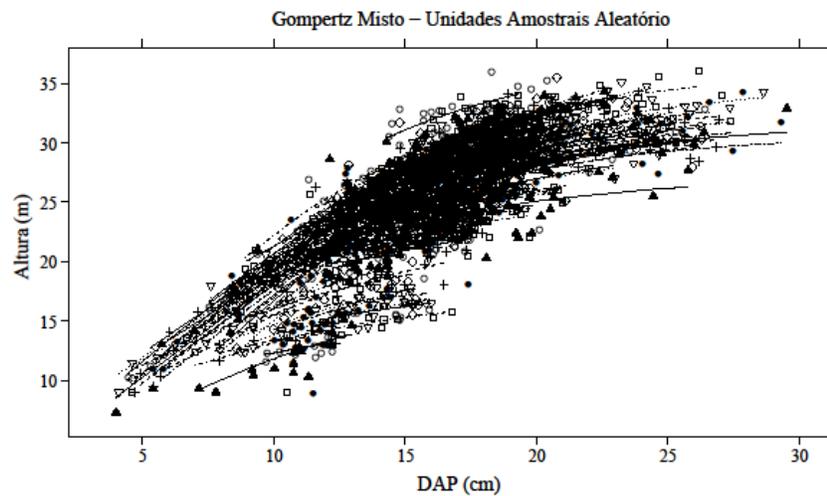
Fonte: Do autor (2018).

No anexo A (Ajuste das curvas de regressão para os quatros modelos de efeitos mistos testados), é observado um comportamento similar nos ajustes dos

modelos, reafirmando a boa escolha de parâmetros como efeitos aleatórios. Todos os modelos testados apresentaram comportamentos e estatísticas adequadas para a predição de alturas de árvores individuais. Contudo, para fins de adição de covariantes e predição será utilizado apenas o modelo de Gompertz.

Na Figura 4, são apresentadas as curvas de regressão do modelo de Gompertz, considerando aleatórios os parâmetros β_0 e β_1 e a unidade amostral.

Figura 4 - Ajuste das curvas de regressão para o modelo de efeitos misto de Gompertz.



Fonte: Do autor (2018).

3.3 Inclusão de covariantes no modelo

Como a variável altura não é apenas explicada pelo diâmetro, procedeu-se à decomposição dos parâmetros do modelo de Gompertz para a inclusão das covariantes: clone, área basal ($G - m^2 ha^{-1}$), sítio (m) e idade (anos). O modelo de Gompertz foi ajustado conforme apresentado pela Equação (12).

$$H_i = \beta_0 \exp[-\beta_1 \exp(-\beta_2 \text{DAP}_i)] + \varepsilon_i \quad (12)$$

Em que: $\beta_0 = \beta_{00} + \beta_{01}G + \sum_{i=1}^3 \beta_{02}S_i + \beta_{03}I$; $\beta_1 = \beta_{10} + \beta_{11}I$; $\beta_2 = \beta_{20} + \beta_{21}G + \beta_{22}I + \sum_{j=1}^4 \beta_{23}\text{Clone}_j$. Onde: $\sum_{i=1}^3 \beta_{02}S_i = \beta_{02}$ = refere-se ao parâmetro associado ao *i-ésimo* índice de sítio e S_i é uma variável indicadora (*dummy*) com valor 1 para o *i-ésimo* sítio e 0 para os demais; $\sum_{j=1}^4 \beta_{23}\text{Clone}_j = \beta_{23}$ = refere-se ao parâmetro associado ao *j-ésimo* clone e Clone_j é uma variável indicadora (*dummy*) com valor 1 para o *j-ésimo* clone e 0 para os demais.

Pode ser verificado na Equação 12 que o parâmetro assíntota (β_0) foi influenciado, significativamente, por todas as covariantes adicionadas ao modelo, com exceção do clone. O parâmetro β_1 foi influenciado apenas pela idade do povoamento. Já o β_2 sofreu influência de todas as covariáveis, exceto a classe de sítio.

Na Tabela 6, são apresentadas as estatísticas de ajuste do modelo de Gompertz misto com a inclusão de covariantes (completo).

Tabela 6 - Parâmetros estimados para o modelo misto de Gompertz com inclusão de covariantes (completo).

Gompertz ($S_{yx} = 1,22m - 4,73\%$)					
Parâmetro	Variável Associada ao Parâmetro	Estimativa	Erro Padrão	t-valor	p-valor
β_0	Intercepto (Sítio - 37)	8,374589	1,002	8,356	< 0,0001
	G	0,229312	0,061	3,775	0,0002
	Sítio - 33	- 1,328811	0,203	-6,553	< 0,0001
	Sítio - 29	- 3,512091	0,427	-8,228	< 0,0001
β_1	Idade	3,583232	0,188	19,015	< 0,0001
	Intercepto	1,297585	0,388	3,339	0,0009
β_2	Idade	0,178279	0,069	2,593	0,0096
	Intercepto (Clones - 1 e 4)	0,221600	0,026	8,460	< 0,0001
	G	-0,001661	0,001	-2,229	0,0259
	Idade	-0,015455	0,003	-4,377	< 0,0001
	Clone - 2	0,065061	0,013	4,823	< 0,0001
	Clone - 3	0,055468	0,013	4,111	< 0,0001

Fonte: Do autor (2018).

Onde: Clone - 1 é *E. urocam*, Clone - 2 é *E. urograndis*; Clone - 3 é *E. urophylla* e Clone - 4 é uma mistura de três materiais: *E. urophylla*, *E. urograndis* e *E. urocam*.

As estatísticas apresentadas, na Tabela 6, mostram que todos os parâmetros foram significativos ($p < 0,05$). A adição de variáveis ao modelo misto não possibilitou uma melhoria de acurácia da estimativa, conservando o mesmo valor de erro padrão residual - 1,22m. Tal conclusão já era esperada, em virtude de os dados não serem longitudinais e os ajustes do modelo misto conseguir captar toda a variabilidade entre as unidades amostrais.

Resultados encontrados utilizando dados longitudinais, classes de *DAP* como variável aleatória e modelagem convencional sem o uso da teoria de modelos mistos mostram uma melhoria na precisão, como em Calegario et al.

(2015b), ao testarem modelos não lineares generalizados para estimativa da altura de árvores de eucalipto. Os autores verificaram uma redução de 53% do erro padrão residual (1,28m para 0,60m).

Mendonça, Carvalho e Calegario (2015), testando modelos não lineares generalizados de efeitos mistos para relação hipsométrica, constataram que o modelo assintótico foi o de maior acurácia com redução do erro padrão relativo em 36% (10,12% para 6,43%).

Alves et al. (2017), testando modelos lineares e não lineares para estimar a altura de árvores de *Toona ciliata* constataram que a adição da covariante altura dominante e codominante ao modelo logístico possibilitou um ganho de precisão de 23%.

Na Tabela 7, são apresentadas as estatísticas de comparação do modelo de Gompertz misto original (sem inclusão de covariantes) e o modelo de Gompertz misto completo (com adição de covariantes).

Tabela 7 - Estatísticas de comparação entre o modelo de Gompertz misto original e o modelo de Gompertz misto completo.

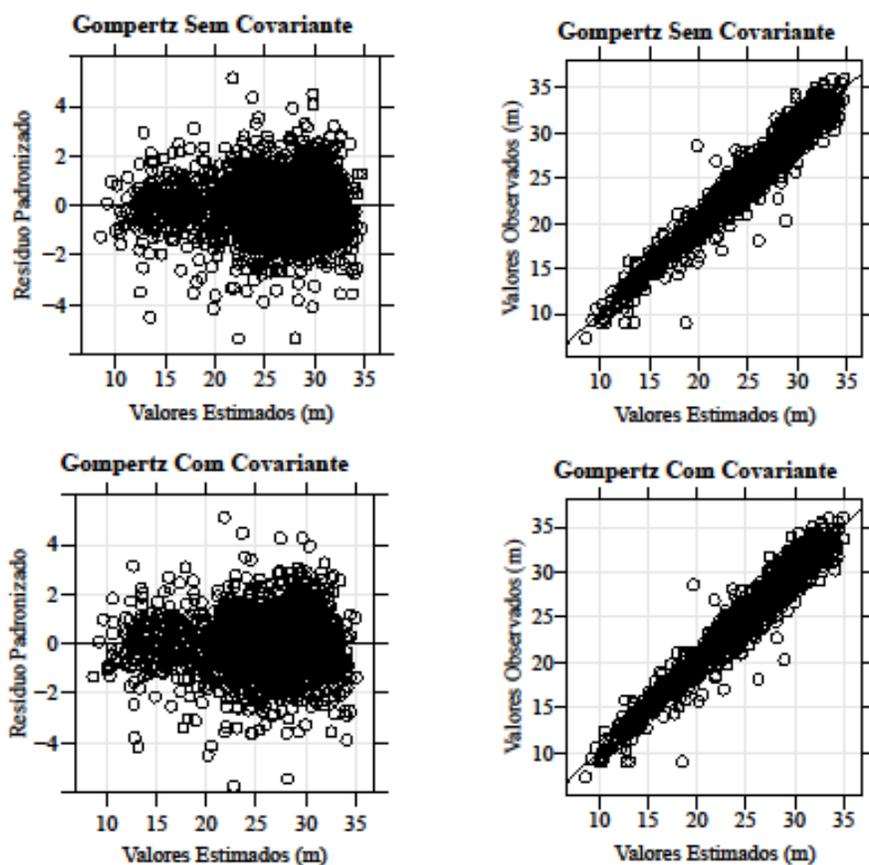
Modelo	G.L.	CIA	CIB	Ln _{mv}	TRMV	p-valor
Gompertz Misto	2138	7814,69	7854,39	-3900,35		
Gompertz Misto Covariante	2129	7426,35	7517,08	-3697,17	406,35	< 0,0001

Fonte: Do autor (2018).

De acordo com os resultados apresentados, na Tabela 7, nota-se uma leve superioridade do modelo de Gompertz completo em relação ao modelo original. Tal superioridade é verificada pelos menores valores de CIA e CIB, maior valor do logaritmo da máxima verossimilhança e pela significância ($p < 0,05$) do teste de razão da máxima verossimilhança.

Na Figura 5, é mostrado o gráfico de resíduos para os modelos de Gompertz misto original e completo e a relação entre as alturas observadas e previstas.

Figura 5 - Gráficos de resíduos e da relação entre as alturas observadas e estimadas para o modelo de Gompertz misto original e completo, respectivamente.



Fonte: Do autor (2018).

Na Figura 5, ao analisar a distribuição dos resíduos do modelo de Gompertz misto original e completo, nota-se um leve problema de heterocedasticidade. A heterogeneidade da variância pode ser confirmada ao se

observar uma leve variação crescente, em forma de funil, da relação entre os valores observados e estimados de altura.

Uma das vantagens atribuídas à técnica de adição de covariantes é a possibilidade de se obter boa acurácia das estimativas com o desenvolvimento de um único modelo. Na prática, a construção de um modelo genérico elimina a necessidade de estratificação do povoamento florestal para o ajuste de várias funções para estimar a altura de árvores de eucalipto. Além disso, o emprego dessa técnica pode representar uma redução dos custos de inventário florestal, seja pela redução da intensidade amostral de alturas mensuradas e não necessidade de uma estratificação, seja pela simplificação do processamento dos dados, em razão da não necessidade de ajuste de vários modelos.

3.4 Modelagem da heterocedasticidade

Na modelagem da heterocedasticidade, considerou-se a variável área basal como covariável (v_{ij}) a ser modelada. Seu uso se justifica pelo comportamento desta, ao longo dos anos, ou seja, com a variação da idade do povoamento, espera-se um aumento da variância na distribuição de G e, portanto, esta deve ser modelada.

A aplicação da estrutura de variância “varPower” de Pinheiro e Bates (2000) possibilitou uma redução na estimativa do erro padrão residual e percentual, quando comparado ao modelo de efeitos mistos com inclusão de covariantes (completo), de aproximadamente 55% (de 1,22m para 0,55m). Além disso, é verificada uma similaridade dos valores dos parâmetros, conforme Tabela 8.

Tabela 8 - Parâmetros estimados para o modelo de Gompertz misto heterocedástico.

Gompertz ($S_{yx} = 0,55m - 2,15\%$)					
Parâmetro	Variável Associada ao Parâmetro	Estimativa	Erro Padrão	t-valor	p-valor
β_0	Intercepto (Sítio - 37)	8,345001	0,991	8,423	< 0,0001
	G	0,230694	0,061	3,781	0,0002
	Sítio - 33	-1,328188	0,203	-6,531	< 0,0001
	Sítio - 29	-3,504187	0,423	-8,279	< 0,0001
	Idade	3,587267	0,189	18,950	< 0,0001
β_1	Intercepto	1,283322	0,370	3,469	0,0005
	Idade	0,178484	0,066	2,707	0,0068
β_2	Intercepto (Clones - 1 e 4)	0,220403	0,025	8,768	< 0,0001
	G	-0,001684	0,001	-2,259	0,0240
	Idade	-0,015383	0,003	-4,430	< 0,0001
	Clone - 2	0,065743	0,013	5,022	< 0,0001
	Clone - 3	0,056189	0,013	4,294	< 0,0001

Fonte: Do autor (2018).

Na Tabela 9, são apresentadas as estatísticas de comparação entre o modelo misto homocedástico e o modelo misto heterocedástico.

Tabela 9 - Estatísticas de comparação entre o modelo de Gompertz misto homocedástico e o modelo de Gompertz misto heterocedástico.

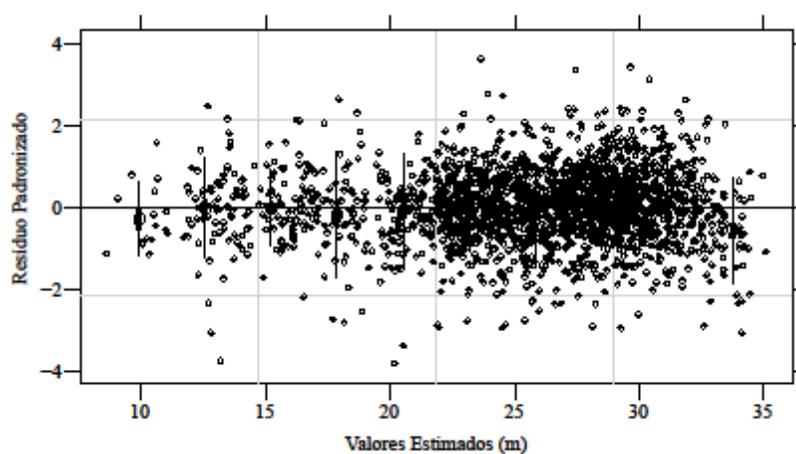
Modelo	G.L.	CIA	CIB	Lnmv	TRM V	p-valor
Gompertz Misto Homocedástico	2129	7826,35	7517,08	-3697,17		
Gompertz Misto Heterocedástico	2128	7420,55	7516,95	-3693,27	7,80	0,0052

Fonte: Do autor (2018).

De acordo com os critérios estatísticos apresentados na Tabela 9, a utilização da função de variância “varPower” proporcionou uma leve superioridade do modelo de Gompertz heterocedástico. Tal superioridade é comprovada pelos menores valores de CIA e CIB, maior valor do logaritmo da máxima verossimilhança e pela significância ($p < 0,05$) do teste de razão da máxima verossimilhança.

Na Figura 6, é apresentada a distribuição de resíduos para o modelo de Gompertz misto heterocedástico.

Figura 6 - Distribuição de resíduos para o modelo de Gompertz misto heterocedástico.



Fonte: Do autor (2018).

Para o modelo de Gompertz misto heterocedástico, é verificada uma melhor distribuição de resíduos em relação ao gráfico de resíduos do modelo de Gompertz original. Os intervalos de confiança (linhas verticais grossas) das classes encontram-se quase todos sobre a linha média do gráfico de resíduos do modelo heterocedástico, indicando uma boa qualidade de ajuste. Além disso, o problema de heterogeneidade da variância verificado no modelo original foi

solucionado pela modelagem da estrutura de variância. Com o uso dessas técnicas, foi possível verificar um comportamento uniforme dos desvios-padrão (linhas verticais finas), ocorrendo uma subestimativa nas maiores alturas.

4 CONCLUSÃO

A utilização da teoria dos modelos não lineares de efeitos mistos possibilitou um aumento da acurácia.

