



Modelagem GAMLSS em dois estágios para seleção de clones de *Eucalyptus Spp.*

GAMLSS modeling in two stages for selection of *Eucalyptus Spp.* clones

Rayane Santos Leite ¹, José Antônio Aleixo da Silva ², Tiago Almeida de Oliveira ¹

RESUMO

A modelagem da distribuição do volume é essencial ao manejo florestal. Este estudo avaliou uma estratégia em dois estágios para modelar a produtividade volumétrica de três clones de *Eucalyptus spp.* em cinco espaçamentos, no Polo Gesseiro do Araripe-PE, utilizando Modelos Aditivos Generalizados para Localização, Escala e Forma (GAMLSS). Foram analisadas 2.896 árvores em 60 parcelas, em delineamento inteiramente casualizado, fatorial 3×5 , com quatro repetições. No primeiro estágio, ajustou-se um GAMLSS com distribuição SHASH ao nível das plantas, mantendo o parâmetro de localização como intercepto e associando escala e forma aos tratamentos; o ajuste apresentou desvio global de -12.025,74 e AIC/GAIC ($k = 2$) de -11.973,74. No segundo, os resíduos quantílicos foram agregados por parcela e submetidos à ANOVA fatorial, atendendo à normalidade (Shapiro-Wilk, $p = 0,433$) e homocedasticidade (Bartlett, $p = 0,790$). Clone ($p = 0,002$) e espaçamento ($p < 0,001$) foram significativos, sem interação ($p = 0,406$). O clone C3 foi mais produtivo ($0,0494 \text{ m}^3$), e o espaçamento $3 \text{ m} \times 2 \text{ m}$ apresentou maior rendimento ($0,0549 \text{ m}^3$). A abordagem acomodou a distribuição do volume e preservou a inferência experimental por parcela, indicando utilidade para comparações de clones e espaçamentos em condições edafoclimáticas do semiárido pernambucano florestal regional.

Palavras-chave: Assimetria; Curtose; Modelos flexíveis; SHASH; Volumétrico.

ABSTRACT

Volume distribution modeling is essential for forest management. This study evaluated a two-stage strategy to model the volumetric productivity of three *Eucalyptus spp.* clones under five spacings in the Araripe Gypsum Hub, Pernambuco, Brazil, using Generalized Additive Models for Location, Scale and Shape (GAMLSS). Overall, 2,896 trees in 60 plots were analyzed, in a completely randomized 3×5 factorial design with four replications. In the first stage, a GAMLSS with the SHASH distribution was fitted at tree level, keeping the location parameter as an intercept and associating scale and shape with treatments; the fit showed a global deviance of -12,025.74 and an AIC/GAIC($k = 2$) of -11,973.74. In the second stage, quantile residuals were aggregated by plot and submitted to factorial ANOVA, meeting normality (Shapiro-Wilk, $p = 0.433$) and homoscedasticity (Bartlett, $p = 0.790$) assumptions. Clone ($p = 0.002$) and spacing ($p < 0.001$) were significant, without interaction ($p = 0.406$). Clone C3 was the most productive (0.0494 m^3), and the $3 \text{ m} \times 2 \text{ m}$ spacing showed highest yield (0.0549 m^3). The approach accommodated volume distribution and preserved plot-level experimental inference, indicating usefulness for comparing clones and spacings under edaphoclimatic conditions of the Pernambuco semiarid regional forestry context.

Keywords: Asymmetry; Kurtosis; Flexible models; SHASH; Volumetric.

¹ Universidade Federal Rural de Pernambuco;

² Universidade Estadual da Paraíba.

Autor correspondente Rayane Santos Leite: rayferreiraleite@gmail.com

Recebido em: 31/03/2025. Aceito em: 06/07/2026.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/>), que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

1 INTRODUÇÃO

O Polo Gesseiro do Araripe se destaca por sua significativa concentração de reservas do minério gipsita, matéria prima para a produção do gesso, sendo reconhecida por sua alta qualidade, contribuindo com a produção de gesso no País, que se destaca como um grande produtor mundial desse material. No entanto, têm-se atrelado a este cenário, os avanços do desmatamento na mata nativa, causando grandes impactos ao meio ambiente, em virtude da biomassa florestal ser a fonte de energia mais utilizada pelas indústrias do Polo Gesseiro do Araripe, Pernambuco.

Diante do crescente consumo de madeira e derivados, a implantação de florestas energéticas com elevada produtividade torna-se fundamental para mitigar os impactos ambientais e gerar ganhos econômicos para o estado de Pernambuco. Nesse contexto, florestas de rápido crescimento do gênero *Eucalyptus* spp. representam uma alternativa viável para reduzir a pressão sobre a vegetação nativa, proporcionando bom volume de madeira em curto período.

A investigação das características da espécie por meio de métodos estatísticos robustos é essencial, visto que o crescimento do eucalipto apresenta relação não linear entre volume (V), diâmetro a 1,3 m do solo (D) e altura total (Ht). Métodos adequados possibilitam modelar essas relações, garantindo estimativas mais precisas do volume de madeira e subsidiando decisões de manejo florestal (Dantas *et al.*, 2020; Azevedo *et al.*, 2020).

Algumas técnicas estatísticas clássicas empregadas na análise de experimentos, como a análise de variância, dependem de pressupostos como normalidade dos resíduos, independência e homogeneidade de variâncias. Quando esses pressupostos são violados, a validade dos testes de significância e das comparações entre tratamentos pode ser comprometida (Glass, Peckham e Sanders, 1972; Montgomery, 2019). Assim, metodologias mais flexíveis e menos restritivas têm sido desenvolvidas. Dentre as técnicas de modelagem, o uso de modelos da classe de regressão tem sido amplamente utilizado para descrever fenômenos biológicos com acurácia (Santos *et al.*, 2025), assim como abordagens preditivas, que têm sido propostas como alternativas.

Rigby e Stasinopoulos (2005) introduziram os Modelos Aditivos Generalizados para Localização, Escala e Forma (GAMLSS), uma classe que reúne modelos lineares e não lineares, Modelos Lineares Generalizados (GLM) (Nelder; Wedderburn, 1972), Modelos Aditivos Generalizados (GAM) (Hastie; Tibshirani, 1990) e modelos mistos, proporcionando maior flexibilidade na especificação da distribuição da variável resposta (Kneib; Silbersdorff; Säfken, 2023; Umlauf; Klein; Zeileis, 2021; Carlan; Kneib; Klein, 2024). Os GAMLSS constituem uma classe de modelos que admite diversas distribuições para a variável resposta, não restritas à família exponencial, oferecendo maior flexibilidade para dados com assimetria e caudas pesadas. Esses modelos têm sido aplicados em diversas áreas,

como atuária, genômica, finanças, medicina e meteorologia (Stasinopoulos *et al.*, 2017).

No agronegócio brasileiro, as florestas plantadas de eucalipto ocupam mais de 7,5 milhões de hectares (IBÁ, 2023), sendo de grande importância econômica na produção de celulose, papel, painéis e carvão vegetal (Eufrade-Junior *et al.*, 2022). A quantificação volumétrica é um parâmetro fundamental em florestas de rápido crescimento, sendo uma das informações mais relevantes para o manejo florestal (Mendonça *et al.*, 2017; Müller *et al.*, 2024). Estimar com precisão a produção de povoamentos florestais é essencial para o planejamento eficiente.

Embora a análise de variância clássica seja amplamente empregada em experimentos florestais, os dados volumétricos frequentemente violam seus pressupostos de normalidade e homogeneidade de variâncias, comprometendo a validade das comparações entre clones; os GAMLSS superam essa limitação ao modelar também a dispersão, a assimetria e a curtose. Hipotetiza-se que uma abordagem em dois estágios, modelagem distribucional via GAMLSS seguida de análise de variância ao nível da unidade experimental descreva a produtividade volumétrica com maior adequação do que a análise de variância clássica, preservando a validade inferencial. Assim, o objetivo deste trabalho foi modelar a produtividade volumétrica de clones de *Eucalyptus* spp. em diferentes espaçamentos, plantados no Polo Gesseiro do Araripe, Pernambuco, identificando o clone mais produtivo por meio de modelagem em dois estágios: o primeiro utilizando modelos GAMLSS e o segundo a análise de variância (ANOVA).

