

Modelagem GAMLSS no afilamento de árvores de *Tectona grandis* L.f

João Paulo Sardo Madi¹ John Paul McTague² Gabriel Agostini Orso¹ Julio Eduardo Arce¹ Allan Libanio Pelissari¹ Alexandre Behling¹ Rômulo Môra³ Juliano Bortolini⁴ Cyro Matheus Cometti Favalessa³

¹Universidade Federal do Paraná, Departamento de Ciências Florestais, Av. Prof. Lothário Meissner, 632, 80210-170, Curitiba-PR, Brasil

²Southern Cross Biometrics, 95172 Bermuda Dr., Fernandina Beach, FL 32034, Estados Unidos da América

³Universidade Federal de Mato Grosso, Faculdade de Engenharia Florestal, Av. Fernando Corrêa da Costa, 2367, Boa Esperança, 78060-900, Cuiabá-MT, Brasil

⁴Universidade Federal de Mato Grosso, Departamento de estatística/ Instituto de Ciências Exatas e da Terra, Av. Fernando Corrêa da Costa, 2367, Boa Esperança, 78060-900, Cuiabá-MT, Brasil

Original Article

*Corresponding author:
joaosardomadi@gmail.com

Palavras-chave:

Teca

Box Cox t

Gráfico de minhoca

Keywords:

Teak

Box Cox t

Worm plot

Received in
2023/07/21

Accepted on
2023/11/13

Published in
2023/12/31



DOI:
<http://dx.doi.org/10.34062/af.s.v10i4.15996>

RESUMO: Propõe-se neste estudo, verificar a utilização do modelo GAMLSS no ajuste de funções afilamento. Foram cubadas 52 árvores de *Tectona grandis* para o ajuste do polinômio do 5º grau por regressão linear múltipla e pela teoria dos modelos generalizados aditivos de parâmetros de locação, escala e forma (GAMLSS). A qualidade dos ajustes deu-se pelas estatísticas: coeficiente de correlação ao quadrado (R^2), erro padrão da estimativa (S_{yx} e $S_{yx}(\%)$), critério de informação Akaike (AIC), gráfico de minhoca, quantil quantil, dispersão residual e estatísticas auxiliares (viés (Bias), Raiz quadrada da soma do quadrado dos resíduos (REQM) e os resíduos percentuais (RP)). Foi verificado por meio da simulação de uma árvore e a aplicação dos diferentes métodos de ajuste, a classificação dos sortimentos estabelecidos, volume das seções classificadas e a estimativa da altura de ocorrência de diâmetros pré especificados. O modelo linear clássico ajustado foi mais eficiente segundo R^2 , S_{yx} e $S_{yx}(\%)$, porém, o modelo GAMLSS ajustado demonstrou maior eficiência pelo AIC, estatísticas auxiliares e pela análise gráfica dos resíduos. Na avaliação do efeito prático dos diferentes métodos utilizados, verificou que não houve diferença na classificação dos produtos, porém, o volume obtido dos mesmos pelo modelo linear clássico foi 5,47% maior em relação ao GAMLSS para o primeiro produto e as alturas de ocorrência dos diâmetros estabelecidos chegou até 37 cm. O modelo GAMLSS ajustado, em geral, apresentou estimativas de diâmetro com maior eficiência em estimar os valores reais

GAMLSS Modeling on stem taper of *Tectona grandis* L.f

ABSTRACT: In this study we propose GAMLSS to fit taper functions. We measured 52 *Tectona grandis* trees for fitting of the 5th degree polynomial by multiple linear regression and by the theory of generalized additive model for location, scale and shape (GAMLSS). The goodness of fit was determined by the statistics: squared coefficient of correlation (R^2), standard error of the estimate (S_{yx} and $S_{yx}(\%)$), Akaike information criterion (AIC), worm plot, quantile-quantile plots, residual dispersion and statistics Deviation (Bias), Root Sum Squared Residues (RMSE) and Percentage Residues (RP). Pre-established assortment sections were assessed by numeric integrals in both adjustment methods, as well as volume and height for a given diameter of each assortment section. The classic linear model was more efficient according to R^2 , S_{yx} and $S_{yx}(\%)$. However, the GAMLSS model was more efficient according to AIC, auxiliary statistics and graphical analysis of residuals. Assessing the practical effect of the different methods, we verified no differences in the classification of the products, moreover, the volume obtained from the classic linear model was 5.47% greater in relation to the GAMLSS for the first product and the heights occurrence of diameters reached up to 37 cm. The GAMLSS model, in general, presented diameter estimates with greater efficiency in estimating the actual values.