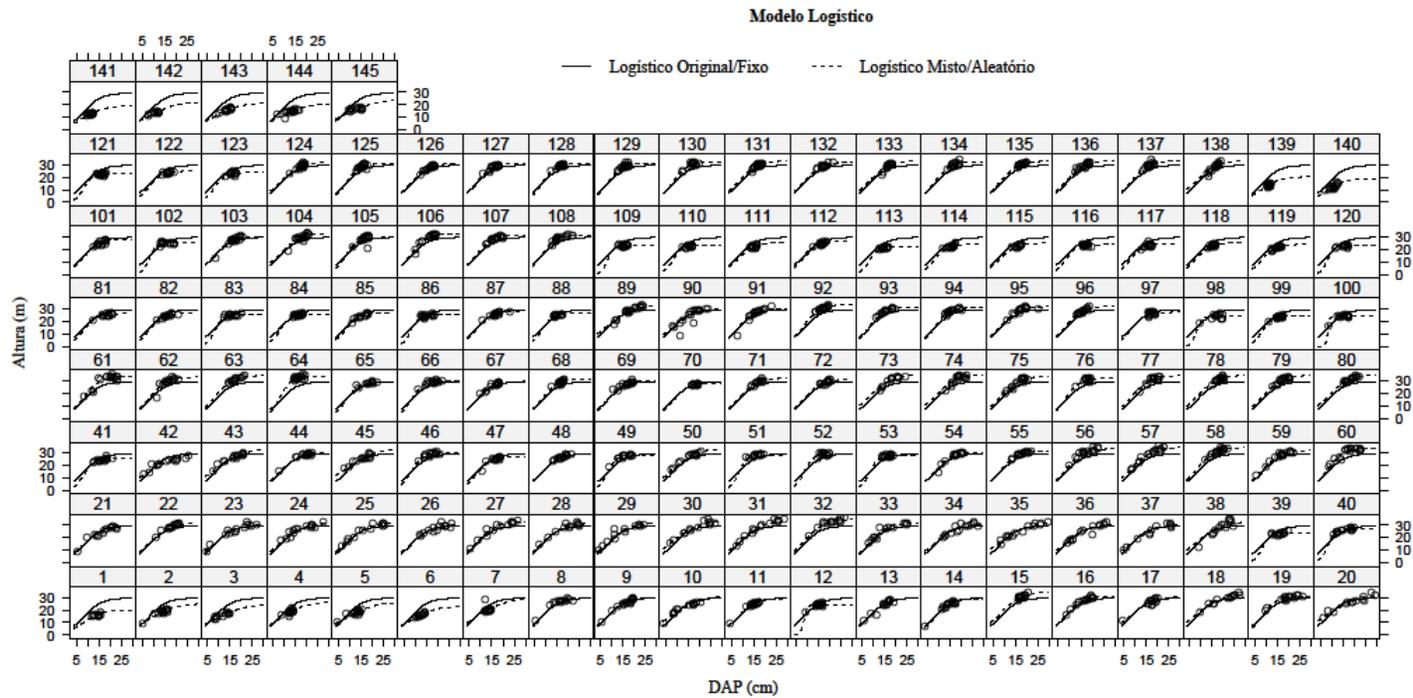
De acordo com as estatísticas de ajuste, todos os modelos testados apresentaram comportamentos semelhantes, entretanto, em razão da leve superioridade apresentada pelo modelo de Gompertz nas estatísticas CIA, CIB e erro padrão residual, este foi utilizado para a predição das alturas.

No modelo de Gompertz misto, todos os parâmetros foram influenciados, significativamente, pelas variáveis do povoamento. A assíntota foi influenciada pelo sítio, área basal (G) e idade. O parâmetro β_1 foi influenciado somente pela idade. Já o parâmetro β_2 foi influenciado pelo clone, área basal e pela idade.

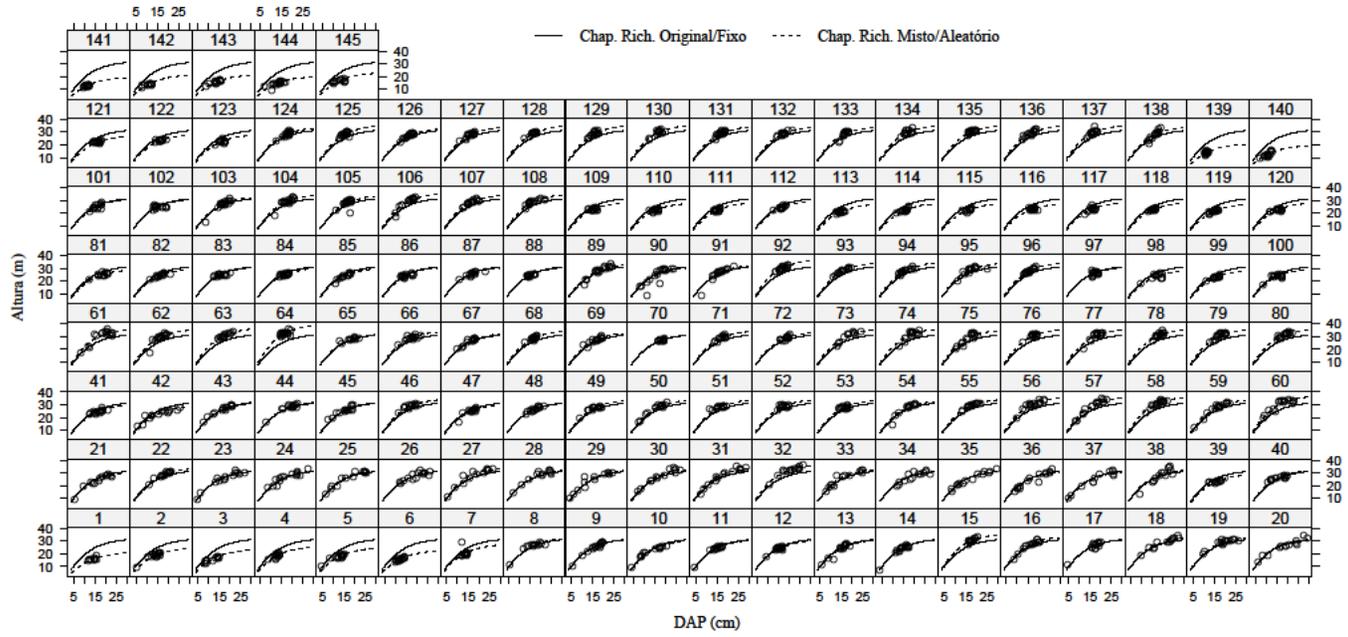
A adição de covariantes ao modelo de Gompertz misto possibilitou uma leve superioridade das estatísticas CIA, CIB, $\ln mv$ e TRMV, e conservou os mesmos valores de erro padrão residual e percentual quando comparado ao modelo misto.

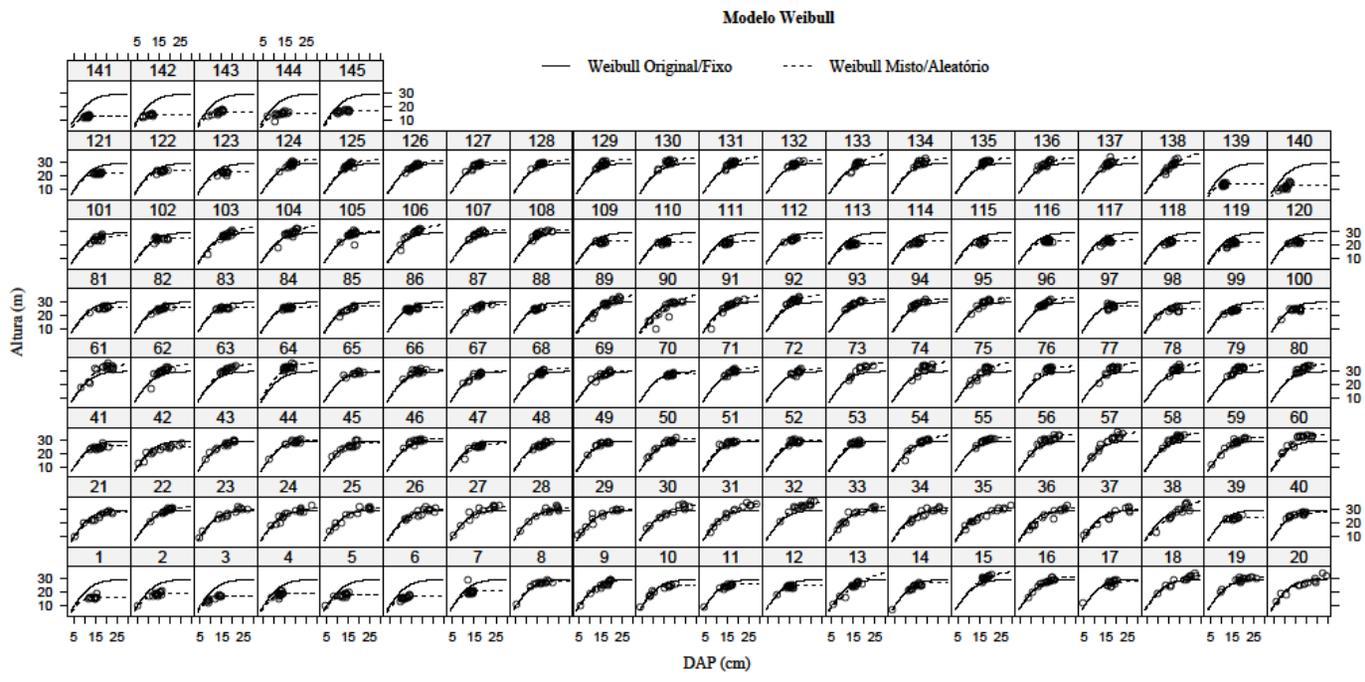
A modelagem da variância proporcionou uma redução do erro padrão residual em aproximadamente 55%, quando comparado com o modelo misto com adição de covariantes (completo) (de 1,22m para 0,55m). A partir da comparação entre as estatísticas CIA, CIB, $\ln mv$ e TRMV, o modelo heterocedástico mostrou-se superior ao modelo completo, tendo em vista os menores valores dos critérios CIA e CIB, o maior valor do $\ln mv$ e a significância do teste de razão da máxima verossimilhança.

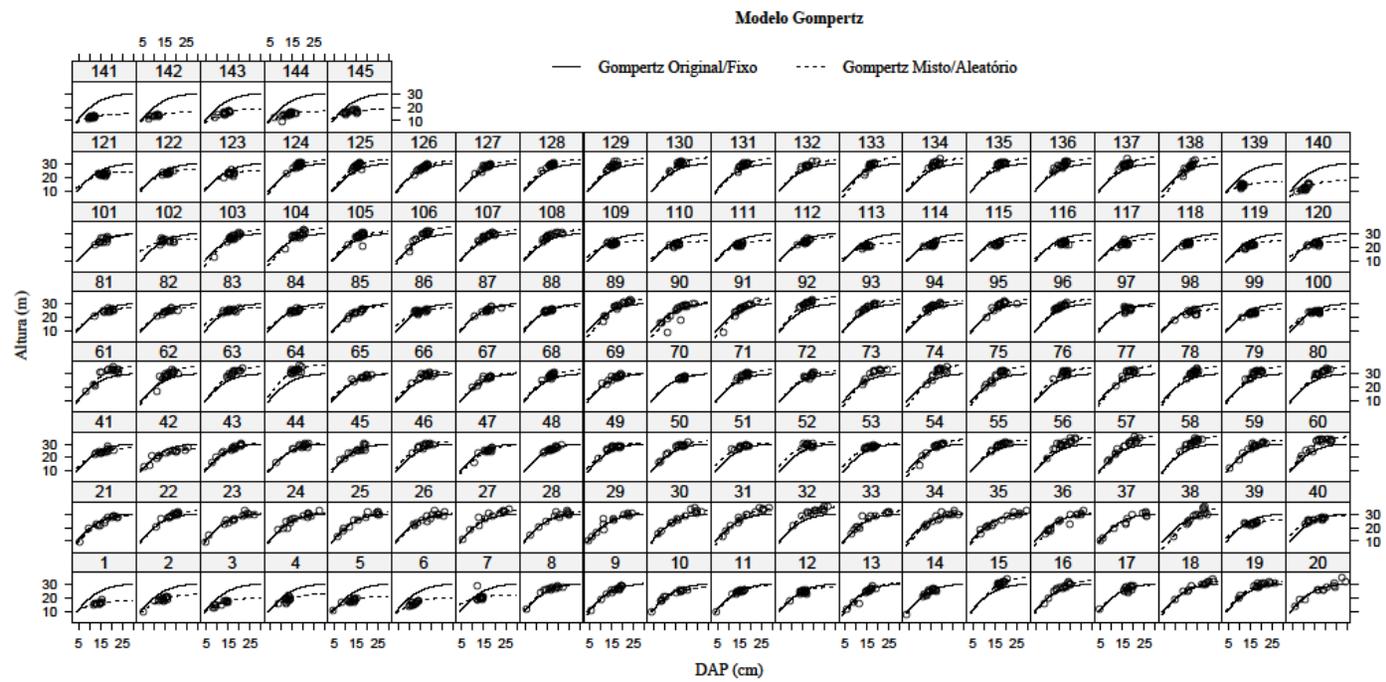
ANEXO A - AJUSTE DAS CURVAS DE REGRESSÃO PARA OS QUATROS MODELOS DE EFEITOS MISTOS TESTADOS.



Modelo Chapman e Richards







Fonte: Do autor (2018).

REFERÊNCIAS

- ALVES, J. de A. et al. Equações hipsométricas para *Toona ciliata* com inclusão de covariantes. **Ciência Florestal**, Santa Maria, v. 27, n. 2, p. 581-595, abr./jun. 2017.
- CALEGARIO, N. et al. Estimativa do crescimento de povoamentos de *Eucalyptus* baseada em modelos lineares em multiníveis de efeito misto. **Revista Árvore**, Viçosa, v. 29, n. 2, p. 251-261, mar./abr. 2005a.
- _____. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não lineares generalizados. **Scientia Forestalis**, Piracicaba, n. 69, p. 38-50, dez. 2005b.
- CARVALHO, S. de P. C. e et al. Modelos não lineares generalizados aplicados na predição da área basal de volume de *Eucalyptus* clonal. **Cerne**, Lavras, v. 17, n. 4, p. 541-548, out./dez. 2011.
- GREGOIRE, T. G.; SCHABENBERGER, O. Nonlinear mixed-effects modeling of cumulative bole volume with spatially correlated within-tree data. **Journal of Agricultural, Biological, and Environmental Statistics**, Alexandria, v. 1, n. 1, p. p. 107-119, Mar. 1996.
- LAPPI, J. A longitudinal analysis of height/diameter curves. **Forest Science**, Bethesda, v. 43, n. 4, p. 555-570, Nov. 1997.
- _____. Calibration of height and volume equations with random parameters. **Forest Science**, Bethesda, v. 37, n. 3, p. 781-801, Aug. 1991.
- MCCULLOCH, C. E.; SEARLE, S. R. **Generalized, linear, and mixed models**. New York: Wiley Series in Probability and Statistics, 2001. v. 1, 325 p.
- MEHTÄTALO, L. A longitudinal height-diameter model for Norway spruce in Finland. **Canadian Journal of Forest Research**, Ottawa, v. 34, n. 1, p. 131-140, 2004.
- MENDONÇA, A. R. de et al. Modelos hipsométricos e de crescimento em altura das árvores dominantes e codominantes para *Pinus caribaea* var. hondurensis. **Scientia Forestalis**, Piracicaba, v. 39, n. 90, p. 151-160, jun. 2011.

MENDONÇA, A. R. de; CARVALHO, S. de P. C. e; CALEGARIO, N. Modelos hipsométricos generalizados mistos na predição da altura de *Eucalyptus* sp. **Cerne**, Lavras, v. 21, n. 1, p. 107-114, jan./mar. 2015.

PINHEIRO, J. C.; BATES, D. M. **Mixed-Effects Models in S and S-PLUS**. New York: Springer, 2000. 528 p.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. **Akaike information criterion statistics**. Dordrecht: D. Reidel, 1986. 290 p.

SCHWARZ, G. Estimating the dimension of a model. **Annals of Statistics**, Hayward, v. 6, n. 2, p. 461-464, Mar. 1978.

TRINCADO, G.; BURKHART, H. E. A generalized approach for modeling and localizing stem profile curves. **Forest Science**, Bethesda, v. 52, n. 6, p. 670-682, June 2006.

YANG, Y.; HUANG, S. Comparison of different methods for fitting nonlinear mixed forest models and for making predictions. **Canadian Journal of Forest Research**, Ottawa, v. 41, n. 8, p. 1671-1686, Aug. 2011.

ZHAO, D.; WILSON, M.; BORDER, B. E. Modeling response curves and testing treatment effects in repeated measures experiments: a multilevel nonlinear mixed-effects model approach. **Canadian Journal of Forest Research**, Ottawa, v. 35, p. 122-132, 2005.