2 MATERIAL E MÉTODOS

2.1 DESCRIÇÃO DO EXPERIMENTO

Os dados utilizados são oriundos de um experimento com clones de *Eucalyptus* spp., instalado em março de 2010, na Estação Experimental do Instituto Agrônomo de Pernambuco (IPA), com coordenadas 07° 28' 04,2" S e 40° 24' 37,7" W, a 816 metros de altitude. O solo da área é classificado como latossolo amarelo + latossolo vermelho-amarelo (Gouveia *et al.*, 2015) no município de Araripina-PE. O clima é classificado como BSh (semiárido quente), caracterizado por longos períodos de seca e chuvas irregulares, de acordo com a classificação de Köppen. A precipitação média anual é de aproximadamente 740 mm, com elevados índices de evaporação, a temperatura média é de 25 °C (Silva, 2020).

O experimento se constitui de três clones de eucaliptos: (C11 - C1 - Híbrido de *Eucalyptus brassiana*, C39 - C2 - Híbrido de *Eucalyptus urophylla* e C41 - C3 - Híbrido de *Eucalyptus urophylla*), obtidos por cruzamento natural, plantados em cinco espaçamentos distintos (2m × 1m; 2m × 2m; 3m × 2m; 3m × 3m e 4m × 2m). O delineamento foi inteiramente casualizado, em um esquema fatorial (3 × 5), com quatro repetições. Os dados foram coletados em março de 2018, totalizando um povoamento de 96 meses, em uma área de 2,2 ha, composta por 60 parcelas de

dimensão (14m × 21m) e área de (294 m²), contendo 49 plantas por parcela e 25 plantas na área útil, eliminando o efeito de bordadura.

Ao todo, foram analisadas 2.896 árvores (C1 = 941; C2 = 994; C3 = 961). As variáveis coletadas semestralmente foram: diâmetro medido a 1,30 metros do solo (D), obtido pela fórmula ($D = C/\pi$), em que C é a circunferência medida a 1,3 metros do solo, altura total da árvore (Ht) e volume (m³) de todas as plantas contidas na área útil de cada parcela. A altura total (Ht) não foi incluída na modelagem, pois em povoamentos clonais equiâneos, é fortemente colinear com o diâmetro (D) e sua mensuração em árvores altas está sujeita a maior erro operacional. O diâmetro, por sua vez, também não foi empregado como covariável no parâmetro de locação: por ser fortemente determinado pelos tratamentos, sua inclusão absorveria o efeito de clone e de espaçamento sobre a produção, comprometendo a inferência. A variável volume foi, portanto, modelada diretamente. A possibilidade de estimar o volume sem perda relevante de acurácia em sistemas de eucalipto sem a inclusão da altura é corroborada por Müller *et al.* (2024).

Após o abate das 2896 árvores, realizou-se o procedimento de cubagem das árvores pelo método de Smalian (Silva; Neto, 1979). Para isto, foram medidos os diâmetros ao longo do fuste nos intervalos: 0,30 m; 0,50 m; 0,70 m; 0,90 m; 1,10 m; 1,30 m; 1,50 m; 1,70 m; 1,90 m e 2,30 m. A partir deste ponto as medições foram realizadas em intervalos de 1 metro até o topo das árvores. As alturas reais foram medidas com o uso de uma trena de 30 m. O volume foi determinado pelo método Smalian para cada seção, sendo obtido por:

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

em que, v_i , é o volume da seção i (m³); g_i , área basimétrica no início da seção i (m²); g_{i+1} , área basimétrica no fim da seção i (m²); l_i , comprimento da seção i ; g_i calculado por meio da seguinte expressão: $g_i = \frac{\pi \cdot d_i^2}{4}$, sendo d_i , o diâmetro da seção na posição que se deseja calcular.

2. 2 MODELAGEM EM DOIS ESTÁGIOS

A análise foi estruturada em dois estágios para conciliar dois níveis de informação presentes no experimento. No primeiro estágio, as 2.896 árvores foram utilizadas no ajuste do GAMLSS, pois as medições individuais contêm a informação necessária para caracterizar a distribuição do volume, incluindo heterogeneidade, assimetria e comportamento das caudas. Essa etapa não foi usada para inferência experimental sobre clones e espaçamentos; sua finalidade foi produzir resíduos quantílicos normalizados a partir de uma distribuição flexível. No segundo estágio, esses resíduos foram agregados por parcela, que é a unidade experimental do delineamento, resultando em 60 observações independentes para a ANOVA fatorial 3 × 5 e para o teste de Tukey.

Essa estratégia separa explicitamente dois objetivos analíticos e apoia-se em fundamentos consolidados na análise de experimentos. A análise em estágios (*stage-wise analysis*) é amplamente empregada em ensaios agrônômicos e de melhoramento, nos quais uma primeira etapa resume a informação em nível apropriado e uma segunda etapa conduz a inferência sobre os tratamentos (Piepho et al., 2012). No primeiro estágio, o GAMLSS atua como etapa de normalização distribucional no nível das plantas, permitindo que a dispersão, a assimetria e o peso das caudas do volume variem conforme os tratamentos. No segundo estágio, a inferência experimental é realizada somente após a agregação por parcela, respeitando a unidade experimental e evitando a pseudorrepetição que resultaria do uso das árvores como observações independentes nos testes de clone e espaçamento (Hurlbert, 1984). Por se tratar de delineamento balanceado, a análise sobre os resíduos agregados por parcela preserva os graus de liberdade do experimento e é coerente com a lógica das abordagens de estágio único com a parcela como unidade experimental, cuja equivalência às análises em dois estágios, em condições balanceadas, é bem documentada (Möhring e Piepho, 2009; Damesa et al., 2017).

2. 2. 1 Modelagem primeiro estágio

A formulação geral dos GAMLSS, com preditores aditivos para os parâmetros de locação (μ), escala (σ) e forma (ν , τ), é apresentada na Equação 2. No ajuste final, adotou-se a distribuição Sinh-Arcsinh (SHASH; Jones e Pewsey, 2009), adequada para dados contínuos assimétricos e com diferentes pesos de cauda. A especificação adaptada ao experimento é definida a seguir:

$$Y^{ind} \sim D(\mu, \sigma, \nu, \tau) \begin{cases} \eta_1 = g_1(\mu) = \mathbf{X}_1 \boldsymbol{\beta}_1 + s_{11}(x_{11}) + \dots + s_{1J_1}(x_{1J_1}) \\ \eta_2 = g_2(\sigma) = \mathbf{X}_2 \boldsymbol{\beta}_2 + s_{21}(x_{21}) + \dots + s_{2J_2}(x_{2J_2}) \\ \eta_3 = g_3(\nu) = \mathbf{X}_3 \boldsymbol{\beta}_3 + s_{31}(x_{31}) + \dots + s_{3J_3}(x_{3J_3}) \\ \eta_4 = g_4(\tau) = \mathbf{X}_4 \boldsymbol{\beta}_4 + s_{41}(x_{41}) + \dots + s_{4J_4}(x_{4J_4}) \end{cases} \quad (2)$$

em que, $D(\mu, \sigma, \nu, \tau)$ da equação (2) é uma distribuição de quatro parâmetros, μ é um parâmetro de locação, σ é um parâmetro de escala, ν e τ são os parâmetros de forma da distribuição, comumente associados à assimetria e curtose, respectivamente. As matrizes de incidência são modeladas a nível de cada parâmetro da distribuição utilizada com $\mathbf{X}_i$ $i = 1 \dots 4$ que podem ou não coincidir, ou seja, o preditor de cada parâmetro da distribuição recebe diferentes variáveis explicativas e S são funções de suavização (*smooth curves*) (Rigby; Stasinopoulos, 2005). Assim, a notação adaptada a realidade do experimento ficou:

$$Y_{ijk} \sim SHASH(\mu, \sigma, \nu, \tau) \quad (3)$$

A especificação utilizada foi com $\mu_i = \beta_0$, $\log(\sigma_i) = \gamma_0 + \text{Clone} + \text{Espaçamento} + \text{Clone} \times \text{Espaçamento}$, $\log(\nu_i) = \delta_0 + \text{Espaçamento}$ e $\log(\tau_i) = \kappa_0 + \text{Espaçamento}$. A decisão de manter o parâmetro de locação (μ) apenas como intercepto foi central para

a análise em dois estágios. Como μ representa o nível central da distribuição, incluir clone, espaçamento, interação ou diâmetro nesse parâmetro removeria, no primeiro estágio, parte do efeito médio que se desejava avaliar no segundo estágio. Assim, os efeitos de tratamento foram utilizados para modelar escala e forma, permitindo acomodar heterogeneidade, assimetria e curtose, mas preservando as diferenças de produtividade para a ANOVA ao nível de parcela.