**ARTIGO 2 - MODELAGEM DO CRESCIMENTO E PRODUÇÃO DE
POVOAMENTOS DE EUCALIPTO UTILIZANDO MODELO DE
REGRESSÃO E APRENDIZAGEM DE MÁQUINA**

**Artigo redigido conforme a NBR 6022 (ABNT, 2003) e formatado de acordo
com o Manual da UFLA de apresentação de teses e dissertações.**

RESUMO

Conhecer o crescimento e a produção de povoamentos florestais é um dos principais objetivos do manejador. A busca por técnicas capaz de produzir melhor acurácia nas estimativas é justificável, possibilitando um melhor planejamento do empreendimento florestal. Neste trabalho, objetivou-se analisar o desempenho de técnicas de aprendizagem de máquinas (Redes Neurais Artificiais - RNA e Máquina de Vetor de Suporte - MVS) na projeção da área basal e volume em comparação ao método tradicional de análise de regressão, utilizando o modelo de Clutter. Além disso, o estudo buscou a construção de variações do sistema de Clutter e extração do modelo da rede Perceptron com seus respectivos valores de pesos e bias (viés). A base de dados é proveniente de um inventário florestal contínuo (IFC) de povoamentos clonais do híbrido *Eucalyptus grandis* x *E. urophylla*, compreendendo sete materiais genéticos. Foram amostradas 2.550 unidades amostrais permanentes, distribuídas em diferentes idades e locais. Destas 2.550 parcelas foram criadas, aleatoriamente, dois conjuntos de dados: a) 1.525 (60%) unidades amostrais para treinamento e b) 1.025 (40%) unidades amostrais para teste. Os métodos de aprendizagem de máquinas para projeção da área basal proporcionaram boa capacidade de treinamento e generalização. Para projeção volumétrica, os métodos apresentaram baixa capacidade de generalização. As estimativas produzidas pelo sistema de Clutter e variações deste foram superiores às técnicas de aprendizagem de máquinas, tanto para os dados de treinamento como para os dados de teste. A extração do modelo da rede Perceptron de Múltiplas Camadas possibilitou a estimativa da área basal e do volume. Dessa forma, o modelo poderá ser utilizado para qualquer povoamento, considerando o mesmo material genético, sem a necessidade do treinamento constante de redes.

Palavras-chave: Aprendizagem de Máquinas. Redes Neurais Artificiais. Máquina de Vetor de Suporte. Regressão. Modelo de Clutter.

ARTICLE 2 - MODELING OF THE GROWTH AND PRODUCTION OF EUCALYPTUS STANDS USING A REGRESSION MODEL AND MACHINE LEARNING

ABSTRACT

Knowing the growth and production of forest stands is one of the main objectives of the forest manager. The search for techniques capable of producing better estimation accuracy is justifiable, allowing better planning of the forestry enterprise. The objective of this work was to analyze the performance of machine learning techniques (Artificial Neural Networks - ANN and Support Vector Machine - SVM) in the projection of basal area and volume and comparing them to the traditional method of regression analysis using the Clutter model. Moreover, we sought to construct variations of the Clutter system and extract the Perceptron model with their respective weights and bias values. The database derives from a continuous forest inventory (CFI) of clonal stands of the *Eucalyptus grandis* x *E. urophylla* hybrid, comprising seven genetic materials. We sampled 2,550 permanent sample units, distributed at different ages and sites. Of the 2,550 plots, we randomly created two sets of data: a) 1,525 (60%) training sample units and b) 1,025 (40%) testing sample units. The machine learning methods for projecting basal area provided good training and generalization skills. For volumetric projection, the methods presented low generalization capacity. The estimates produced by the Clutter system and its variations were superior to the machine learning techniques for both training data and test data. The extraction of the Multiple Layers Perceptron network model allowed the estimation of basal area and volume. Thus, the model can be used for any population, considering the same genetic material, without the need for constant network training.

Keywords: Machine Learning. Artificial neural networks. Support Vector Machine. Regression. Clutter model.

1 INTRODUÇÃO

De acordo com Fausett (1994), as técnicas de aprendizagem de máquinas - redes neurais artificiais (RNAs) e máquina de vetor de suporte (MVS) - são sistemas de processamento que tentam replicar o funcionamento do cérebro humano pela forma em que os neurônios biológicos são conectados entre si. Algumas características das RNAs são atribuídas por Fausett (1994) e Haykin (2001), como: o processamento de informações ocorre em estruturas simples representadas pelos neurônios artificiais; os sinais são passados por links de conexão; cada link tem um peso associado e; cada neurônio artificial possui uma função de ativação que determina um sinal ou valor de saída.

A motivação do uso da aprendizagem de máquinas está relacionada com vantagens dessa técnica em relação ao sistema de modelagem convencional. As principais vantagens são mencionadas por Sando et al. (2005), como: não necessidade de atender às pressuposições da análise de regressão; capacidade de modelar relações não lineares complexas entre os dados de entrada e saída das redes. Além disso, as redes possuem a capacidade de generalização, ou seja, quando treinadas são capazes de produzir soluções a conjunto de dados não conhecidos. No entanto, Peng e Wen (1999) e Sando et al. (2005) mencionam algumas desvantagens relacionadas ao consumo de tempo maior para treinamento, exigindo, conseqüentemente, um maior esforço computacional. Outra desvantagem apontada por Özçelik et al. (2010), é atribuída aos pesos obtidos durante o treinamento. Para esses pesquisadores, os pesos não são interpretáveis, em razão da presença da(s) camada(s) ocultas e não linearidade da função de ativação.

Assim como as redes perceptrons de múltiplas camadas (PMC) e a função de base radial (RBF), a máquina de vetor de suporte (MVS) é uma ferramenta de modelagem, a qual também tem sido utilizada para aproximação

de funções. Diferentemente das redes PMC e RBF, a máquina de vetor de suporte apresenta vantagens relacionadas à obtenção do ótimo global, da boa capacidade de generalização, da facilidade de implementação e independência dimensional (SMOLA et al., 1999; VAPNIK, 1995).

Vários trabalhos na área florestal são desenvolvidos com base em técnicas de aprendizagem de máquinas para modelagem de variáveis do povoamento. Dentre os trabalhos que modelam as variáveis, pode-se citar: estimação do volume de árvores de *Pinus brutia* Ten., *Pinus sylvestris* L., *Abies cilicica* Carr., *Cedrus libani* A. Rich., *Pinus pinaster*, *Acacia mangium*, eucalipto e *Tectona grandis* L.f. utilizando as redes PMC (DIAMANTOPOULOU, 2005; GORGENS et al., 2009, 2014; ÖZÇELIK et al., 2010), RBF (CASTELLANOS; BLANCO; PALENCIA, 2007) e MVS (CORDEIRO et al., 2015); modelagem da altura de árvores de Crimean juniper e eucalipto utilizando a rede PMC (BINOTI; BINOTI; LEITE, 2013; ÖZÇELIK et al., 2013); modelagem do afilamento de fustes de eucalipto e de vegetação nativa (NUNES; GORGENS, 2016; SCHIKOWSKI; CORTE; SANQUETTA, 2015); projeção da área basal e volume em nível de povoamento de clones de eucalipto utilizando as redes PMC e RBF (BINOTI et al., 2015).

Apesar da técnica de aprendizagem de máquinas apresentar um papel importante na área florestal, o uso de modelos de regressão, especial o sistema de projeção de Clutter, ainda é o mais utilizado pelas empresas florestais. Tal fato se deve ao bom desempenho em termos de acurácia e consistência biológica do sistema, bem como por predizer e projetar tanto o crescimento como a produção em área basal e volume (CLUTTER, 1963; CLUTTER et al., 1983). Variações do sistema de Clutter são possíveis e, às vezes, necessárias pelo fato de algumas variáveis integrantes do modelo, após o ajuste, não serem significativas e o sinal das estimativas paramétricas não apresentarem consistência biológica. Dessa forma, é justificável construir um modelo que

apresente significância em seus parâmetros e que o sinal das estimativas paramétricas seja condizente com a relação das variáveis.

Nesse contexto, objetivou-se, com este trabalho, analisar o desempenho de técnicas de aprendizagem de máquinas na projeção da área basal e volume em comparação ao modelo proposto por Clutter. Além disso, o estudo busca a obtenção de variações do sistema de Clutter e extração do modelo da rede PMC com seus respectivos valores de pesos e bias.

2 MATERIAL E MÉTODOS

2.1 Caracterização da área de estudo e base de dados

Os dados para a realização do estudo são provenientes de 25.500 ha de reflorestamento pertencente à International Paper, localizados nas regiões norte de São Paulo (Ribeirão Preto) e no extremo sul de Minas Gerais. O clima das regiões enquadra-se nos tipos Aw e Cwa, respectivamente, que, de acordo com a classificação climática de Köppen, apresenta dois períodos distintos, um chuvoso e outro seco, com temperatura média anual variando entre 20 e 23°C e precipitação média anual com valores entre 1300 e 1600mm.

A base de dados é composta por sete materiais genéticos (H13, H15, H17, H27, H34, H35 e H144) do híbrido *Eucalyptus grandis* x *E. urophylla*, implantados em seis espaçamentos: 3 x 2; 3 x 2,5; 3 x 2,75; 3 x 3; 3,3 x 2,5 e 3,6 x 2,5m. O regime de manejo adotado pela empresa predomina o de alto fuste, visando à produção de madeira para celulose, com idade de corte variando entre 6 e 7 anos.

Os dados são provenientes de inventário florestal contínuo (IFC). O método de amostragem foi a casual simples (ACS), com parcelas de área variando entre 300 a 460m². No total, foram amostradas 10.312 unidades amostrais, compreendendo parcelas temporárias e permanentes, distribuídas em diferentes idades e locais. Uma análise descritiva das variáveis do povoamento é apresentada na Tabela 1.

Tabela 1 - Estatísticas descritivas relacionadas às variáveis do povoamento.

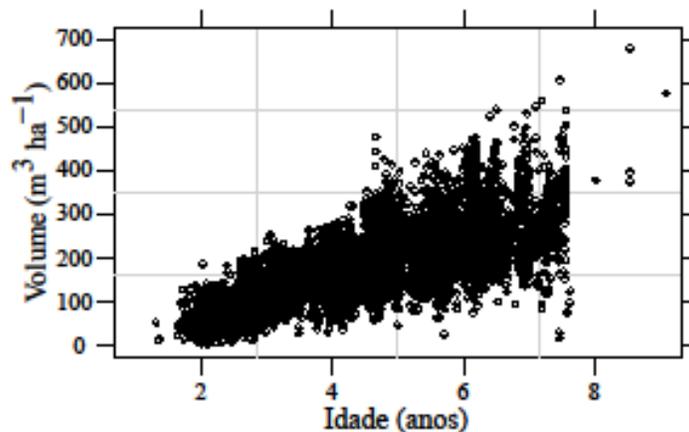
Variável do Povoamento	Mínimo	Máximo	Média	CV (%)
I	1,31	9,09	4,06	40,43
$\overline{\text{HDC}}$	6,40	39,65	20,56	28,44
S	7,61	51,89	26,76	20,07
G	1,73	46,05	15,92	40,44
V	5,65	680,85	160,27	59,72

Fonte: Do autor (2018).

Em que: I = idade (anos); $\overline{\text{HDC}}$ = altura média das árvores dominantes e codominantes (m); S = índice de sítio (m); G = área basal ($\text{m}^2 \text{ha}^{-1}$); V = volume ($\text{m}^3 \text{ha}^{-1}$); CV = coeficiente de variação.

Na Figura 1, é apresentada a relação entre o volume ($\text{m}^3 \text{ha}^{-1}$) e a idade (anos) dos povoamentos de eucalipto.

Figura 1 - Relação entre o volume ($\text{m}^3 \text{ha}^{-1}$) e a idade (anos) dos povoamentos de eucalipto.



Fonte: Do autor (2018).

2.2 Modelagem do crescimento e produção em área basal e volume

Para a modelagem do crescimento e produção em área basal e volume, foram utilizadas e comparadas as técnicas de aprendizagem de máquinas (RNA

e MVS) e modelagem convencional, utilizando o sistema de Equações proposto por Clutter (1963) e suas modificações. Para a manipulação dos dados, treinamento das RNAs, ajuste dos modelos e análises estatísticas, utilizou-se o *software R*.

2.2.1 Preparação dos dados

As variáveis do povoamento utilizadas para modelagem do crescimento e produção florestal foram: a idade (anos), o índice de sítio (m), a área basal ($\text{m}^2 \text{ha}^{-1}$) o volume ($\text{m}^3 \text{ha}^{-1}$) e o material genético. O material genético foi introduzido aos modelos como variável *dummy*.

Para aplicação da IA, como os dados apresentam grandezas diferentes, procedeu-se com a normalização ou padronização das variáveis, utilizando a transformação que permite que os valores de entrada e saída da rede estejam entre 0,1 e 0,9 (EQUAÇÃO 1).

$$VT_i = (((0,9 - 0,1)(x_i - \min_i)) / (\max_i - \min_i)) + 0,1 \quad (1)$$

Em que: V_{Ti} = valor transformado da *i-ésima* variável; min e max = valor mínimo e máximo da *i-ésima* variável, respectivamente.

Tanto para a aplicação da técnica de aprendizagem de máquinas (RNA e MVS) como para o ajuste da regressão pelo sistema de Clutter e suas modificações, foram utilizados somente dados de parcelas permanentes. No total, são 2.550 parcelas permanentes que foram divididas, aleatoriamente, em dois conjuntos de dados: a) 1.525 (60%) unidades amostrais para o treinamento de máquinas e ajuste dos modelos de regressão e b) 1.025 (40%) unidades amostrais para o teste de máquinas e dos modelos de regressão.

2.3 Modelagem pela aprendizagem de máquinas

Para modelagem da área basal e volume foram testadas as máquinas perceptron de múltiplas camadas (PMC), a função de base radial (RBF) e a máquina de vetor de suporte (MVS).

2.3.1 Perceptron de múltiplas camadas (PMC)

Para o ajuste de pesos (parâmetros) e bias (viés), por meio da rede perceptron de múltiplas camadas, empregou-se a função do *software* R *neuralnet*. Utilizou-se a arquitetura *feedforward* e o algoritmo de retropropagação resiliente (*resilient backpropagation*), representado pela função *rprop+*, com o treinamento realizado de forma supervisionada.

Esse algoritmo foi descrito por Fausett (1994), Haykin (2001) e Rumelhart, Hinton e Williams (1986). Segundos esses autores, o treinamento dessa rede é realizado mediante a execução das fases: *forward* e *backward*. Essas fases são executadas até que sejam encontradas estimativas de pesos e bias que minimizem a soma do erro quadrático (RMSE) produzido pelas respostas da rede frente àquelas desejadas.

Na fase *forward*, valores iniciais aleatórios de pesos e bias são produzidos e utilizados pelo algoritmo para se obter as saídas da rede. Em seguida, as respostas produzidas pela saída da rede são comparadas com as respectivas respostas desejadas. Caso os desvios (erros) produzidos entre a saída da rede e a saída desejada não forem satisfatórios, aplica-se a fase *backward*. Na fase *backward*, ocorrem as alterações dos valores de pesos e bias de todos os neurônios da rede, por meio do processo de retropropagação do erro para os neurônios das camadas anteriores (BRAGA; CARVALHO; LUDERMIR, 2011;

FAUSETT, 1994; HAYKIN, 2001; RUMELHART; HINTON; WILLIAMS, 1986; SILVA; SPATTI; FLAUZINO, 2010).

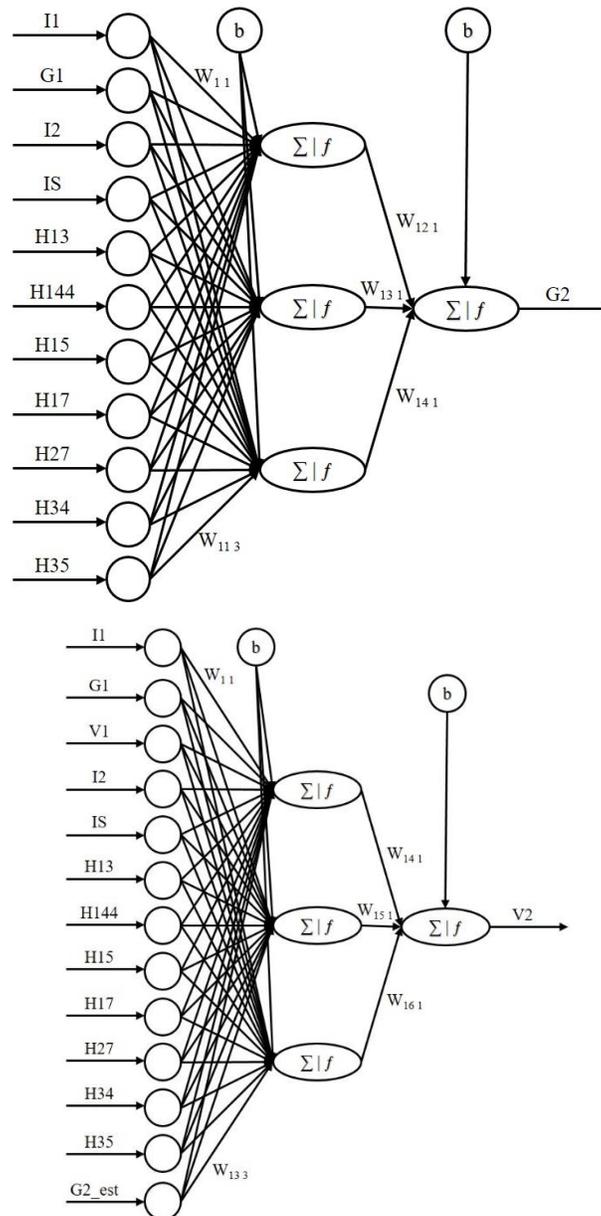
Foram utilizados três neurônios com função de ativação sigmoide logística na camada intermediária (EQUAÇÃO 2) e na camada de saída, a função de ativação linear, a qual desempenha um papel de combinação linear das funções.

$$g(u) = \frac{1}{1 + e^{-u}} \quad (2)$$

Em que: u = potencial de ativação.

Na Figura 2, é apresentada a arquitetura da rede PMC para projeção da produção em área basal e volume, respectivamente.

Figura 2 - Arquitetura da rede PMC para projeção da produção em área basal (G_2) e volume (V_2).



Fonte: Do autor (2018).

2.3.2 Função de base radial (RBF)

Com características semelhantes às redes PMCs, as redes RBFs também utilizam a arquitetura *feedforward* e o treinamento é realizado de forma não supervisionada e supervisionada, respectivamente. No R foi utilizado o pacote *RSNNS* e para o processamento, a função *rbf*.

O treinamento da rede RBF é realizado em duas etapas distintas. A primeira etapa está relacionada ao ajuste dos pesos (c) e da variância da função gaussiana (σ^2) dos neurônios da camada intermediária. Nessa etapa, o treinamento não é supervisionado e está diretamente relacionado com a alocação das funções de base radial e formação de clusters. Já na segunda etapa, o treinamento é realizado de forma supervisionada e objetiva estimar os valores de pesos e bias dos neurônios da camada de saída (FAUSETT, 1994; HAYKIN, 2001; SILVA; SPATTI; FLAUZINO, 2010).

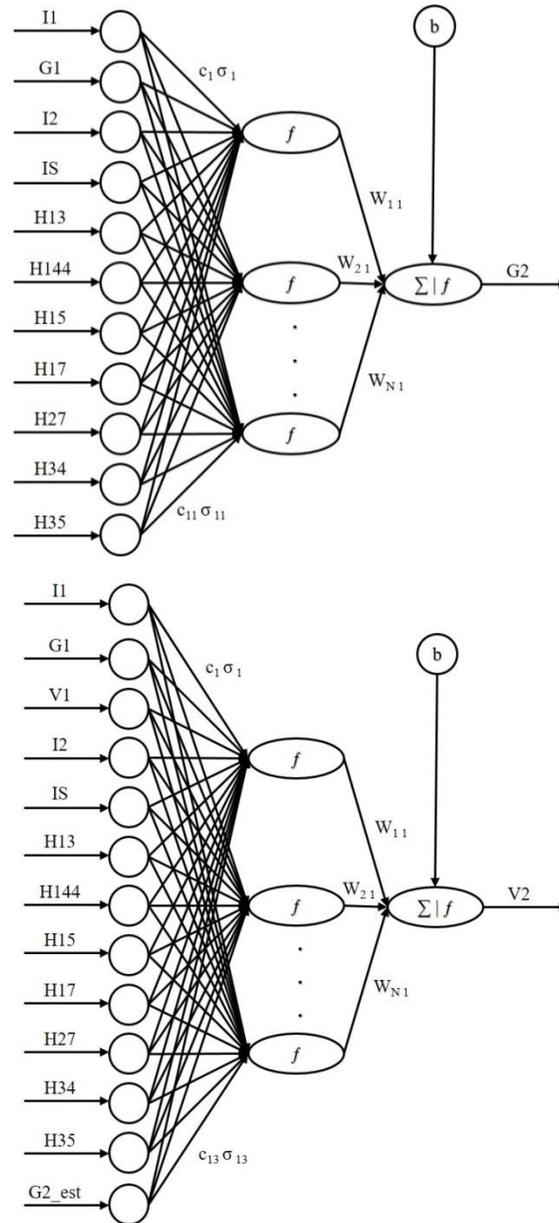
Assim, os valores de entrada serão as respostas produzidas pelas funções de ativação gaussiana dos neurônios da camada intermediária e o neurônio da camada de saída (linear) realizará somente uma combinação linear das funções gaussianas. A função sigmoide gaussiana é representada pela Equação 3.

$$g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}} \quad (3)$$

Em que: c = centro da função gaussiana, associado aos próprios pesos; σ^2 = variância da função gaussiana (raio).

Na Figura 3, é apresentada a arquitetura da rede RBF para projeção da produção em área basal e volume.

Figura 3 - Arquitetura da rede RBF para projeção da produção em área basal (G2) e volume (V2).



Fonte: Do autor (2018).

Para o treinamento da rede RBF, foram considerados 20 neurônios com função de ativação gaussiana na camada intermediária.

2.3.3 Máquina de vetor de suporte (MVS)

Na MVS, a construção do algoritmo de treinamento por vetor de suporte é dependente da formação do núcleo do produto interno entre um vetor de suporte, dado por um pequeno subconjunto dos dados de treinamento, e um vetor retirado do espaço de entrada, usando a função de Kernel. Dependendo de como esse núcleo de produto interno é gerado, pode-se construir diferentes máquinas de aprendizagem, como: a) Máquinas de aprendizagem polinomial, b) Redes de função de base radial (RBF), e c) Perceptron de, pelo menos, uma camada intermediária (HAYKIN, 2001).

Segundo Haykin (2001), o algoritmo de treinamento por vetor de suporte minimiza uma função de perda (erro) insensível, que é uma extensão do critério do erro médio absoluto da teoria *minimax* proposta por Huber em 1964 e 1981. Isso se dá pela necessidade de um estimador robusto e que seja insensível a pequenas variações no modelo. Assim, um estimador de mínimos quadrados torna-se inadequado pela sua sensibilidade à presença de pontos afastados (*outliers*), necessitando, portanto, de um algoritmo robusto.

Diferentemente das redes PMC e RBF, o número de neurônios (Kernels) da camada intermediária e seus respectivos valores de peso e bias é determinado, automaticamente, a partir de conjuntos de dados de treinamento (vetores de suporte).

O ajuste foi realizado no R, utilizando o pacote *e1071*. A função de Kernel utilizada foi a RBF gaussiana, representada pela Equação 4.

$$K(u, v) = e^{-\frac{|u-v|^2}{2\sigma^2}} \quad (4)$$

Em que: $K(u, v)$ = função de Kernel; u = vetor de entrada; v_i = vetor de suporte; $2\sigma^2$ = parâmetro *gamma*.

No primeiro momento, para o ajuste, foi utilizada a função *svm*. Com o intuito de melhorar a acurácia do ajuste, foi aplicada a *cross-validation*. Nesse caso, foi utilizada a função *tune*.

Na Figura 4, é apresentada a arquitetura da máquina de aprendizado por vetor de suporte para projeção da área basal e volume.

Figura 4 - Arquitetura da máquina de aprendizagem por vetor de suporte para projeção da área basal (G2) e volume (V2).

(Continua)

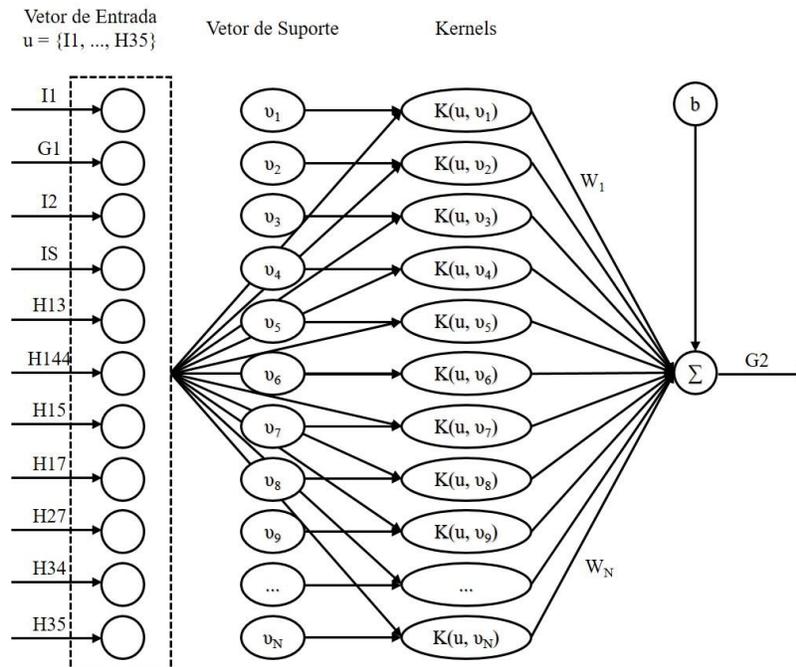
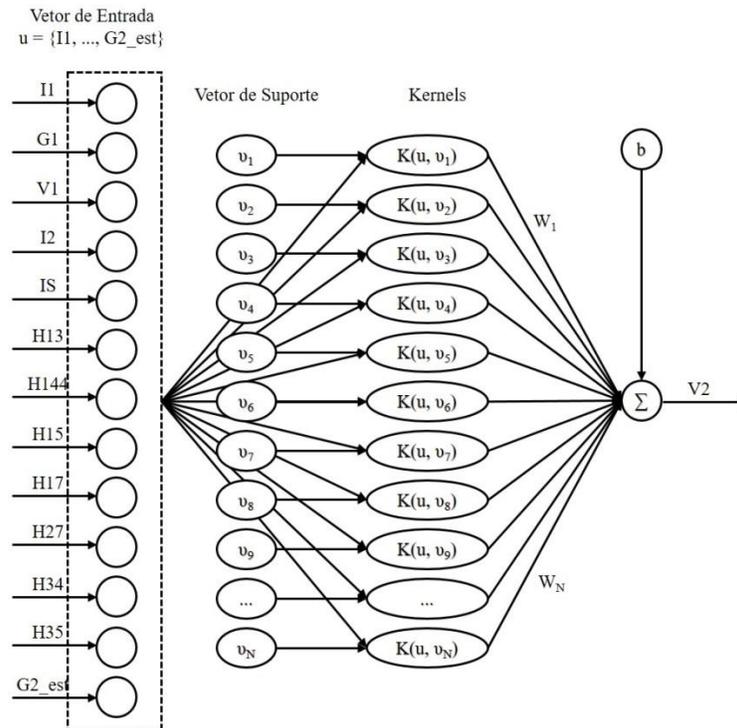


Figura 4 - Arquitetura da máquina de aprendizagem por vetor de suporte para projeção da área basal (G2) e volume (V2).

(Conclusão)



Fonte: Do autor (2018).

2.4 Modelagem pelo sistema de Equações de Clutter

Na segunda técnica utilizada, trata-se do modelo de Clutter, o qual é bastante utilizado para estudos de predição e projeção do crescimento e da produção em área basal e volume. É composto por um sistema de Equações consistente capaz de prever com alta precisão a área basal e volume atual e futuro (CLUTTER, 1963; CLUTTER et al., 1983).

Esse sistema de Equações iniciou-se com uma função desenvolvida por Schumacher e Coile (1960), para a modelagem da produção em área basal, Equação (5).

$$\ln(G_i) = \alpha_0 + \alpha_1 I_i^{-1} + \alpha_2 S_i + \alpha_3 I_i^{-1} S_i + \alpha_4 I_i^{-1} \ln(G_{i10}) + \varepsilon_i \quad (5)$$

Em que: G_i associada ao α_4 = área basal da *i-ésima* unidade amostral ($m^2 ha^{-1}$) na idade de 10 anos.

Para obter-se a taxa de crescimento em área basal, correspondente a uma determinada idade, é aplicada a diferencial implícita.

$$\frac{1}{G_i} \frac{\partial G}{\partial I} = -\alpha_1 I_i^{-2} - \alpha_3 S_i I_i^{-2} - \alpha_4 \ln(G_i) I_i^{-2}$$

Desenvolvendo a Equação acima, tem-se:

$$\frac{\partial G}{\partial I} = -I_i^{-1} G_i (\alpha_1 I_i^{-1} + \alpha_3 I_i^{-1} S_i + \alpha_4 I_i^{-1} \ln(G_i)) \quad (5.1)$$

Substituindo na função (5) a Equação (5.1), tem-se a taxa de crescimento em área basal, Equação (6).

$$\frac{\partial G}{\partial I} = I_i^{-1} G_i [\alpha_0 + \alpha_2 S_i - \ln(G_i)] \quad (6)$$

Com o objetivo de estabelecer um sistema de Equações, para a modelagem da projeção em área basal, a Equação (6) será integrada nos intervalos de G_1 a G_2 e I_1 a I_2 , resultando no modelo (7).

$$\int_{G_1}^{G_2} \partial G = \int_{I_1}^{I_2} \left(\frac{1}{I_i} \right) G_i [\alpha_0 + \alpha_2 S_i - \ln(G_i)] \partial I$$

$$\ln(G_{2i}) = \ln G_{1i} \left(\frac{I_{1i}}{I_{2i}} \right) + \alpha_0 \left(1 - \frac{I_{1i}}{I_{2i}} \right) + \alpha_2 \left(1 - \frac{I_{1i}}{I_{2i}} \right) S_i + \varepsilon_i \quad (7)$$

A função de projeção da área basal (7), juntamente com o modelo volumétrico desenvolvido por Schumacher (1939), Equação (8), são importantes para o estabelecimento de um sistema capaz de estimar (V_2).

$$\ln(V_{2i}) = \beta_0 + \beta_1 \left(\frac{1}{I_{2i}} \right) + \beta_2 S_i + \beta_3 \ln(G_{2i}) + \varepsilon_i \quad (8)$$

O sistema de Clutter é concretizado quando é incorporada na Equação (8) a função (7), originando o modelo (9).

$$\ln(V_{2i}) = \beta_0 + \beta_1 \left(\frac{1}{I_{2i}} \right) + \beta_2 S_i + \beta_3 \left(\frac{I_{1i}}{I_{2i}} \right) \ln(G_{1i}) + \beta_4 \left(1 - \frac{I_{1i}}{I_{2i}} \right) + \beta_5 \left(1 - \frac{I_{1i}}{I_{2i}} \right) S_i + \varepsilon_i \quad (9)$$

Em que: V_{2i} = volume total na idade (I_2) da i -ésima unidade amostral ($m^3 \text{ ha}^{-1}$);
 \ln = logaritmo neperiano; β_s = parâmetros do modelo; S_i = índice de sítio (m) da

i -ésima unidade amostral; I_{1i} e I_{2i} = idades de medição atual e futura da i -ésima unidade amostral, respectivamente (anos); $\hat{\beta}_{4*} = \hat{\beta}_{3*}\hat{\alpha}_0$; $\hat{\beta}_{5*} = \hat{\beta}_{3*}\hat{\alpha}_2$; $\varepsilon_i = \varepsilon_1 + \varepsilon_2$ (erro aleatório).

Como a Equação (9) é dependente da área basal futura, torna-se necessária a modelagem do sistema de Clutter, via método de ajuste em dois estágios. Esse método de ajuste foi realizado por meio do pacote *systemfit* no R.