Dentre as distribuições GAMLSS ajustadas ao volume (m^3) de clones de *Eucalyptus* spp. a *Sinh-Arcsinh* de quatro parâmetros é considerada como caso canônico, pois a locação $\mu \in \mathbb{R}$ e a escala $\sigma > 0$ são removidos, e podem ser reintegrados utilizando a densidade da nova família $\sigma^{-1}f_{\epsilon,\delta}(\sigma^{-1}(x - \mu))$ em que $f_{\epsilon,\delta}(x)$ é a densidade de um membro da nova família. Logo $\epsilon \in \mathbb{R}$ se tornará um parâmetro de assimetria e $\delta > 0$ controlará o peso das caudas. Associando as variáveis aleatórias Z e $X_{\epsilon,\delta}$ com a densidade da normal padrão $\emptyset$ e $f_{\epsilon,\delta}$, respectivamente. Obtém-se a definição *Sinh-Arcsinh*: $Z = S_{\epsilon,\delta}(X_{\epsilon,\delta}) \equiv \sinh\{\epsilon + \delta \sinh^{-1}(X_{\epsilon,\delta})\}$.

A densidade completa da distribuição *Sinh-Arcsinh* unimodal é apresentada na Equação (4). Os parâmetros ϵ e δ podem ser interpretados como assimetria e curtose, respectivamente. Segundo Jones e Pewsey (2009), ao ajustá-los separadamente, acomodam-se caudas mais pesadas ou mais leves do que as da distribuição normal.

$$f_{\epsilon,\delta}(x) = \frac{1}{\sqrt{2\pi}} \frac{\delta C_{\epsilon,\delta}(x)}{\sqrt{1+x^2}} \exp\left\{\frac{1}{2} S_{\epsilon,\delta}^2(x)\right\} \quad (4)$$

em que, $C_{\epsilon,\delta}(x) = \cosh(\epsilon + \delta \sinh^{-1}(x)) = \sqrt{1 + S_{\epsilon,\delta}^2(x)}$.

2. 2. 2 Modelagem segundo estágio

Os resíduos quantílicos normalizados do GAMLSS foram calculados para cada árvore e, em seguida, agregados por parcela por meio da média aritmética. A parcela, e não a árvore individual, foi considerada a unidade experimental do delineamento; portanto, o banco do segundo estágio continha 60 linhas, correspondentes às combinações de clone, espaçamento e repetição.

A análise de variância foi conduzida segundo o modelo fatorial $y_{ijk} = \mu + C_i + E_j + (CE)_{ij} + \epsilon_{ijk}$, em que y_{ijk} representa o resíduo quantílico médio da parcela k submetida ao clone i e ao espaçamento j ; μ é a média geral; C_i é o efeito de clone; E_j é o efeito de espaçamento; $(CE)_{ij}$ é a interação clone $\times$ espaçamento; e ϵ_{ijk} é o erro experimental. Como o delineamento é balanceado, com três clones, cinco espaçamentos e quatro repetições, o erro da ANOVA possui 45 graus de liberdade.

Essa estrutura evita pseudorrepetição, pois os testes F e o teste de Tukey são realizados com base nas parcelas, respeitando o nível em que os tratamentos foram casualizados. A independência é assumida entre parcelas, enquanto a variação entre árvores dentro de cada parcela é incorporada ao primeiro estágio como informação distribucional. Os pressupostos do segundo estágio foram avaliados pelos resíduos da ANOVA, por meio do teste de Shapiro-Wilk para normalidade, do teste de Bartlett para

homogeneidade de variâncias e de gráficos diagnósticos. Quando a interação não foi significativa, a interpretação baseou-se nos efeitos principais de clone e espaçamento.

A escolha dos resíduos quantílicos como resposta do segundo estágio possui fundamentação formal. Trata-se dos resíduos quantílicos aleatorizados de Dunn e Smyth (1996), definidos por $r_i = \Phi^{-1}\{F(y_i; \mu, \sigma, \nu, \tau)\}$, em que F é a função de distribuição acumulada ajustada e Φ^{-1} é a função quantílica da normal padrão. Sob especificação correta do modelo, esses resíduos têm distribuição normal padrão independentemente da distribuição original da resposta, propriedade que os consagra como o resíduo de referência dos GAMLSS (Stasinopoulos et al., 2017) e que justifica sua avaliação pelo worm plot (van Buuren e Fredriks, 2001). Como o parâmetro de locação foi mantido como intercepto, a distribuição ajustada no primeiro estágio não absorve as diferenças de nível entre os tratamentos: estas permanecem como deslocamentos sistemáticos dos resíduos quantílicos, que são preservados na agregação por parcela e testados no segundo estágio. A modelagem de escala e forma (σ, ν, τ) pelos tratamentos assegura que a transformação remova apenas os artefatos distribucionais, tais como: heterocedasticidade, assimetria e curtose; e não o sinal de produtividade. Conceitualmente, o procedimento equivale a uma normalização baseada em modelo, análoga às transformações estabilizadoras de variância clássicas (Box e Cox, 1964), porém mais flexível por acomodar simultaneamente dispersão, assimetria e peso de caudas; após essa normalização, a inferência fatorial padrão ao nível de parcela é válida. Para os modelos estatísticos os seguintes pressupostos foram avaliados: normalidade dos resíduos (teste de Shapiro-Wilk). Todas as análises estatísticas foram realizadas no software R, versão 4.2.1 (Team Posit, 2025). A modelagem distribucional do primeiro estágio foi conduzida com o pacote *gamlss* (Stasinopoulos e Rigby, 2007); a análise de variância fatorial e as comparações múltiplas pelo teste de Tukey, no segundo estágio, foram realizadas com os pacotes *ExpDes* (Ferreira et al., 2014), *multcomp* (Hothorn, Bretz e Westfall, 2008) e *BSagri* (Schaarschmidt, 2023).

3 RESULTADO E DISCUSSÃO

A Figura 1 apresenta os gráficos boxplot para os efeitos de clones (a), espaçamento (b) e interação (c). Na Figura 1(a), observa-se variabilidade nos clones C1 e C2, com leve assimetria no C3. O clone C3 apresentou maior média de volume, seguido de C1 e C2, com desempenho competitivo entre si. Na Figura 1(b), o espaçamento E3 (3m × 2m) apresentou a maior média volumétrica. Na interação [Figura 1(c)], os clones apresentam maior média dentro do espaçamento E3, independentemente do clone, sugerindo papel fundamental do espaçamento no desempenho volumétrico.