Ainda, com o intuito de modificação do sistema de Clutter, visando a uma melhor acurácia, foi realizada a adição da variável clone tanto para o modelo de área basal como para o modelo de volume. Interações entre as variáveis bases do modelo foram testadas apenas para o modelo de volume. Para as Equações de área basal e volume modificou-se a relação do índice de sítio - S_i para S_i^{-1} . Tais relações são demonstradas pelas Equações (10), (11) e (12), respectivamente.

$$\ln(V_{2i}) = \beta_0 + \beta_1 \left(\frac{1}{I_{2i}} \right) + \beta_2 \frac{1}{S_i} + \beta_3 \ln(G_{2i}) + \sum_{j=1}^7 \beta_4 \text{Clone}_j + \varepsilon_i \quad (10)$$

$$\ln(V_{2i}) = \beta_0 + \beta_1 I_{2i}^{-1} + \beta_2 S_i^{-1} + \beta_3 \ln(G_{2i}) + \sum_{j=1}^7 \beta_4 \text{Clone}_j + \beta_5 I_{2i}^{-1} S_i^{-1} + \beta_6 I_{2i}^{-1} \ln(G_{2i}) + \beta_7 S_i^{-1} \ln(G_{2i}) + \beta_8 I_{2i}^{-1} S_i^{-1} \ln(G_{2i}) + \varepsilon_i \quad (11)$$

$$\ln(G_{2i}) = \ln G_{1i} \left(\frac{I_{1i}}{I_{2i}} \right) + \alpha_0 \left(1 - \frac{I_{1i}}{I_{2i}} \right) + \alpha_2 \left(1 - \frac{I_{1i}}{I_{2i}} \right) S_i^{-1} + \sum_{j=1}^7 \alpha_3 \text{Clone}_j + \varepsilon_i \quad (12)$$

Onde: $\sum_{j=1}^7 \beta_4$ ou $\alpha_3 \text{Clone}_j = \beta_4$ e $\alpha_3 =$ refere-se ao parâmetro associado ao *j*-ésimo clone e Clone_j é uma variável indicadora (*dummy*) com valor 1 para o *j*-ésimo clone e 0 para os demais.

2.5 Análises das estimativas geradas pela aprendizagem de máquinas e pelos modelos de regressão

Após a configuração de cada topologia de máquina, a mesma será treinada e testada com os respectivos dados de treinamento e teste para estimar a área basal e volume futuro. Tal procedimento será feito também com o método de regressão tradicional, baseado no sistema de Clutter. A qualidade de ajuste dos métodos será realizada com base nas estatísticas: erro médio absoluto (MAE); erro quadrático médio (MSE); no valor da raiz do erro quadrático médio (RMSE); na correlação entre os valores observados e estimados; e pela análise de resíduos.

O MAE e o MSE avaliam o desvio médio absoluto e o desvio quadrático médio entre os valores observados e estimados, respectivamente. Quanto menores os seus valores, melhor é a acurácia das estimativas. O MAE e o MSE são representados pelas Equações 13 e 14.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i| \quad (13) \quad \text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \quad (14)$$

Em que: n = número de observações; \hat{Y}_i = valor estimado da *i*-ésima unidade amostral; Y_i = valor observado da *i*-ésima unidade amostral.

O RMSE avalia o desvio quadrático médio entre os valores observados e os valores estimados. Quanto menor o seu valor, melhor a acurácia da estimativa

(MEHTÄTALO; MALTAMO; KANGAS, 2006). A correlação indica o grau de associação entre os valores observados e estimados. Quanto mais próxima de um, maior é a correlação entre as variáveis. O RMSE (%) e a correlação é obtida pelas Equações 15 e 16, respectivamente.

$$\text{RMSE}(\%) = \frac{100}{\bar{Y}} \sqrt{\frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{n}} \quad (15) \quad r_{Y\hat{Y}} = \frac{\text{cov}(Y, \hat{Y})}{\sqrt{S^2(Y)S^2(\hat{Y})}} \quad (16)$$

Em que: \bar{Y} = média dos valores observados; S^2 = variância; cov = covariância.

3 RESULTADOS E DISCUSSÃO

3.1 Perceptron de múltiplas camadas

3.1.1 Modelagem da área basal

Na Tabela 2, são apresentadas as estimativas paramétricas para o modelo de projeção da área basal.

Tabela 2 - Estimativas paramétricas para o modelo de projeção da área basal.

(Continua)

Área basal (PMC) - RMSE = 2,19m ² ha ⁻¹ (12,9%)		
Neurônio 1		
Parâmetro	Variável Associada	Estimativa
β_{10}	Intercepto (bias)	-2,374437266866
β_{11}	I1	4,905031964
β_{12}	G1	-2,494933072
β_{13}	I2	-1,501498312
β_{14}	S	-0,620862373
β_{15}	H13	0,763656595
β_{16}	H144	1,301510844
β_{17}	H15	0,821506761
β_{18}	H17	0,82589286
β_{19}	H27	4,48011471
β_{110}	H34	-708,1059768
β_{111}	H35	37,335197
Neurônio 2		
β_{20}	Intercepto (bias)	2,018167009
β_{21}	I1	4,486688174
β_{22}	G1	-1,839094651
β_{23}	I2	-1,155746918
β_{24}	S	-0,229385547
β_{25}	H13	-1,021956292
β_{26}	H144	-0,618555997
β_{27}	H15	-1,036457965
β_{28}	H17	-0,968929686
β_{29}	H27	-0,217535496
β_{210}	H34	-1,496987871
β_{211}	H35	-1,888113781

Tabela 2 - Estimativas paramétricas para o modelo de projeção da área basal.

(Conclusão)

Área basal (PMC) - RMSE = 2,19m ² ha ⁻¹ (12,9%)		
Neurônio 1		
Parâmetro	Variável Associada	Estimativa
Neurônio 3		
β_{30}	Intercepto (bias)	0,360027564
β_{31}	I1	2,554002974
β_{32}	G1	0,43789238
β_{33}	I2	-0,626769534
β_{34}	S	-0,262142026
β_{35}	H13	-1,323975118
β_{36}	H144	-1,005274667
β_{37}	H15	-1,327606926
β_{38}	H17	-1,278332625
β_{39}	H27	0,288787569
β_{310}	H34	-1,956144469
β_{311}	H35	-0,586856645
Camada de Saída (G2)		
β_{40}	Intercepto (bias)	1.135777147
β_{41}	Neurônio 1	-0.897796558
β_{42}	Neurônio 2	-1.947027343
β_{43}	Neurônio 3	2.270327828

Fonte: Do autor (2018).

A Equação 17 corresponde ao modelo de projeção da área basal extraído da máquina perceptron de múltiplas camadas (PMC).

$$\begin{aligned}
G_{2i} = & \left\{ 1 / \left[1 + \exp \left(\frac{\beta_{10} + \beta_{11}I_{1i} + \beta_{12}G_{1i} + \beta_{13}I_{2i} + \beta_{14}S_i + \sum_{j=5}^{11} \beta_{1j}Clone_j}{\sum_{j=5}^{11} \beta_{1j}Clone_j} \right) \right] \right\} \beta_{41} + \\
& \left\{ 1 / \left[1 + \exp \left(\frac{\beta_{20} + \beta_{21}I_{1i} + \beta_{22}G_{1i} + \beta_{23}I_{2i} + \beta_{24}S_i + \sum_{j=5}^{11} \beta_{2j}Clone_j}{\sum_{j=5}^{11} \beta_{2j}Clone_j} \right) \right] \right\} \beta_{42} + \\
& \left\{ 1 / \left[1 + \exp \left(\frac{\beta_{30} + \beta_{31}I_{1i} + \beta_{32}G_{1i} + \beta_{33}I_{2i} + \beta_{34}S_i + \sum_{j=5}^{11} \beta_{3j}Clone_j}{\sum_{j=5}^{11} \beta_{3j}Clone_j} \right) \right] \right\} \beta_{43} + \\
& \beta_{40} + \varepsilon_i
\end{aligned} \tag{17}$$

Em que: I_{1i} e I_{2i} = idade presente e futura associada à *i-ésima* unidade amostral (anos); G_{1i} e G_{2i} = área basal presente e futura associada à *i-ésima* unidade amostral ($m^2 ha^{-1}$); S_i = índice de sítio associado à *i-ésima* unidade amostral (m); ε_i = erro aleatório da *i-ésima* unidade amostral. Onde: $\sum_{j=5}^{11} \beta_{xj}Clone_j = \beta_{xj}$ = refere-se ao parâmetro associado ao neurônio x (variando de 1 a 3) e ao *j-ésimo* clone e $Clone_j$ é uma variável indicadora (*dummy*) com valor 1 para o *j-ésimo* clone e 0 para os demais.

Na Tabela 3, são apresentadas as estatísticas de ajuste para os dados de treinamento e de teste da RNA (PMC), na estimação da área basal futura.

Tabela 3 - Estatísticas de ajuste para os dados de treinamento e de teste da rede PMC na estimativa da área basal futura.

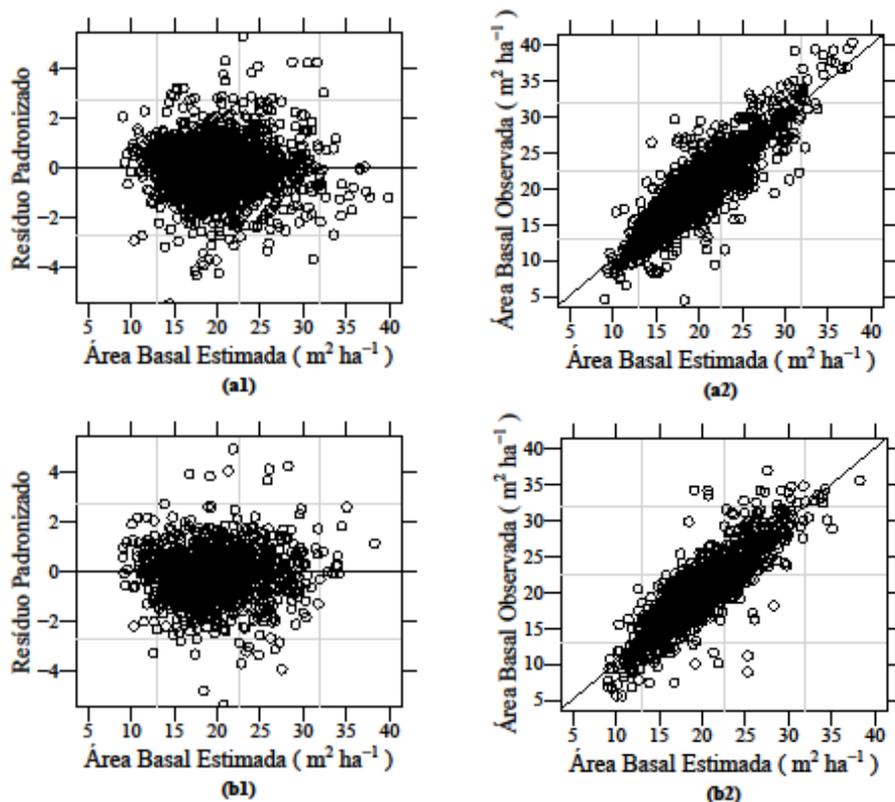
Modelo	Estatísticas de Ajuste			
	MAE	MSE	RMSE	Correlação
Treinamento	1,57	4,81	2,19	0,89
Teste	1,72	5,83	2,41	0,88

Fonte: Do autor (2018).

De posse das estatísticas de ajuste apresentadas na Tabela 3, nota-se uma boa qualidade de treinamento da RNA, com um RMSE de $2,19\text{m}^2 \text{ha}^{-1}$ e uma correlação de 89%. Observa-se também que a rede teve uma boa capacidade de generalização. Essa característica foi confirmada com a proximidade das estatísticas de ajuste para os dados de treinamento e teste, pelos gráficos de resíduos e da relação dos valores de área basal observada e estimada.

Na Figura 5 (a1), (a2), (b1) e (b2), são apresentados os gráficos de dispersão dos resíduos e a relação dos valores observados e estimados da área basal para os dados de treinamento e teste da RNA, respectivamente.

Figura 5 - Dispersão dos resíduos e relação dos valores de área basal estimada e observada para os dados de treinamento (a1 e a2) e teste (b1 e b2).



Fonte: Do autor (2018).

Ao analisar a Figura 5, observa-se uma boa distribuição dos resíduos padronizados, variando entre -2 e 2, tanto para os dados de treinamento (a1) quanto para os dados de teste (b1). Para os dados de treinamento (a1), houve uma subestimativa dos valores de área basal para povoamentos com produção acima de $32\text{m}^2 \text{ ha}^{-1}$. Esse fato pode ser confirmado ao observar a relação dos valores de área basal observados e estimados (a2). A boa qualidade de ajuste pode ser afirmada ao observar a relação dos valores observados e estimados (a2 e b2). Nota-se que a maioria dos valores concentram-se sobre a linha de 45° , indicando um bom ajuste da rede PMC para estimativa da área basal futura.

3.1.2 Modelagem do volume

Na Tabela 4, são apresentadas as estimativas paramétricas para o modelo de projeção do volume.

Tabela 4 - Estimativas paramétricas para o modelo de projeção volumétrica.

(Continua)

Volume (PMC) - RMSE = $29,82\text{m}^3 \text{ ha}^{-1}$ (16,9%)		
Neurônio 1		
Parâmetro	Variável Associada	Estimativa
β_{10}	Intercepto (bias)	-1,382157614
β_{11}	I1	-0,323922929
β_{12}	G1	-0,308922076
β_{13}	V1	1,508141406
β_{14}	I2	0,053719121
β_{15}	S	0,059297628
β_{16}	H13	0,275936509
β_{17}	H144	0,232405908
β_{18}	H15	0,270941993
β_{19}	H17	-0,076465799
β_{110}	H27	0,285424443
β_{111}	H34	0,262272674
β_{112}	H35	0,21558776
β_{113}	G2 _{estimada}	1,513436996

Tabela 4 - Estimativas paramétricas para o modelo de projeção volumétrica.

(Conclusão)

Volume (PMC) - RMSE = 29,82m ³ ha ⁻¹ (16,9%)		
Neurônio 1		
Parâmetro	Variável Associada	Estimativa
Neurônio 2		
β_{20}	Intercepto (bias)	0,700026368
β_{21}	I1	11,34915777
β_{22}	G1	-22,22006448
β_{23}	V1	1,063056768
β_{24}	I2	7,192904862
β_{25}	S	-1,848044947
β_{26}	H13	3,102842987
β_{27}	H144	6,944877296
β_{28}	H15	1,955210085
β_{29}	H17	-5,092147354
β_{210}	H27	4,643198006
β_{211}	H34	5,466863006
β_{212}	H35	4,312676914
β_{213}	G2 _{estimada}	-6,325679173
Neurônio 3		
β_{30}	Intercepto (bias)	-1,65143358
β_{31}	I1	-3,126410004
β_{32}	G1	9,699351206
β_{33}	V1	9,671842711
β_{34}	I2	-10,19936694
β_{35}	S	1,64419055
β_{36}	H13	-0,610881183
β_{37}	H144	-4,091787876
β_{38}	H15	0,416743559
β_{39}	H17	-5,372891524
β_{310}	H27	-1,766662698
β_{311}	H34	-2,462161809
β_{312}	H35	-3,141641453
β_{313}	G2 _{estimada}	2,937476931
Camada de Saída (V2)		
β_{40}	Intercepto (bias)	-0,285204792
β_{41}	Neurônio 1	1,786513062
β_{42}	Neurônio 2	-0,15552491
β_{43}	Neurônio 3	-0,192002264

Fonte: Do autor (2018).

A Equação 18 corresponde ao modelo de projeção do volume extraído da RNA perceptron de múltiplas camadas (PMC).

$$\begin{aligned}
 V_{2i} = & \left\{ 1 / \left[1 + \exp \left(\frac{\beta_{10} + \beta_{11}I_{1i} + \beta_{12}G_{1i} + \beta_{13}V_{1i} + \beta_{14}I_{2i} +}{\beta_{15}S_i + \sum_{j=6}^{12} \beta_{1j}Clone_j + \beta_{113}G_{2iest.}} \right) \right] \right\} \beta_{41} + \\
 & \left\{ 1 / \left[1 + \exp \left(\frac{\beta_{20} + \beta_{21}I_{1i} + \beta_{22}G_{1i} + \beta_{13}V_{1i} + \beta_{24}I_{2i} +}{\beta_{25}S_i + \sum_{j=6}^{12} \beta_{2j}Clone_j + \beta_{213}G_{2iest.}} \right) \right] \right\} \beta_{42} + \\
 & \left\{ 1 / \left[1 + \exp \left(\frac{\beta_{30} + \beta_{31}I_{1i} + \beta_{32}G_{1i} + \beta_{13}V_{1i} + \beta_{34}I_{2i} +}{\beta_{35}S_i + \sum_{j=6}^{12} \beta_{3j}Clone_j + \beta_{313}G_{2iest.}} \right) \right] \right\} \beta_{43} + \\
 & \beta_{40} + \varepsilon_i
 \end{aligned} \tag{18}$$

Em que: V_{1i} e V_{2i} = volume presente e futuro associado à *i-ésima* unidade amostral ($m^3 \text{ ha}^{-1}$); $G_{2iest.}$ = área basal futura associada à *i-ésima* unidade amostral estimada pela RNA ($m^2 \text{ ha}^{-1}$).

Na Tabela 5, são apresentadas as estatísticas de ajuste para os dados de treinamento e teste para a RNA (PMC) na estimativa do volume futuro.

Tabela 5 - Estatísticas de ajuste para os dados de treinamento e teste da rede PMC na estimativa do volume futuro.