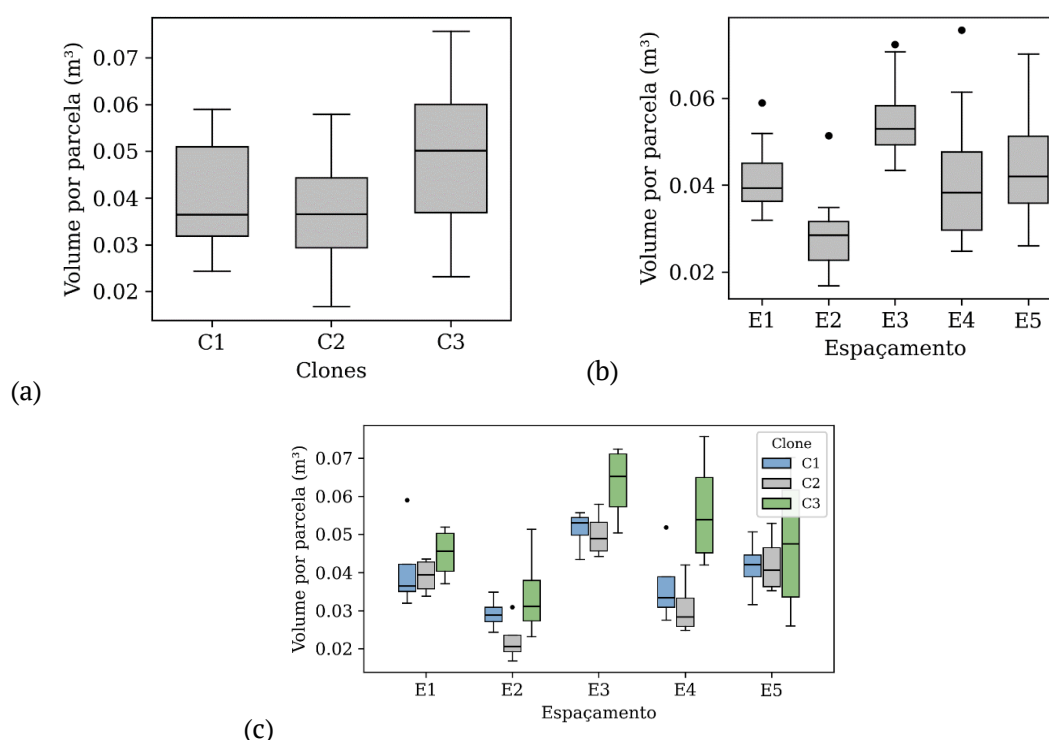


Figura 1. Boxplot dos efeitos de clone (a), espaçamento (b) e interação clone x espaçamento (c) do volume por parcela ($n = 60$), unidade experimental do delineamento.

Fonte: os autores (2025)

A competitividade entre os clones pode ser justificada pelas características das espécies. Silva, Silva e Tavares (2022), utilizando os mesmos clones e espaçamentos na mesma Estação Experimental sob condições de seca severa, verificaram que o clone de *Eucalyptus urophylla* apresentou os maiores valores de incremento médio anual. Como antecedentes históricos do potencial do gênero em ambientes secos, estudos no Níger demonstraram que o *Eucalyptus brassiana* dentre trinta e cinco espécies de *Eucalyptus* estudadas, foi uma das mais promissoras, para condições do semiárido sudano-saheliana citado por Silva *et al.* (1980), e em Moçambique o gênero foi apontado como de grande potencialidade para reflorestamentos em clima seco a semiárido (Bila; Issufo, 1993). Esses relatos pioneiros são aqui tomados apenas como contexto histórico. O *Eucalyptus urophylla* espécie de rápido crescimento e grande porte, conforme caracterização clássica de (Lorenzi *et al.*, 2003), foi a espécie com maior volume no presente experimento, em concordância com os resultados recentes de Silva, Silva e Tavares (2022) na mesma estação experimental. O *Eucalyptus urophylla*, de rápido crescimento (IBÁ, 2023), é uma das espécies mais cultivadas no Brasil, por seu grande potencial de crescimento em termos de área plantada, trazendo uma boa produtividade, utilizada em diversos fins: carvão, celulose e papel, chapas duras, serraria, dentre outros (Junior; Garcia, 2004).

A maior produtividade volumétrica do clone C3 (Figura 1a), híbrido de *Eucalyptus urophylla*, é consistente com o desempenho superior dessa espécie em ambientes com restrição hídrica, associado ao rápido crescimento, à eficiência no uso da água e à plasticidade adaptativa (Silva, Silva e Tavares, 2022; Aguiar Júnior *et al.*,

2022). André *et al.* (2021) e Fernandes *et al.* (2023) também reportaram que diferentes espaçamentos afetaram significativamente a produção volumétrica de clones de eucalipto, corroborando os resultados encontrados neste estudo.

Silva, Silva e Tavares (2022), em estudo na mesma área experimental com os mesmos clones sob condições de seca severa, concluíram que o arranjo espacial influencia fortemente a produtividade, com valores de IMA de até $15,92 \text{ m}^3 \text{ ha}^{-1} \text{ ano}^{-1}$ para o clone C39. Stape *et al.* (2022) também demonstraram que o espaçamento e o arranjo geométrico são determinantes na produtividade de plantios clonais de eucalipto.

Gouveia *et al.* (2015), na mesma Estação Experimental e com o mesmo clone C3, compararam modelos clássicos com modelos mistos não lineares, obtendo melhores resultados com estes últimos para a modelagem do volume. A assimetria encontrada nos dados era esperada para esta cultura (Fumes *et al.*, 2017). Diversas abordagens têm sido propostas para tratar assimetrias em dados florestais, como a modelagem distribucional via GAMLSS (Kneib; Silbersdorff; Säfken, 2023; De Bastiani *et al.*, 2023) e a modelagem de curvas de crescimento com efeitos aleatórios (Fumes *et al.*, 2017).

Os resultados da modelagem linear com arranjo fatorial não atenderam ao pressuposto de normalidade dos resíduos (teste de Shapiro-Wilk): modelo linear fatorial com a variável original, $W = 0,9198$ ($p < 0,001$); após transformação logarítmica, $W = 0,9500$ ($p < 0,001$); após transformação inversa, $W = 0,6236$ ($p < 0,001$); e após transformação arco-seno, $W = 0,9751$ ($p < 0,001$). Nenhuma das transformações aproximou os resíduos da normalidade — a logarítmica e a arco-seno reduziram o afastamento sem reverter a rejeição, ao passo que a inversa o agravou. Assim, adotou-se a modelagem GAMLSS como alternativa, selecionando a distribuição com base no GAIC e na interpretabilidade dos parâmetros.

Entre as distribuições comparadas por meio do GAIC (Tabela 1), a SHASH foi selecionada por combinar ajuste adequado, estabilidade computacional e interpretação direta dos parâmetros de escala, assimetria e curtose. (funções *fitDist* e *chooseDist* do *gamlss*), a SHASH figurou entre as de menor valor (GAIC ≈ -11.467 ; o menor foi a SEP3, ≈ -11.472) e foi selecionada por aliar bom ajuste à interpretabilidade direta dos parâmetros de assimetria e curtose. O ajuste final (Tabela 2) apresentou desvio global de $-12.025,74$, AIC/GAIC($k = 2$) de $-11.973,74$ e SBC de $-11.818,49$, com 26 graus de liberdade do modelo e 2.870 graus de liberdade residuais. A análise residual indicou que a modelagem dos efeitos apenas no parâmetro de locação não era adequada para estes dados; por isso, no primeiro estágio, os tratamentos foram utilizados para modelar escala e forma, enquanto o parâmetro de locação foi mantido como intercepto para preservar as diferenças médias de produtividade a serem avaliadas no segundo estágio.

Tabela 1. Critério de Informação Generalizado (GAIC) das distribuições candidatas ajustadas ao volume (m^3) ao nível de árvore (dez menores valores de AIC), sob três penalizações: $k = 2$ (AIC), $k = 3,84$ (equivalente a um teste qui-quadrado) e $k = \log(n)$ (BIC).