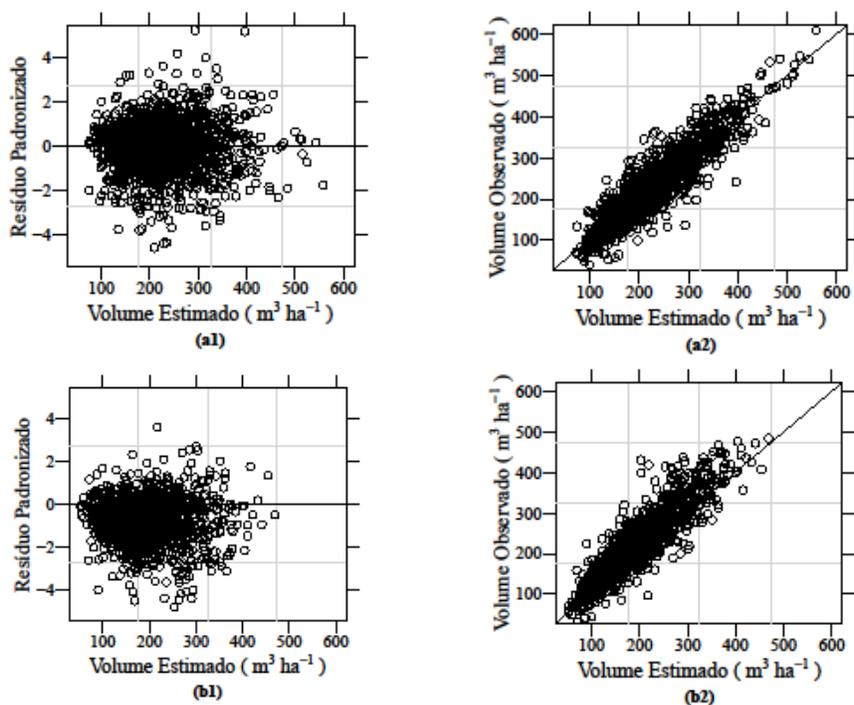
Modelo	Estatísticas de Ajuste			
	MAE	MSE	RMSE	Correlação
Treinamento	22,21	883,85	29,82	0,92
Teste	33,30	1850,80	42,44	0,90

Fonte: Do autor (2018).

Ao observar as estatísticas de ajuste da Tabela 5, nota-se uma boa qualidade de treinamento da RNA, com um RMSE de $29,82\text{m}^3\text{ ha}^{-1}$ e uma correlação de 92%. Observa-se também que a rede não teve uma boa capacidade de generalização, com aumento do RMSE de 29,82 para $42,44\text{m}^3\text{ ha}^{-1}$ (42,32%) em comparação aos dados de treinamento. A incapacidade de generalização pode ser confirmada pelos gráficos de resíduos e da relação dos valores de volume observado e estimado (FIGURA 6).

Na Figura 6 (a1), (a2), (b1) e (b2), são apresentados os gráficos de dispersão dos resíduos e a relação dos valores observados e estimados do volume para os dados de treinamento e teste da RNA, respectivamente.

Figura 6 - Dispersão dos resíduos e relação dos valores de volume estimado e observado para os dados de treinamento (a1 e a2) e teste (b1 e b2).



Fonte: Do autor (2018).

Para o conjunto de dados de treinamento é observado em (a1) um bom ajuste da rede, com os resíduos bem distribuídos, variando entre -2 e 2. Nota-se uma tendência de subestimar a produção de parcelas com volume maior do que $450\text{m}^3 \text{ ha}^{-1}$. A análise gráfica de (a2) corrobora com a boa distribuição dos resíduos em (a1), uma vez que é observada uma concentração da relação volume estimado e observado sobre a linha de 45° . O mesmo não se pode afirmar acerca dos dados de teste. No gráfico de resíduos (b1), é notável uma tendência de subestimar os valores de volume em todas as classes de produção, o que é confirmado em (b2).

3.2 Redes de Função de Base Radial (RBF)

3.2.1 Modelagem da área basal

Na Tabela 6, são apresentadas as estatísticas de ajuste para os dados de treinamento e de teste da rede RBF na estimativa da área basal futura.

Tabela 6 - Estatísticas de ajuste para os dados de treinamento e de teste da rede RBF na estimativa da área basal futura.

Modelo	Estatísticas de Ajuste			
	MAE	MSE	RMSE	Correlação
Treinamento	2,15	8,23	2,87	0,80
Teste	2,10	8,56	2,93	0,82

Fonte: Do autor (2018).

Com base nas estatísticas de ajuste apresentadas na Tabela 6, nota-se uma boa qualidade de treinamento da RNA, com um RMSE de $2,87\text{m}^2 \text{ ha}^{-1}$ e uma correlação de 80%. Observa-se também que a rede teve uma boa capacidade de generalização, pela proximidade das estatísticas de ajuste para os dados de

treinamento e teste, pelos gráficos de resíduos e da relação dos valores de área basal observada e estimada.

Ao comparar os valores das estatísticas produzidas pelas redes RBF e PMC é observada uma proximidade dos valores de RMSE ($2,87$ e $2,93\text{m}^2 \text{ha}^{-1}$) e ($2,19$ e $2,41\text{m}^2 \text{ha}^{-1}$) para os dados de treinamento e teste para ambas as redes testadas, respectivamente. Para as estatísticas MAE e MSE é observado um aumento de, aproximadamente, $36,9\%$ e $71,1\%$, respectivamente. Com relação aos valores de correlação, houve uma redução de aproximadamente 10% . Essa comparação dos valores de estatísticas demonstram uma inferioridade da rede RBF em comparação à rede PMC.

Na Figura 7 (a1), (a2), (b1) e (b2), são apresentados os gráficos de dispersão dos resíduos e a relação dos valores observados e estimados da área basal para os dados de treinamento e teste da rede RBF, respectivamente.

Figura 7 - Dispersão dos resíduos e relação dos valores de área basal estimada e observada para os dados de treinamento (a1 e a2) e teste (b1 e b2) da rede RBF.

(Continua)

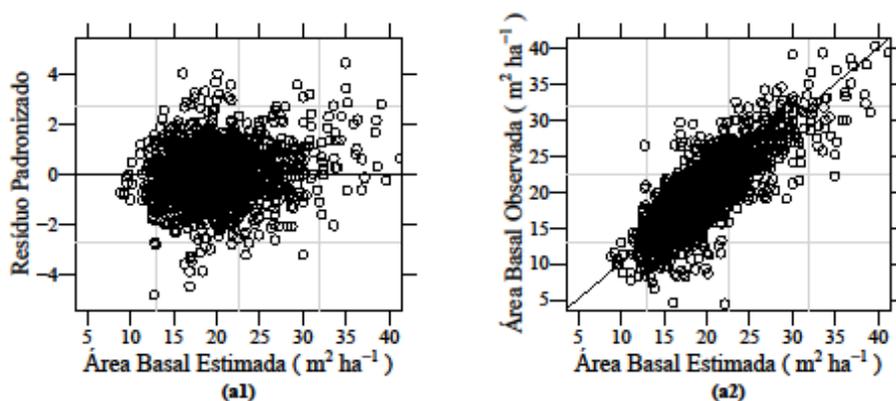
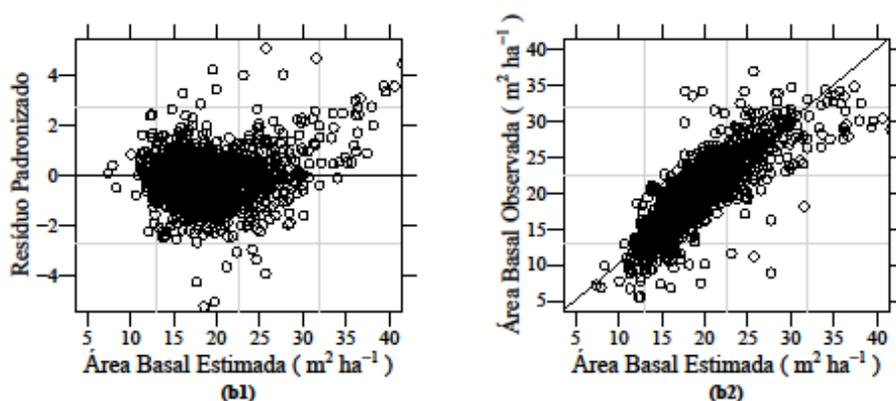


Figura 7 - Dispersão dos resíduos e relação dos valores de área basal estimada e observada para os dados de treinamento (a1 e a2) e teste (b1 e b2) da rede RBF.

(Conclusão)



Fonte: Do autor (2018).

Ao analisar a Figura 7, observa-se uma boa distribuição dos resíduos, variando entre -2 e 2, tanto para os dados de treinamento (a1) quanto para os dados de teste (b1). Ao contrário da rede PMC, observou-se uma superestimação dos dados de treinamento e teste para G2 acima de $32 \text{m}^2 \text{ha}^{-1}$ (a1 e b1). Esse fato pode ser confirmado ao observar a relação dos valores de área basal observados e estimados (a2 e b2). De maneira geral, é possível afirmar que houve uma boa qualidade de ajuste para estimativa da área basal futura.

3.2.2 Modelagem do Volume

Na Tabela 7, são apresentadas as estatísticas de ajuste para os dados de treinamento e de teste para a rede RBF na estimativa do volume futuro.

Tabela 7 - Estatísticas de ajuste para os dados de treinamento e teste da rede RBF na estimativa do volume futuro.

Modelo	Estatísticas de Ajuste			
	MAE	MSE	RMSE	Correlação
Treinamento	32,04	1746,80	41,80	0,83
Teste	39,42	2610,84	51,10	0,83

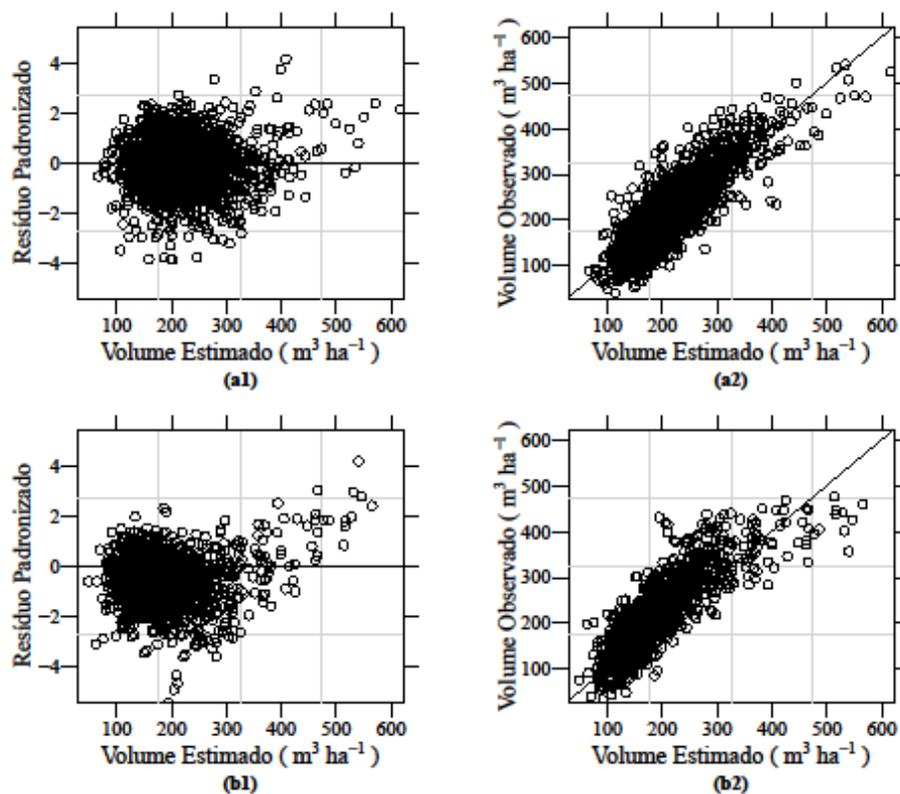
Fonte: Do autor (2018).

Com base nas estatísticas apresentadas na Tabela 7, nota-se uma perda de acurácia da rede RBF em comparação à rede PMC. Verifica-se um aumento no valor de RMSE de 40,17% (29,82 para 41,80m³ ha⁻¹) para os dados de treinamento e 20,40% (42,44 para 51,10m³ ha⁻¹) para os dados de teste. As estatísticas MAE, MSE e correlação tiveram uma perda de acurácia de 44,26, 97,63 e 9,78%, para os dados de treinamento, respectivamente.

Além da perda de acurácia verificada, é possível afirmar que a rede não teve uma boa capacidade de generalização, com aumento do RMSE de 41,80 para 51,10m³ ha⁻¹ (22,25%) em comparação aos dados de treinamento. Esse fato corrobora com os gráficos de resíduos e com a relação dos valores de volume observado e estimado (FIGURA 8).

Na Figura 8 (a1), (a2), (b1) e (b2), são apresentados os gráficos de dispersão dos resíduos e a relação dos valores observados e estimados do volume para os dados de treinamento e teste da rede RBF, respectivamente.

Figura 8 - Dispersão dos resíduos e relação dos valores de volume estimado e observado para os dados de treinamento (a1 e a2) e teste (b1 e b2) da rede RBF.



Fonte: Do autor (2018).

No treinamento, é observado (a1) um bom ajuste da rede, com os resíduos bem distribuídos, variando entre -2 e 2. Diferentemente da rede PMC, verifica-se uma tendência de superestimação nas parcelas com volume maior do que 450m³ ha⁻¹. A análise gráfica de (a2) corrobora com a boa distribuição dos resíduos em (a1). O mesmo não se pode afirmar para os dados de teste. No gráfico de resíduos (b1) é notável uma tendência de subestimar o volume em todas as classes de produção, o que é confirmado em (b2).

3.3 Máquina de vetor de suporte (MVS)

3.3.1 Modelagem da área basal

Na Tabela 8, são apresentadas as estatísticas de ajuste para os dados de treinamento e de teste da rede MVS na estimativa da área basal futura.

Tabela 8 - Estatísticas de ajuste para os dados de treinamento e de teste da rede MVS na estimativa da área basal futura.

Modelo	Estatísticas de Ajuste			
	MAE	MSE	RMSE	Correlação
Treinamento	1,46	4,36	2,09	0,90
Teste	1,72	5,94	2,44	0,87

Fonte: Do autor (2018).

Na Tabela 8, observa-se uma boa qualidade de treinamento da MVS, com um RMSE de $2,09\text{m}^2 \text{ ha}^{-1}$ e uma correlação de 90%. É verificada uma boa capacidade de generalização da MVS, pelo fato da proximidade das estatísticas de ajuste para os dados de treinamento e teste, pelos gráficos de resíduos e da relação dos valores de área basal observada e estimada.

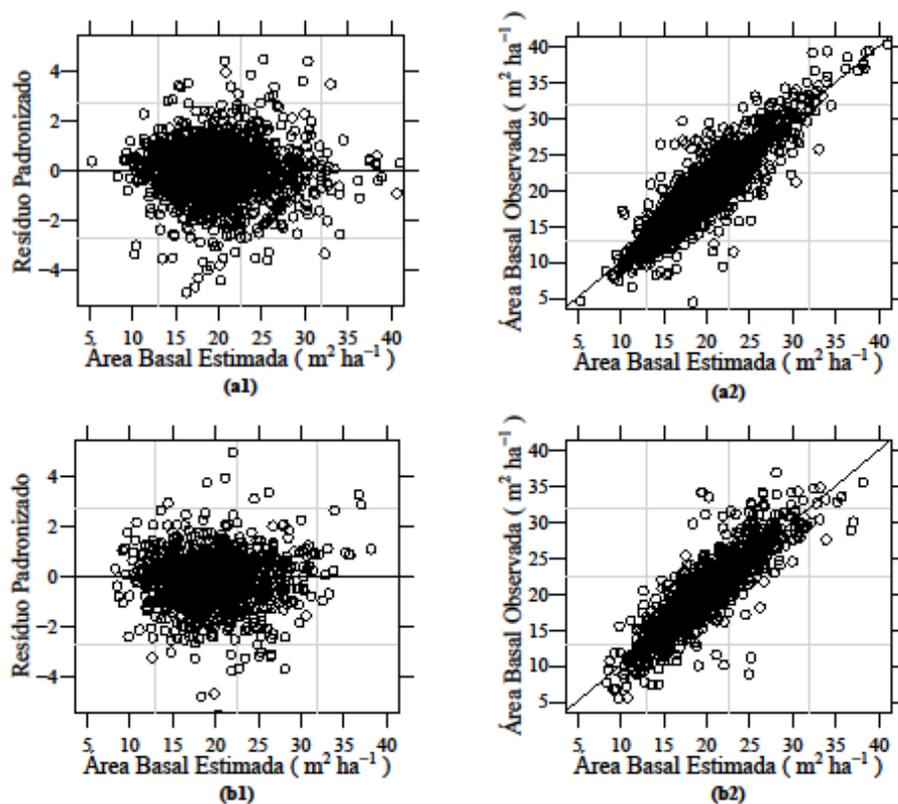
A partir da comparação entre MVS e as redes PMC e RBF, é condizente afirmar, com base nas estatísticas produzidas, que a rede MVS é ligeiramente superior às demais. Haja vista a superioridade da rede PMC frente à RBF e comparando os valores das estatísticas produzidas pelas redes MVS e PMC, observa-se uma proximidade dos valores de RMSE ($2,09$ e $2,44\text{m}^2 \text{ ha}^{-1}$) e ($2,19$ e $2,41\text{m}^2 \text{ ha}^{-1}$) para os dados de treinamento e teste para ambas as técnicas analisadas, respectivamente.

Para as estatísticas MAE e MSE, é observada uma redução de aproximadamente 7,0% e 9,3%, respectivamente. Com relação aos valores de correlação houve um ligeiro aumento de 1,1% (de 0,89 para 0,90). Com base na

análise dos valores das estatísticas referentes aos dados de teste, é condizente afirmar que a rede MVS possui uma menor capacidade de generalização em relação à rede PMC.

Na Figura 9 (a1), (a2), (b1) e (b2), são apresentados os gráficos de dispersão dos resíduos e a relação dos valores observados e estimados da área basal para os dados de treinamento e teste da rede MVS, respectivamente.

Figura 9 - Dispersão dos resíduos e relação dos valores de área basal estimada e observada para os dados de treinamento (a1 e a2) e teste (b1 e b2).



Fonte: Do autor (2018).

Analisando a Figura 9, observa-se uma boa distribuição dos resíduos, variando entre -2 e 2, tanto para os dados de treinamento (a1) quanto para os dados de teste (b1). A boa qualidade de ajuste pode ser afirmada, ao observar a relação dos valores observados e estimados (a2 e b2). Nota-se que a maioria dos valores se concentram sobre a linha de 45°, indicando um bom ajuste da MVS para estimar a área basal futura.

3.3.2 Modelagem do volume

Na Tabela 9, são apresentadas as estatísticas de ajuste para os dados de treinamento e de teste para a rede MVS na estimativa do volume futuro.

Tabela 9 - Estatísticas de ajuste para os dados de treinamento e de teste da rede MVS na estimativa do volume futuro.

Modelo	Estatísticas de Ajuste			
	MAE	MSE	RMSE	Correlação
Treinamento	20,31	761,40	27,59	0,93
Teste	33,17	1841,13	42,90	0,90

Fonte: Do autor (2018).

De posse das estatísticas apresentadas na Tabela 9, e com a informação da superioridade da rede PMC frente a RBF, é confirmada uma superioridade da MVS em relação as demais.

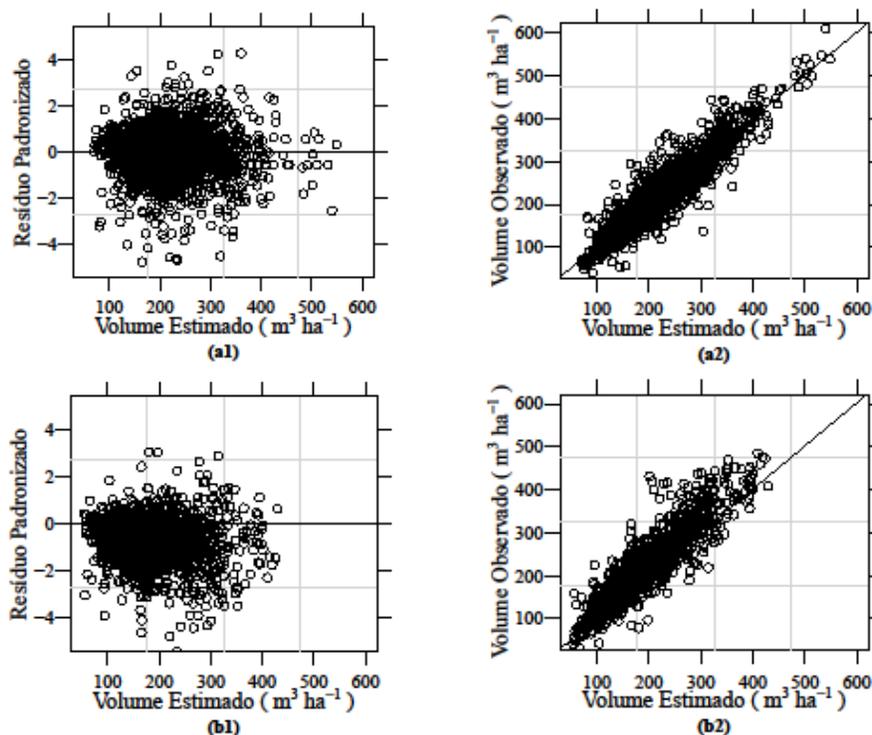
Quando se compara MVS e PMC, observa-se um decréscimo no valor de RMSE de 7,5% (29,82 para 27,59m³ ha⁻¹) para os dados de treinamento. Ao se comparar o valor das estatísticas para os dados de teste é notável uma similaridade dos valores. Essa similaridade dos valores remete à baixa eficiência da rede MVS em obter boa generalização para os dados de teste. Esse fato também foi verificado na projeção da área basal (G2). As estatísticas MAE,

MSE e correlação tiveram um ganho de acurácia de 8,5, 13,8 e 1,1%, para os dados de treinamento e teste, respectivamente.

Ao analisar a Tabela 9, é possível afirmar que a MVS não teve uma boa capacidade de generalização, com aumento do RMSE de 27,59 para $42,90\text{m}^3\text{ ha}^{-1}$ (55,5%). Esse fato corrobora com os gráficos de resíduos e com a relação dos valores de volume observado e estimado (FIGURA 10).

Na Figura 10 (a1), (a2), (b1) e (b2), são apresentados os gráficos de dispersão dos resíduos e a relação dos valores observados e estimados do volume para os dados de treinamento e de teste da rede MVS, respectivamente.

Figura 10 - Dispersão dos resíduos e relação dos valores de volume estimado e observado para os dados de treinamento (a1 e a2) e de teste (b1 e b2) da rede MVS.



Fonte: Do autor (2018).

No treinamento, é observado em (a1) um bom ajuste da rede, com os resíduos bem distribuídos, variando entre -2 e 2. A análise gráfica de (a2) corrobora com a boa distribuição dos resíduos em (a1). O mesmo não se pode afirmar acerca dos dados de teste. No gráfico de resíduos (b1), é notável uma tendência de subestimativa dos valores de volume em todas as classes de produção, o que é confirmado em (b2). Esse fato é observado para todas as redes testadas.

Diante dos resultados obtidos na projeção da produção em área basal e volume, nota-se que as redes PMC e MVS foram as que apresentaram as melhores estatísticas de treinamento. Porém, ambas apresentaram incapacidade de obter uma boa generalização na projeção volumétrica.

Para a rede PMC, a incapacidade de generalização poderá estar relacionada com a definição da topologia da rede. Segundo Silva, Spatti e Flauzino (2010), a capacidade de generalização de uma RNA está diretamente ligada à definição de uma boa topologia, a qual está relacionada com a complexidade do problema. Porém, destaca-se que várias topologias foram testadas, variou-se o número de neurônios até dez na camada intermediária e aumentou-se, para duas, o número de camadas ocultas. Houve ganho de acurácia pouco expressivo, na medida em que se aumenta o número de neurônios da camada oculta. Mesmo o ganho de acurácia sendo pouco expressivo, talvez o aumento do número de neurônios da camada intermediária melhoraria a capacidade de generalização da rede PMC. Para esse caso, existe a possibilidade de novas pesquisas com o intuito de avaliar qual a melhor topologia para a projeção da produção volumétrica utilizando técnicas de inteligência artificial PMC.

Apesar da rede MVS apresentar as melhores estatísticas de acurácia para a área basal e volume para os dados de treinamento, observou-se uma baixa capacidade de generalização para os dados de teste, atingindo valores bem

próximos dos obtidos pela rede PMC. Esse melhor desempenho para os dados de treinamento se deve ao fato de o número de vetores de suporte e, conseqüentemente, o número de funções de kernels (neurônios) serem definidos automaticamente pelo algoritmo de treinamento. Além disso, pelo fato de apresentar uma topologia com apenas uma camada intermediária, a MVS minimiza a função de perda (erro), o que permite atingir o ótimo global.

O fato da MVS apresentar fundamentações matemáticas que permitem a escolha de funções com bom desempenho e, conseqüentemente, consegue maximizar a generalização, não foram suficientes para alcançar uma boa generalização. Além disso, a aplicação da técnica *cross-validation* também não foi suficiente para obter-se uma boa generalização e tampouco necessária, quando relacionado ao tempo de processamento. A aplicação da *cross-validation* possibilita um desempenho mais lento da MVS, o que leva a um aumento substancial do tempo de processamento.

O desempenho inferior das redes RBF também pode estar relacionado com a topologia definida. Assim como na rede PMC testada, preocupou-se em variar o número de neurônios com função de ativação gaussiana na camada intermediária. Com o mesmo número de neurônios (3) definido na rede PMC, percebeu-se uma instabilidade da rede em obter boa acurácia para os dados de treinamento. Na medida em que se aumenta o número de neurônios (máximo de 20) percebia-se uma ligeira melhora da estabilidade dos resíduos, ou seja, um bom comportamento visual do gráfico de resíduos. Ao se atingir um bom ajuste para os dados de treinamento, procedeu-se com a generalização para os dados de teste. Nessa ocasião, observou-se que a rede não obteve uma boa generalização, apresentando o mesmo comportamento das demais redes testadas para a projeção volumétrica.

Binoti et al. (2015) ao empregar RNAs na prognose da produção em área basal e volume constataram boa capacidade de treinamento e generalização

quando fez-se o uso da rede PMC e RBF. Neste estudo, os pesquisadores variaram o número de neurônios na camada intermediária, até três, para as redes PMC. Para as redes RBF, na estimativa de G2 o número de neurônios variou até 197 e na estimativa de V2 até 966. Nesse caso, há indícios de que a baixa qualidade de ajuste da rede RBF para a projeção de V2 esteja ligada ao reduzido número de neurônios.

3.4 Sistema de Equações de Clutter

Na Tabela 10, são apresentadas as estatísticas de ajuste e as estimativas paramétricas para o sistema de Clutter usual, modificado e o modificado com interações.

Tabela 10 - Estimativas paramétricas e estatísticas de ajuste para o sistema de Clutter usual, modificado e modificado com interações.

(Continua)

Clutter usual - (Syx = 24,14m ³ ha ⁻¹ - 11,10%)					
Parâmetro	Variável Associada	Estimativa	Erro Padrão	t-valor	p-valor
β0	Intercepto	1,323727238	0,05	25,36	< 0,001
β1	I2 ⁻¹	-0,822384615	0,08	- 10,81	< 0,001
β2	IS	0,002705363	0,00	5,24	< 0,001
β3	Ln(G2)	1,383308030	0,01	95,35	< 0,001
Clutter modificado - (Syx = 23,20m ³ ha ⁻¹ - 10,66%)					
β0	Intercepto Clone - H13	1,41033957	0,06	24,75	< 0,001
β1	I2 ⁻¹	-0,67641347	0,08	-8,69	< 0,001
β2	IS ⁻¹	-0,99632237	0,34	-2,96	0,003
β3	Ln(G2)	1,37761641	0,01	93,99	< 0,001
β4	H144	0,02686569	0,01	2,60	< 0,001
β5	H15	-0,02849327	0,01	-3,39	< 0,001
β6	H17	0,05985383	0,01	6,28	< 0,001
β7	H27	0,04247731	0,01	4,15	< 0,001
β8	H35	0,07758379	0,01	6,11	< 0,001

Tabela 10 - Estimativas paramétricas e estatísticas de ajuste para o sistema de Clutter usual, modificado e modificado com interações.

(Conclusão)

Clutter usual - ($S_{yx} = 24,14\text{m}^3 \text{ha}^{-1}$ - 11,10%)					
Clutter modificado com interação - ($S_{yx} = 22,85\text{m}^3 \text{ha}^{-1}$ - 10,50%)					
β_0	Intercepto	10,44692782051	0,59	17,75	< 0,001
	Clone - H13				
β_1	I_2^{-1}	-38,67431291946	2,72	-14,23	< 0,001
β_2	IS^{-1}	-216,66809081331	13,73	-15,78	< 0,001
β_3	$\text{Ln}(G_2)$	-1,59503735306	0,20	-8,07	< 0,001
β_4	H15	-0,03097669708	0,01	-4,14	< 0,001
β_5	H17	0,06463241407	0,01	7,51	< 0,001
β_6	H35	0,06136402137	0,01	5,41	< 0,001
β_7	$I_2^{-1}:IS^{-1}$	942,08814392576	61,37	15,35	< 0,001
β_8	$I_2^{-1}:\text{Ln}(G_2)$	12,57931958882	0,93	13,58	< 0,001
β_9	$IS^{-1}:\text{Ln}(G_2)$	71,65939687606	4,63	15,47	< 0,001
β_{10}	$I_2^{-1}:IS^{-1}:\text{Ln}(G_2)$	-315,50076730451	21,05	-14,98	< 0,001

Fonte: Do autor (2018).

De acordo com as estatísticas apresentadas, na Tabela 10, observa-se que todos os parâmetros dos modelos, baseados no sistema de Clutter, para projeção volumétrica foram significativos ($p < 0,05$). Tomando como base o critério paramétrico, ou seja, sua significância, e a estatística erro padrão (RMSE), pode-se dizer que todos os três modelos testados podem ser utilizados para projeção do volume. Nota-se, na Tabela 11, que o modelo de Clutter modificado com interação apresentou o menor RMSE, $22,85\text{m}^3 \text{ha}^{-1}$ (10,50%), seguido pelo modelo modificado $23,20\text{m}^3 \text{ha}^{-1}$ (10,66%) e pelo modelo usual $24,14\text{m}^3 \text{ha}^{-1}$ (11,10%).

Tabela 11- Estatísticas de ajuste para os dados de treinamento e teste do sistema de Clutter usual, modificado e modificado com interações.

	Modelo	Estatísticas de Ajuste			
		MAE	MSE	RMSE	Correlação
Treinamento	Usual (V2)	16,28	582,84	24,14	0,95
	Modificado (V2)	15,72	538,50	23,20	0,95
	Mod. com interações (V2)	16,34	522,36	22,85	0,95
Teste	Usual (V2)	16,71	544,58	23,34	0,95
	Modificado (V2)	16,12	513,61	22,66	0,96
	Mod. com interações (V2)	17,23	547,18	23,39	0,95

Fonte: Do autor (2018).

Com relação aos dados de teste, observa-se uma boa capacidade dos modelos baseados no sistema de Clutter em predizer o volume futuro dos povoamentos de eucalipto. Isso é confirmado pela proximidade dos valores das estatísticas MAE, MSE, RMSE e correlação para os dados de treinamento e teste. Uma ligeira melhora foi verificada nos valores da estatística RMSE para o modelo de Clutter usual ($23,34\text{m}^3 \text{ ha}^{-1}$) e modificado ($22,66\text{m}^3 \text{ ha}^{-1}$), em relação aos dados de treinamento.

Como observado na Tabela 12, os valores do coeficiente de determinação ajustado foram 87% para o modelo de Clutter usual; 88% para o modelo modificado e 90% para o modelo modificado com interações. Nesse contexto, pode-se dizer que a adição do clone como variável *dummy* e as interações testadas nos modelos alternativos ao tradicional (usual), proporcionaram uma leve melhora, com redução do RMSE e aumento do $R^2_{\text{ajust.}}$.

Valores próximos da estatística RMSE (%) e coeficiente de determinação ajustado foram encontrados por Silva (2017), ao testar diferentes relações funcionais do sistema de Clutter em povoamentos de eucaliptos

desbastados e não desbastados. Neste trabalho, os valores de RMSE (%) e R^2_{ajust} , variaram entre 8,58% e 9,05%, e 58,07% e 97,63%, respectivamente.

Quando se levou em consideração a consistência paramétrica para a seleção do modelo, apenas os modelos usual e modificado apresentaram os sinais paramétricos em conformidade com as relações das variáveis independentes, o que possibilita que se afirme, portanto, que apresentaram consistência biológica. Apesar de o modelo modificado com interações apresentar os melhores valores das estatísticas RMSE e de seleção (TABELA 11 e 12), foi o que obteve inconsistência biológica nos parâmetros relacionados ao $\ln(G_2)$ e nas interações: $I_2^{-1}:IS^{-1}$, $I_2^{-1}:\ln(G_2)$, $IS^{-1}:\ln(G_2)$ e $I_2^{-1}:IS^{-1}:\ln(G_2)$.