Distribuição	AIC (k = 2)	Qui-quadrado (k = 3,84)	BIC (k = log n)
SEP1	-11.460,55	-11.453,19	-11.436,66
SHASHo2	-11.457,21	-11.449,85	-11.433,33
ST2	-11.456,70	-11.449,34	-11.432,82
SST	-11.455,75	-11.448,39	-11.431,86
ST3	-11.453,69	-11.446,33	-11.429,81
ST1	-11.453,57	-11.446,21	-11.429,68
SEP2	-11.450,28	-11.442,92	-11.426,40
SEP3	-11.440,00	-11.432,64	-11.416,11
SHASHo	-11.435,57	-11.428,21	-11.411,69
SN2	-11.430,87	-11.425,35	-11.412,95

Nota: SHASHo e SHASHo2 são parametrizações da distribuição Sinh-Arcsinh (SHASH); a linha em negrito indica a família adotada. Embora a SEP1 tenha apresentado o menor GAIC, a diferença para a família SHASH foi pequena ($\Delta AIC \approx 3$ em relação à SHASHo2, a segunda colocada). A SHASH foi adotada por aliar ajuste competitivo à interpretação direta dos parâmetros de assimetria (ϵ) e de curtose (δ) — aspecto central para a caracterização do comportamento distribucional do volume no primeiro estágio — e por sua estabilidade computacional no ajuste com os efeitos de tratamento sobre escala e forma.

Fonte: os autores (2025)

Essa estratégia separa explicitamente dois objetivos analíticos. No primeiro estágio, o GAMLSS atua como etapa de normalização distribucional no nível das plantas, permitindo que a variabilidade, a assimetria e o peso das caudas do volume variem conforme os tratamentos. Rigby e Stasinopoulos (2005) enfatizam a importância de uma análise detalhada dos resíduos nos modelos GAMLSS, e ajustes subsequentes podem ser feitos quando resíduos indicarem inadequações. No segundo estágio, a inferência experimental é realizada somente após a agregação por parcela, respeitando a unidade experimental e evitando a pseudorrepetição que resultaria do uso das árvores como observações independentes nos testes de clone e espaçamento (Hurlbert, 1984). Por se tratar de delineamento balanceado, essa análise sobre os resíduos agregados por parcela preserva os graus de liberdade do experimento e é coerente com a lógica de modelos mistos de estágio único com parcela como unidade experimental (Piepho et al., 2012; Möhring; Piepho, 2009; Damesa et al., 2017). Aguiar Júnior *et al.* (2022) aplicaram abordagem semelhante na modelagem de clones de eucalipto em sistemas agrossilvipastoris

3. 1 RESULTADO DO MODELO GAMLSS (PRIMEIRO ESTÁGIO)

As estimativas do modelo SHASH ajustado ao nível das plantas são apresentadas na Tabela 2. O parâmetro de locação (μ) foi mantido como intercepto, enquanto o parâmetro de escala (σ) foi modelado por clone, espaçamento e interação clone $\times$ espaçamento, e os parâmetros de forma (ν e τ) por espaçamento. Essa especificação permitiu acomodar diferenças de dispersão e formato da distribuição sem retirar do segundo estágio as diferenças médias entre tratamentos.

Tabela 2. Estimativas dos parâmetros do modelo GAMLSS-SHASH ajustado ao volume das plantas no primeiro estágio

Parâmetro	Efeito	Estimativa	Erro padrão	t	Valor p
μ (locação) – ligação identidade					
	Intercepto	0,02209	0,00078	28,500	<0,001
σ (escala) – ligação log					
	Intercepto	-2,91666	0,13932	-20,935	<0,001
	Clone C2	-0,04050	0,06402	-0,633	0,527
	Clone C3	-0,04895	0,06624	-0,739	0,460
	Espaçamento E2	0,25324	0,26470	0,957	0,339
	Espaçamento E3	0,79882	0,25829	3,093	0,002
	Espaçamento E4	0,60727	0,25309	2,399	0,016
	Espaçamento E5	0,63173	0,25741	2,454	0,014
	C2 × E2	0,06466	0,08422	0,768	0,443
	C3 × E2	0,12243	0,08942	1,369	0,171
	C2 × E3	0,07109	0,08340	0,852	0,394
	C3 × E3	0,22904	0,08834	2,593	0,010
	C2 × E4	0,01613	0,08060	0,200	0,841
	C3 × E4	0,15750	0,08795	1,791	0,073
	C2 × E5	-0,04919	0,08600	-0,572	0,567
	C3 × E5	0,08694	0,08866	0,981	0,327
v (assimetria) – ligação log					
	Intercepto	1,1971	0,1354	8,840	<0,001
	Espaçamento E2	0,3549	0,2588	1,371	0,170
	Espaçamento E3	0,9783	0,2530	3,867	<0,001
	Espaçamento E4	0,7525	0,2486	3,027	0,002
	Espaçamento E5	0,6708	0,2549	2,632	0,009
τ (curtose) – ligação log					
	Intercepto	-0,06727	0,07962	-0,845	0,398
	Espaçamento E2	0,28491	0,17981	1,585	0,113
	Espaçamento E3	0,57481	0,19219	2,991	0,003
	Espaçamento E4	0,36308	0,18097	2,006	0,045
	Espaçamento E5	0,45147	0,18821	2,399	0,017

Nota: μ = locação; σ = escala; v = assimetria; τ = curtose. O modelo foi ajustado com 2.896 observações; graus de liberdade do ajuste = 26; graus de liberdade residuais = 2.870; desvio global = -12.025,74; AIC/GAIC(k = 2) = -11.973,74; SBC = -11.818,49.

Fonte: os autores (2025)

Os efeitos dos tratamentos sobre os parâmetros distribucionais concentraram-se nos espaçamentos. No parâmetro de escala (σ), foram significativos os espaçamentos E3, E4 e E5 ($p < 0,05$) e a interação C3 × E3 ($p = 0,010$), indicando que a

heterogeneidade de tamanho das árvores variou entre os arranjos espaciais e, em particular, na combinação do clone C3 com o espaçamento 3 m × 2 m. Nos parâmetros de forma, os espaçamentos E3, E4 e E5 foram significativos tanto para a assimetria (v) quanto para a curtose (τ). Os efeitos de clone sobre a escala e os coeficientes associados ao espaçamento E2 não foram significativos; por isso, a interpretação privilegia os efeitos estatisticamente relevantes, mantendo-se os demais na Tabela 2 apenas por transparência.

A modelagem dos parâmetros de forma agrega interpretação silvicultural à análise. O parâmetro de escala (σ) reflete a heterogeneidade de tamanho das árvores dentro de cada combinação clone × espaçamento, associada à competição e à uniformidade do povoamento; a assimetria (v) indica o predomínio relativo de árvores dominadas ou dominantes; e a curtose (τ) reflete a frequência de indivíduos atípicos, ou seja, de árvores muito menores ou muito maiores que a média. Do ponto de vista prático, clones e espaçamentos que reduzem σ e aproximam v de zero originam povoamentos mais homogêneos — desejáveis para a colheita mecanizada e para a regularidade do suprimento de biomassa energética. Essa leitura distribucional, inacessível à análise de variância clássica centrada apenas na média, constitui uma das contribuições da abordagem GAMLSS no primeiro estágio.

A adequação global do modelo SHASH no primeiro estágio é indicada pelas estatísticas dos resíduos quantílicos: média de 0,1460, variância de 1,0667, assimetria de 0,0127, curtose de 2,4081 e coeficiente de Filliben de 0,9862. Esses valores indicam centralidade e variabilidade próximas às esperadas para resíduos normalizados, com assimetria praticamente nula. A curtose inferior a 3 e o worm plot indicam desvios nas caudas, aspecto esperado em dados volumétricos com valores extremos; ainda assim, o ajuste foi considerado aceitável como etapa de normalização distribucional anterior à ANOVA por parcela.

Os diagnósticos gráficos do primeiro estágio (gráficos de resíduos quantílicos e worm plot) são apresentados na Figura 2(a-b), em conjunto com os resíduos da ANOVA fatorial do segundo estágio [Figura 2(c)]. O worm plot mostra boa acomodação na região central da distribuição, com pequenas discrepâncias nas caudas; os gráficos de resíduos quantílicos indicam ausência de padrão sistemático relevante ao longo do índice das observações. Esse comportamento reforça a necessidade de utilizar uma distribuição flexível, como a SHASH, antes da inferência experimental ao nível da parcela.

Após a modelagem via GAMLSS com distribuição SHASH, procedeu-se à análise de variância fatorial dos resíduos quantílicos agregados por parcela, em delineamento inteiramente casualizado (segundo estágio).

3. 2 RESULTADO DO MODELO DE ANOVA (SEGUNDO ESTÁGIO)

A análise de variância ao nível de parcela, conduzida sobre os resíduos quantílicos do primeiro estágio, indicou efeitos significativos de clone ($F_c = 7,063$; $p = 0,002$) e de espaçamento ($F_c = 13,098$; $p < 0,001$), enquanto a interação clone × espaçamento não foi significativa ($F_c = 1,062$; $p = 0,406$). A ausência de interação

indica que o desempenho relativo dos clones foi consistente entre os espaçamentos no segundo estágio.

Os pressupostos da análise de variância foram atendidos ao nível de parcela: normalidade dos resíduos (Shapiro-Wilk, $W = 0,980$; $p = 0,433$) e homogeneidade de variâncias (Bartlett, $K^2 = 9,608$; $gl = 14$; $p = 0,790$). Como a interação não foi significativa, a interpretação baseou-se nos efeitos principais de clone e de espaçamento (Tabela 3), comparados pelo teste de Tukey (Quadros 1 e 2).

Tabela 3. Análise de variância fatorial dos resíduos quantílicos do GAMLSS agregados ao nível de parcela

Fonte de Variação	GL	Soma de Quadrados	Quadrado Médio	Fc	Valor P
Clone	2	0,6895	0,3447	7,063	0,002
Espaçamento	4	2,5573	0,6393	13,098	<0,001
Clone × Espaçamento	8	0,4149	0,0519	1,062	0,406
Resíduo	45	2,1964	0,0488		
Total	59	5,8581			

GL: Graus de liberdade; Fc: Teste F.

Fonte: os autores (2025)

Para o efeito de clone (Quadro 1), o teste de Tukey aplicado aos resíduos quantílicos agregados por parcela classificou o C3 como grupo superior (a), enquanto C1 e C2 compuseram o mesmo grupo estatístico (b). Para manter a interpretação silvicultural, o quadro apresenta também as médias observadas de volume por parcela, nas quais o C3 apresentou $0,0494 \text{ m}^3$, seguido por C1 ($0,0399 \text{ m}^3$) e C2 ($0,0369 \text{ m}^3$).

Quadro 1. Médias dos resíduos quantílicos por clone, agrupamento pelo teste de Tukey e médias observadas de volume

Clone	Média dos resíduos quantílicos	Grupo Tukey	Volume médio observado (m^3)
C3	0,2781	a	0,0494
C1	0,0974	b	0,0399
C2	0,0227	b	0,0369

Fonte: os autores (2025)

Para o efeito de espaçamento (Quadro 2), o teste de Tukey aplicado aos resíduos quantílicos indicou o espaçamento E3 como grupo superior (a). Os espaçamentos E1 e E5 foram classificados como grupo intermediário (ab), E4 como grupo bc e E2 como grupo inferior (c). Em termos de médias observadas de volume por parcela, E3 também apresentou o maior valor ($0,0549 \text{ m}^3$), seguido de E5 ($0,0439 \text{ m}^3$), E1 ($0,0417 \text{ m}^3$), E4 ($0,0413 \text{ m}^3$) e E2 ($0,0286 \text{ m}^3$).

O melhor desempenho do espaçamento $3 \text{ m} \times 2 \text{ m}$ reflete o compromisso entre a área útil disponível por árvore e a ocupação eficiente do sítio. Em densidades muito elevadas, a competição por água, luz e nutrientes restringe o crescimento individual;

em densidades muito baixas, os recursos do sítio são subutilizados. Arranjos intermediários tendem, assim, a maximizar a produção volumétrica por unidade de área, padrão amplamente reportado em plantios clonais de eucalipto (Stape et al., 2022; Fernandes et al., 2023; André et al., 2021). Esse efeito é particularmente relevante no semiárido, em que a limitação hídrica amplifica os efeitos da competição intraespecífica e torna o arranjo espacial determinante da produtividade (Silva, Silva e Tavares, 2022). A ausência de interação clone $\times$ espaçamento, por sua vez, indica que o desempenho relativo dos clones se mantém estável entre os espaçamentos, de modo que a seleção do clone e a definição do espaçamento podem ser otimizadas de forma independente — o que simplifica a recomendação de manejo.

Esse padrão é consistente com as maiores médias volumétricas observadas para o clone C3 e para o espaçamento E3. Entretanto, como a interação clone $\times$ espaçamento não foi significativa, a recomendação principal deve ser apresentada com base nos efeitos principais, e não no desdobramento da interação.

Quadro 2. Médias dos resíduos quantílicos por espaçamento, agrupamento pelo teste de Tukey e médias observadas de volume

Espaçamento	Média dos resíduos quantílicos	Grupo Tukey	Volume médio observado (m ³)
E3	0,4147	a	0,0549
E1	0,2156	ab	0,0417
E5	0,2041	ab	0,0439
E4	0,0297	bc	0,0413
E2	-0,2005	c	0,0286

As médias seguidas pela mesma letra nas colunas não diferem estatisticamente entre si, ao nível de 5% de significância pelo teste de Tukey.

Fonte: os autores (2025)

A análise de resíduos do modelo fatorial do segundo estágio é apresentada na Figura 2. No gráfico Residuals vs Fitted, os resíduos distribuem-se sem tendência sistemática evidente em torno de zero; no gráfico Q-Q, os pontos mantêm-se próximos à diagonal, com pequenas discrepâncias nas extremidades. Esses resultados são coerentes com os testes de Shapiro-Wilk e Bartlett, que indicaram atendimento aos pressupostos da ANOVA ao nível de parcela.

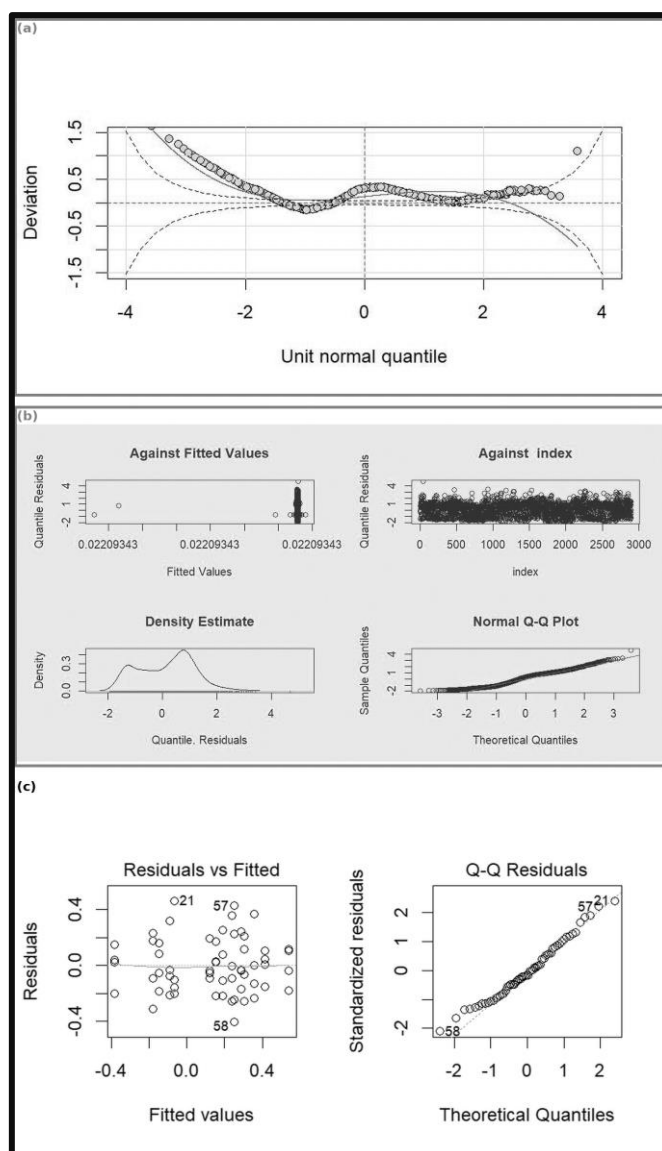


Figura 2. Diagnóstico dos resíduos quantílicos do primeiro estágio (GAMLSS-SHASH) — worm plot (a) e resíduos versus valores ajustados, versus índice, densidade e Q-Q plot (b) — e diagnóstico dos resíduos da ANOVA fatorial do segundo estágio aplicada aos resíduos quantílicos agregados por parcela (c).