Na Tabela 12, são apresentadas as estatísticas de seleção dos modelos de Clutter usual, modificado e modificado com interações.

Tabela 12 - Estatísticas de seleção dos modelos testados: Clutter usual, modificado e modificado com interações.

Modelo	Estatísticas de Seleção			
	CIA	CIB	Lnmv	R^2_{ajust} .
Clutter Usual	-5914,95	-5870,57	2964,48	0,87
Clutter Modificado	-5870,40	-5756,29	2953,20	0,88
Clutter Modificado com Interações	-6085,77	-5977,99	3059,88	0,90

Fonte: Do autor (2018).

De acordo com a Tabela 12, observa-se que os modelos apresentaram valores de estatísticas de seleção próximos. Dada a inconsistência biológica do modelo de Clutter modificado com interações, e quando comparados os modelos de Clutter usual e modificado, nota-se uma leve superioridade do modelo modificado em relação ao usual. Tal fato pode ser comprovado pelo menor valor de RMSE ($23,20\text{m}^3\text{ ha}^{-1}$) e maior valor de R^2_{ajust} . (88%). Os critérios de seleção CIA, CIB, Lnmv não corroboram com as estatísticas RMSE e R^2_{ajust} .

Considerando esses critérios, o modelo usual é ligeiramente superior ao modelo modificado.

Na Figura 11, são apresentados os gráficos de resíduos e a relação volume observado e estimado para os modelos testados.

Figura 11 - Resíduos padronizados e relação volume observado e estimado para o conjunto de dados de ajuste para os modelos de Clutter usual (a), modificado (b) e modificado com interações (c).

(Continua)

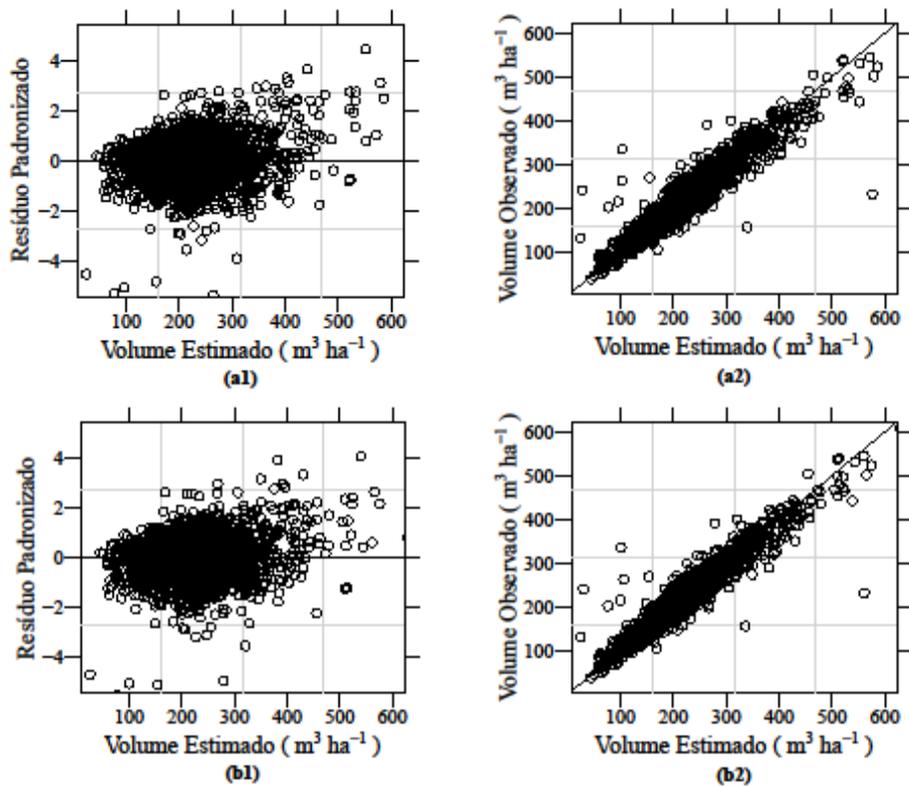
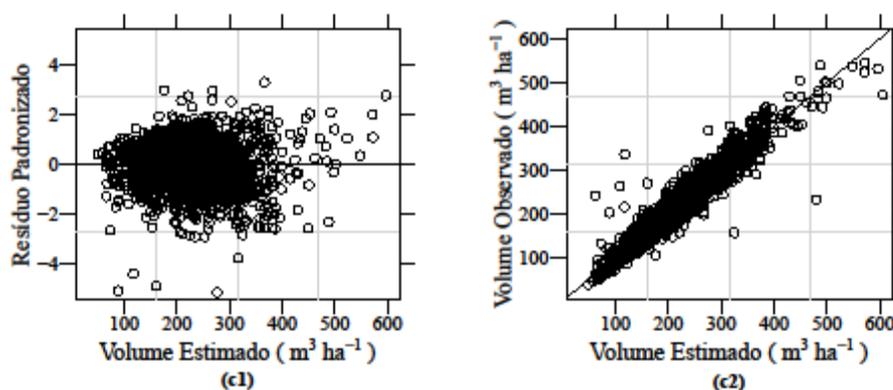


Figura 11 - Resíduos padronizados e relação volume observado e estimado para o conjunto de dados de ajuste para os modelos de Clutter usual (a), modificado (b) e modificado com interações (c).

(Conclusão)

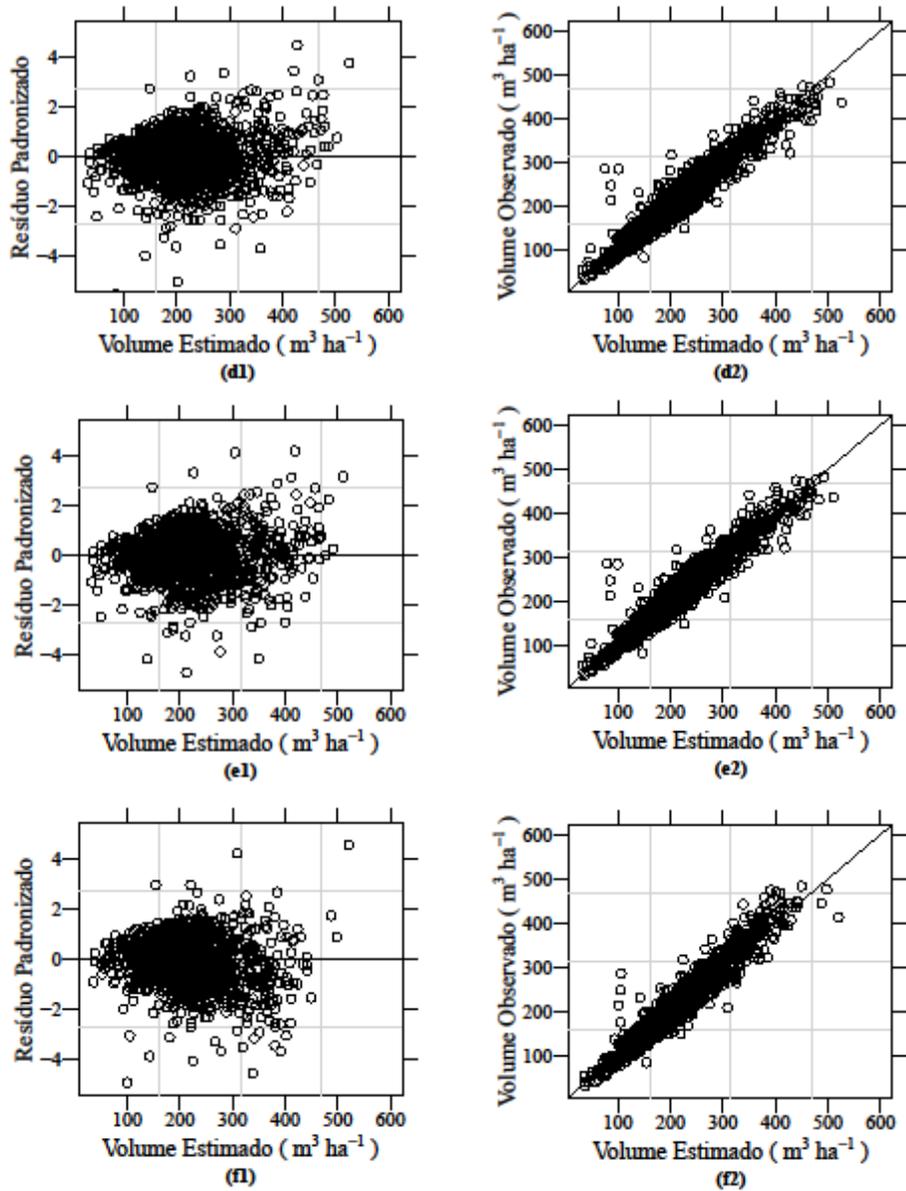


Fonte: Do autor (2018).

Analisando a Figura 11, é possível observar uma boa distribuição dos resíduos em todos os modelos testados (a1, b1 e c1). Para os valores de volume maior do que $500m^3 ha^{-1}$ houve superestimativa da produção. Esse fato está relacionado à ocorrência de um número reduzido de unidades amostrais com produção igual ou maior do que $500m^3 ha^{-1}$ na base de dados utilizada para o ajuste. Os gráficos (a2, b2 e c2) corroboram com a análise de resíduos. Verifica-se um bom ajuste dos modelos testados, uma vez que é observada uma concentração da relação volume estimado e observado sobre a linha de 45° .

Na Figura 12, apresentam-se os gráficos de resíduos e a relação de valores de volume estimados e observados dos modelos testados para os dados de teste.

Figura 12 - Resíduos padronizados e relação volume observado e estimado para o conjunto de dados de teste para os modelos de Clutter usual (d), modificado (e) e modificado com interações (f).



Fonte: Do autor (2018).

Apesar da boa distribuição dos resíduos, é possível notar indícios de uma leve heterocedasticidade da variância nos sistemas Clutter usual (d1), modificado (e1) e modificado com interações (f1). Tais indícios podem comprometer toda a inferência da regressão. Observa-se, nos gráficos (d2, e2 e f2), uma boa qualidade de predição, com concentração dos volumes estimados e observados sobre a linha de 45°.

Os modelos testados referentes ao sistema de Clutter, para os dados de treinamento e teste, foram superiores em todas as estatísticas de ajuste (MAE, MSE, RMSE e correlação) quando comparados à técnica de inteligência artificial (PMC, RBF e MVS). Para os dados de treinamento, os modelos referentes ao sistema de Clutter proporcionaram uma redução média de 21,5% do RMSE, em relação à rede PMC. Já em comparação à MVS a redução foi menor, com média de 15,2%. O modelo modificado com interações foi o mais preciso, seguido pelo modelo modificado e pelo modelo usual, tanto para a rede PMC quanto para a MVS.

Para os dados de teste, os modelos testados de Clutter apresentaram uma superioridade em torno de 45% em relação à técnica de I.A., utilizando a rede PMC e a MVS. Os modelos de Clutter obtiveram estatísticas de ajuste muito próximas. Portanto, não apresentaram diferença significativa de superioridade entre os ajustes.

4 CONCLUSÃO

As RNAs testadas para projeção da área basal, proporcionaram boa capacidade de treinamento e generalização. Para projeção volumétrica as redes apresentaram baixa capacidade de generalização.

A aplicação da rede MVS proporcionou melhoria na acurácia da estimativa da área basal e volume quando comparada às redes PMC e RBF.

Os modelos testados referentes ao sistema de Clutter, para os dados de treinamento e teste, foram superiores, quando comparados à técnica de inteligência artificial na estimativa do volume.

Nos modelos alternativos ao sistema de Clutter, a adição do material genético como variável *dummy* e as interações entre as variáveis independentes, proporcionaram uma leve melhora na precisão dos modelos em relação ao modelo usual.

Apenas o modelo usual e o modificado apresentaram relações paramétricas consistentes.

A obtenção de estimativas da área basal e volume foi possível com a extração dos modelos da rede PMC. Esses modelos podem ser aplicados em qualquer povoamento florestal de mesmo material genético, sem a necessidade do treinamento constante das redes, quando estas apresentarem boa capacidade de generalização.

REFERÊNCIAS

- BINOTI, M. L. M. da S. et al. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, Lavras, v. 21, n. 1, p. 97-105, jan./mar. 2015.
- BINOTI, M. L. M. da S.; BINOTI, D. H. B.; LEITE, H. G.. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, Viçosa, v. 37, n. 4, p. 639-645, jul./ago. 2013.
- BRAGA, A. P.; CARVALHO, A. P. de L. F. de; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2011. 226 p.
- CASTELLANOS, A.; BLANCO, A. M.; PALENCIA, V. Applications of radial basis neural networks for area forest. **International Journal Information Theories and Applications**, Piscataway, v. 14, p. 218-222, 2007.
- CLUTTER, J. L. Compatible growth and yield models for loblolly pine. **Forest Science**, Bethesda, v. 9, n. 3, p. 354-371, Sept. 1963.
- CLUTTER, J. L. et al. **Timber management: a quantitative approach**. New York: Wiley & Sons, 1983. 333 p.
- CORDEIRO, M. A. et al. Estimativa do volume de *Acacia mangium* utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. **Pesquisa Florestal Brasileira**, Colombo, v. 35, n. 83, p. 255-261, jul./set. 2015.
- DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. **Computers and Electronics in Agriculture**, New York, v. 48, n. 3, p. 235-244, Sept. 2005.
- FAUSETT, L. **Fundamentals of neural networks: architectures, algorithms and applications**. New York: Prentice Hall, 1994. 461 p.
- GORGENS, E. B. et al. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, Viçosa, v. 33, n. 6, p. 1141-1147, nov./dez. 2009.
- _____. Influência da arquitetura na estimativa de volume de árvores individuais por meio de redes neurais artificiais. **Revista Árvore**, Viçosa, v. 38, n. 2, p. 289-295, mar./abr. 2014.

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

HUBER, P. J. Robust estimation of a location parameter. **Annals of Mathematical Statistics**, Ann Arbor, v. 35, p. 73-101, 1964.
_____. **Robust statistics**. New York: Wiley & Sons, 1981. 380 p.

MEHTÄTALO, L.; MALTAMO, M.; KANGAS, A. The use of quantile trees in the prediction of the diameter distribution of a stand. **Silva Fennica**, Helsinki, v. 40, n. 3, p. 501-516, Jan. 2006.

NUNES, M. H.; GORGENS, E. B. Artificial intelligence procedures for tree taper estimation within a complex vegetation mosaic in Brazil. **PLoS One**, San Francisco, v. 11, n. 5, p. 16, May 2016.

ÖZÇELİK, R. et al. Estimating crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, Amsterdam, v. 306, p. 52-60, Oct. 2013.

_____. Estimating tree bole volume using artificial neural network models for four species in Turkey. **Journal of Environmental Management**, London, v. 91, n. 3, p. 742-753, Jan./Feb. 2010.

PENG, C.; WEN, X. Recent applications of artificial neural networks in forest resource management: an overview. In: AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE WORKSHOP, 1999, Orlando. **Proceedings...** Orlando: SAS, 1999. p. 1-8.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations by error propagation. In: McCLELLAND, J. L.; RUMELHART, D. E.; HINTON, G. E. **Parallel distributed processing**. Cambridge: MIT Press, 1986. v. 1, chap. 8, p. 318-362.

SANDO, T. et al. Advantages and disadvantages of different crash modeling techniques. **Journal of Safety Research**, Chicago, v. 36, n. 5, p. 485-487, 2005.

SCHIKOWSKI, A. B.; CORTE, A. P. D.; SANQUETTA, C. R. Estudo da forma do fuste utilizando redes neurais artificiais e funções de afilamento. **Pesquisa Florestal Brasileira**, Colombo, v. 35, n. 82, p. 119-127, abr./jun. 2015.

SCHUMACHER, F. X. A new growth curva and its application to timber-yield studies. **Journal of Forest**, Bethesda, v. 37, n. 4, p. 819-820, 1939.

SCHUMACHER, F. X.; COILE, T. S. **Growth and yield of natural stands of the southern pines**. Durham: T.S. Coile, 1960. 115 p.

SILVA, G. C. C. **Modelagem do crescimento e da produção florestal em povoamentos de eucalipto desbastado e não desbastado**. 2017. 129 p.
Dissertação (Mestrado em Engenharia Florestal) - Universidade Federal de Lavras, Lavras, 2017.

SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais: para engenharia e ciências aplicadas**. São Paulo: Arliber, 2010. 399 p.

SMOLA, A. J. et al. Introduction to large margin classifiers. In: _____. (Ed.). **Advances in large margin classifiers**. Cambridge: MIT Press, 1999. p. 1-28.

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