Fonte: os autores (2025)

A análise gráfica confirma que o modelo satisfaz os pressupostos de homocedasticidade e normalidade dos resíduos, validando o ajuste para a análise proposta. A interpretação conjunta dos dois estágios evidencia a contribuição da abordagem distribucional. No primeiro estágio, o fato de os tratamentos atuarem significativamente sobre a escala e a forma da distribuição — e não apenas sobre a média — mostra que clone e espaçamento afetam também a heterogeneidade e a assimetria dos tamanhos das árvores dentro das parcelas, informação que a análise de variância clássica desconsidera. Essa riqueza distribucional tem motivado a adoção dos GAMLSS em estudos florestais, tanto na descrição da distribuição de variáveis dendrométricas (Oliveira et al., 2020; De Bastiani et al., 2023) quanto na modelagem do crescimento (Riebl, Klein e Kneib, 2023; Kneib, Silbersdorff e Säfken, 2023). No segundo estágio, ao transferir a inferência para o nível de parcela, a estratégia concilia

essa riqueza descritiva com a validade experimental: os efeitos de tratamento são testados na unidade de casualização, com pressupostos atendidos, evitando as conclusões espúrias que adviriam tanto da violação dos pressupostos quanto da pseudorrepetição. Tal compatibilidade entre modelagem distribucional e inferência experimental é especialmente útil em ensaios florestais, em que a subamostragem de árvores dentro de parcelas é a regra.

Do ponto de vista prático, a estratégia de modelagem proposta, GAMLSS em dois estágios associado à análise ao nível de parcela, constitui ferramenta inferencial robusta para descrever a produtividade volumétrica de povoamentos clonais quando os dados exibem forte dispersão, assimetria e curtose. Ao caracterizar não apenas a média, mas também o comportamento da dispersão e da forma da distribuição do volume, a abordagem auxilia a seleção de clones e o manejo de espaçamentos para fins energéticos no semiárido, orientando a escolha de arranjos como o 3 m × 2 m, que maximizou o rendimento volumétrico. Por respeitar a unidade experimental, a análise ao nível de parcela fornece inferências confiáveis para subsidiar o planejamento de plantios clonais de *Eucalyptus* na região do Polo Gesseiro do Araripe.

4 CONCLUSÃO

Os GAMLSS mostraram-se adequados para modelar a produtividade volumétrica de clones de *Eucalyptus* spp., por acomodarem características dos dados que violam os pressupostos da análise de variância clássica. A distribuição SHASH foi utilizada no primeiro estágio e produziu resíduos quantílicos com média 0,1460, variância 1,0667, assimetria 0,0127 e coeficiente de Filliben 0,9862, indicando ajuste global aceitável, embora com pequenos desvios nas caudas.

No primeiro estágio, o modelo GAMLSS com distribuição SHASH apresentou convergência e estatísticas dos resíduos adequadas, conforme visualizado no *worm plot*. A técnica foi eficiente para acomodar a dispersão, a assimetria e a curtose dos dados, e os resíduos quantílicos normalizados — preservando o efeito dos tratamentos — permitiram quantificar as diferenças entre clones e espaçamentos no segundo estágio.

No segundo estágio, a ANOVA fatorial ao nível de parcela apresentou pressupostos satisfeitos, com normalidade dos resíduos (Shapiro-Wilk, $p = 0,433$) e homogeneidade de variâncias (Bartlett, $p = 0,790$). O teste F indicou efeitos significativos de clone e de espaçamento, sem interação significativa entre eles; o teste de Tukey identificou o clone C3 como superior e o espaçamento E3 (3 m × 2 m) como o de maior rendimento volumétrico observado.

A modelagem em dois estágios mostrou-se alternativa adequada para dados com elevada dispersão, assimetria e curtose. O clone C3 (híbrido de *Eucalyptus urophylla*) apresentou o melhor rendimento volumétrico no espaçamento 3m × 2m na região do Polo Gesseiro do Araripe - PE, com média de efeito principal de 0,0494 m³, superando o clone C2 em 33,7% e o C1 em 23,6%.

5 AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Código de Financiamento 001. Universidade Federal Rural de Pernambuco – UFRPE. Instituto Agrônomo de Pernambuco – IPA.

6 REFERÊNCIAS

AGUIAR JÚNIOR, A. L. et al. *Eucalypt clone modelling in agrosilvopastoral systems*. **Pesquisa Agropecuária Brasileira**, v. 57, e02547, 2022. DOI: 10.1590/S1678-3921.pab2022.v57.02547.

ANDRÉ, D. L. et al. Wood volume of Eucalyptus clones established under different spacings in the Brazilian Cerrado. **Forest Science**, v. 67, n. 4, p. 478-487, 2021. DOI: 10.1093/forsci/fxab018.

AZEVEDO, G. B. de et al. **Multi-volume modeling of *Eucalyptus* trees using regression and artificial neural networks**. 2020.

BILA A, ISSUFO, A. K. Comportamento de eucaliptos na região norte e sul de Moçambique. In: Congresso i Florestal Brasileiro, 7, Curitiba. **Anais**. São Paulo, SBS/SBEF. 1993.

BOX, G. E. P.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 26, n. 2, p. 211-243, 1964.

CARLAN, M.; KNEIB, T.; KLEIN, N. Distributional Regression for Data Analysis. **Annual Review of Statistics and Its Application**, v. 11, p. 175-196, 2024. DOI: 10.1146/annurev-statistics-040722-053607.

DAMESA, T. M.; MÖHRING, J.; WORKU, M.; PIEPHO, H.-P. One step at a time: stage-wise analysis of a series of experiments. **Agronomy Journal**, v. 109, n. 3, p. 845-857, 2017. DOI: 10.2134/agronj2016.07.0395.

DANTAS, D. et al. **Multilevel nonlinear mixed-effects model and machine learning for volumetric estimation of *Eucalyptus* spp. clones**. *Cerne*, 2020.

DE BASTIANI, F. et al. Distributional regression modelling via generalized additive models for location, scale, and shape: An overview through a data set from learning analytics. **WIREs Computational Statistics**, e1617, 2023. DOI: 10.1002/wics.1617.

EUFRADE-JUNIOR, H. J. et al. Eucalyptus clone selection for charcoal production: survival, growth, and wood quality. **Biomass and Bioenergy**, v. 157, 106345, 2022. DOI: 10.1016/j.biombioe.2022.106345.

DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**, v. 5, n. 3, p. 236-244, 1996.

FERNANDES, G. L. et al. Effects of Spacing on Early Growth Rate and Yield of Hybrid Eucalyptus Stands. **Pertanika Journal of Tropical Agricultural Science**, v. 46, n. 2, p. 523-540, 2023.

FERRÃO, L. F. V. et al. Accurate genomic prediction of *Coffea canephora* in multiple environments using whole-genome statistical models. **Heredity**, v. 122, n. 3, p. 261-275, 2019.

FERREIRA, E. B. et al. ExpDes: an R package for ANOVA and experimental designs. **Applied Mathematics**, v. 5, n. 19, p. 2952, 2014. DOI: 10.4236/am.2014.519280.

FUMES, G. et al. Growth curves for diameter and height using mixed models: An application in eucalyptus seedling. **Open Journal of Forestry**, v. 7, n. 4, p. 403-415, 2017. DOI: 10.4236/ojf.2017.74024.

GOUVEIA, J. F. et al. Modelos volumétricos mistos em clones de Eucalyptus no polo gesso do Araripe, Pernambuco. **FLORESTA**, v. 45, n. 3, p. 587-598, 2015. DOI: 10.5380/rev.v45i3.36844.

GLASS, G. V.; PECKHAM, P. D.; SANDERS, J. R. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, v. 42, n. 3, p. 237-288, 1972.

HASTIE, T.; TIBSHIRANI, R. **Generalized Additive Models**. London: Chapman and Hall/CRC, 1990.

HOTHORN, T.; BRETZ, F.; WESTFALL, P. Simultaneous Inference in General Parametric Models. **Biometrical Journal**, v. 50, n. 3, p. 346-363, 2008. DOI: 10.1002/bimj.200810425.

HURLBERT, S. H. Pseudoreplication and the design of ecological field experiments. **Ecological Monographs**, v. 54, n. 2, p. 187-211, 1984. DOI: 10.2307/1942661.

IBÁ. Indústria Brasileira de Árvores. **Relatório Anual 2023**. São Paulo: IBÁ, 2023. Disponível em: <https://iba.org/publicacoes>. Acesso em: 20/05/2026.

JONES, M. C.; PEWSEY, A. Sinh-arcsinh distributions. **Biometrika**, v. 96, n. 4, p. 761-780, 2009. DOI: 10.1093/biomet/asp054.

JUNIOR, L. S; GARCIA, J. N. Determination of the physical and mechanical properties of the wood of *Eucalyptus urophylla*. **SCIENTIA FORESTALIS**. v. 65, p. 120-129. 2004.

KNEIB, T.; SILBERSDORFF, A.; SÄFKEN, B. Rage Against the Mean – A Review of Distributional Regression Approaches. **Econometrics and Statistics**, v. 26, p. 99-123, 2023. DOI: 10.1016/j.ecosta.2021.07.006.

LORENZI H, SOUZA HM, TORRES AV, BACHER LB. Árvores exóticas no Brasil: madeiras ornamentais e aromáticas. **Nova Odessa, Instituto Plantarum**. 2003.

MENDONÇA, A. R. de et al. Growth and yield models for eucalyptus stands obtained by differential equations. **Scientia Agricola**, v. 74, n. 5, p. 364–370, 2017. DOI: 10.1590/1678-992X-2016-0035.

MÖHRING, J.; PIEPHO, H.-P. Comparison of weighting in two-stage analysis of plant breeding trials. **Crop Science**, v. 49, n. 6, p. 1977-1988, 2009. DOI: 10.2135/cropsci2009.02.0083.

MONTGOMERY, D. C. **Design and Analysis of Experiments**. 10. ed. Hoboken: Wiley, 2019.

MÜLLER, M. D. et al. Estimating volume of eucalyptus trees for small-scale silvopastoral systems: do we really need height for accurate estimates? **European Journal of Forest Research**, v. 143, p. 349-364, 2024. DOI: 10.1007/s10342-023-01608-2.

NELDER, J. A.; WEDDERBURN, R. W. Generalized linear models. **Journal of the Royal Statistical Society: Series A (General)**, v. 135, n. 3, p. 370-384, 1972. DOI: 10.2307/2344614.

OLIVEIRA, T. A. et al. An Application of Generalized Additive Models of Location, Scale, and Shape (GAMLSS) to estimate the Eucalyptus Height. **Ciência e Natura**, v. 42, e15, p. 1-10, 2020. DOI: 10.5902/2179460X41710.

PIEPHO, H.-P.; MÖHRING, J.; SCHULZ-STREECK, T.; OGUTU, J. O. A stage-wise approach for the analysis of multi-environment trials. **Biometrical Journal**, v. 54, n. 6, p. 844-860, 2012. DOI: 10.1002/bimj.201100219.

R CORE TEAM. **R: A Language and Environment for Statistical Computing.**

Vienna, Austria: R Foundation for Statistical Computing, 2025. Disponível em:

<https://www.R-project.org/>. Acesso em: 20/05/2026.

RIEBL, H.; KLEIN, N.; KNEIB, T. Modelling intra-annual tree stem growth with a distributional regression approach for Gaussian process responses. **Journal of the Royal Statistical Society: Series C**, v. 72, n. 2, p. 414-433, 2023. DOI: 10.1093/jrssc/qlad015.

RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 54, n. 3, p. 507-554, 2005. DOI: 10.1111/j.1467-9876.2005.00510.x.

RIGBY, R. A. et al. **Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in R.** Boca Raton: Chapman and Hall/CRC, 2019. DOI: 10.1201/9780429298547.

SANTOS, v. G. et al. Eucalyptus diameter and volume prediction with deep neural networks: A Long Short-Term Memory model approach. **Expert Systems with Applications**, v. 267, 126529, 2025. DOI: 10.1016/j.eswa.2025.126529.

SCHAARSCHMIDT, F. **BSagri: Safety Assessment in Agricultural Field Trials.** R package version 0.1-10, 2023. Disponível em: <https://CRAN.R-project.org/package=BSagri>. Acesso em: 10/05/2026.

SILVA, J. A. A.; NETO, P. F. **Princípios básicos de dendrometria.** Recife: Imprensa Universitária da UFRPE, 1979. 198 p.

SILVA J. W. L. **Modelos estatísticos e técnicas de inteligência artificial para estimativa do volume de clones de *Eucalyptus* spp. com adição variáveis climáticas**, Tese de doutorado, Departamento de Estatística e Informática, Universidade Federal Rural de Pernambuco, Recife, p. 5-120, 2020. Disponível em: <http://www.tede2.ufrpe.br:8080/tede2/handle/tede2/8753>. Acesso em: 06 Jun. 2022.

SILVA H. D. et al. Comportamento essenciais florestais nas regiões árida e semiárida do Nordeste: resultados preliminares. Em Brasil. **EMBRAPA/FEZ.** p. 4, 1980.

SILVA, J. W. L.; SILVA, J. A. A.; TAVARES, J. A. Volumetric production of *Eucalyptus* spp. clones under different spacing in a severe drought period in the semi-arid region of Pernambuco, Brazil. **Floresta**, v. 52, n. 1, p. 150-158, 2022. DOI: 10.5380/rf.v52i1.78798.

STAPE, J. L. et al. Spacing and geometric layout effects on the productivity of clonal *Eucalyptus* plantations. **Forest Ecology and Management**, v. 510, 120103, 2022. DOI: 10.1016/j.foreco.2022.120103.

STASINOPOULOS, M. D.; RIGBY, R. A. Generalized additive models for location scale and shape (GAMLSS) in R. **Journal of Statistical Software**, v. 23, n. 7, p. 1-46, 2007. DOI: 10.18637/jss.v023.i07.

STASINOPOULOS, M. D. et al. **Flexible Regression and Smoothing: Using GAMLSS in R**. Boca Raton: CRC Press, 2017. DOI: 10.1201/b21973.

UMLAUF, N.; KLEIN, N.; ZEILEIS, A. BAMLSS: Bayesian Additive Models for Location, Scale, and Shape (and Beyond). **Journal of Computational and Graphical Statistics**, v. 20, n. 4, p. 806-820, 2021. DOI: 10.1080/10618600.2021.1909602.

VAN BUUREN, S.; FREDRIKS, M. Worm plot: a simple diagnostic device for modelling growth reference curves. **Statistics in Medicine**, v. 20, n. 8, p. 1259-1277, 2001.