Introdução

Os plantios de *Tectona grandis* L. f. estão entre as espécies florestais cultivadas no Brasil, por conciliar a produtividade com madeira de alta qualidade para diversos fins (Chagas et al. 2014), como utilização em fabricação de móveis, pisos para interior, bancadas, janelas e construção naval (Koirala et al., 2021). Por consequência, a estimação de multiprodutos é essencial ao efetivo manejo dessa cultura, especialmente por meio de funções de afilamento que descrevem a forma do fuste das árvores (Sakici et al. 2008; Vendruscolo et al. 2016).

As funções de afilamento possibilitam a estimativa de diâmetros em diferentes alturas, bem como de alturas em diâmetros específicos, além da predição de volumes totais e parciais para determinação de sortimentos em nível de árvore. Para isso, os modelos polinomiais são largamente aplicados (Schröder et al. 2015; Scolforo et al. 2018; Dobner Junior et al. 2019; Nicoletti et al. 2019), os quais podem apresentar problemas de inferência estatísticas relacionados com a violação de pressupostos da regressão linear, como da normalidade dos resíduos, homogeneidade das variâncias e independência residual.

Com isso, diversas pesquisas foram desenvolvidas buscando a melhoria da qualidade das estimativas da forma das árvores (Mctague e Weiskittel 2021). Dentre os quais destacam-se os modelos de efeitos mistos não lineares (NLME) (Cao and Wang 2015; Özçelik e Alkan 2020; Zhang et al. 2021), modelos aditivos generalizados (GAM's) (Robinson et al. 2011) e redes neurais artificiais (RNA) (Schikowski et al. 2015; Martins et al. 2016).

Atualmente, vislumbra-se na literatura estatística os modelos aditivos generalizados para localização, escala e forma (GAMLSS - *Generalized Additive Model for Location, Scale and Shape*). Eles permitem que a média e os demais parâmetros de uma distribuição condicional da variável resposta, como a escala e a forma, sejam modelados como funções paramétricas e/ou não paramétricas aditivas às variáveis explicativas e/ou efeitos aleatórios (Rigby e Stasinopoulos 2005).

O parâmetro de escala está relacionado à uma medida de variância, ao passo que o parâmetro de forma se relaciona a curtose e a assimetria da variável resposta (Kneib 2013; Stasinopoulos et al. 2018). No caso do modelo do polinômio do 5º grau, a variável resposta é dada pela razão entre o diâmetro na seção i e o diâmetro à altura do peito (DAP). Na modelagem desses parâmetros em relação variável resposta, os modelos GAMLSS possuem mais de 50 famílias de distribuição (Kneib 2013), podendo citar distribuições Skew t (Villar et al. 2014),

Skew t type 3 (Oliveira et al. 2019), Box Cox t (Rigby e Stasinopoulos 2005, 2006) e Box Cox *power exponential* (BCPE) (Voudouris et al. 2012).

A vantagem dos GAMLSS fundamenta-se na manutenção da distribuição de probabilidade condicional da variável resposta, bem como na sua interpretação na escala original (Kneib 2013). Assim, a variável resposta pode assumir uma distribuição de probabilidade condicional distinta da gaussiana ou mesmo da família exponencial de distribuições, como diferentes graus de assimetria e curtose, que podem ser mais apropriadas para modelar um determinado conjunto de observações (Oliveira et al. 2019).

Dado o exposto, o objetivo deste trabalho foi verificar o potencial da modelagem GAMLSS para modelar o afilamento de árvores de *Tectona grandis*. Como hipótese, foi considerado que, ao estimar os parâmetros de locação, escala e formas de uma distribuição que expressa a relação de diâmetros ao longo do fuste há ganhos de ordem de inferência estatística para a predição de sortimentos para multiprodutos da madeira.

Material e Métodos

A pesquisa foi desenvolvida em povoamento de origem seminal da espécie *Tectona grandis* em uma propriedade localizada no município de Cuiabá, estado de Mato Grosso, com coordenadas 55°33'0,311"W 15°37'28,248"S. O clima da região é do tipo Aw (Clima Tropical com Estação seca no inverno) segundo a classificação de Köppen, com precipitação elevada, variando de 1.450 a 1.650 mm anuais, com 260 a 300 mm nos meses mais chuvosos e 10 a 20 mm nos meses mais secos (Alvares et al. 2013). A região de estudo apresenta temperatura média de 25°C, com a estação seca coincidindo com o inverno (Alvares et al. 2013).

A base de dados foi constituída pela cubagem de 52 árvores selecionadas aleatoriamente, de um povoamento de 12 anos de idade. A cubagem foi realizada pelo método de Smalian, em que os diâmetros foram medidos nas alturas de 0,2 m, 0,5 m, 0,7 m, 1,0 m, 1,3 m. As medidas subsequentes foram realizadas a cada um metro até a altura total da árvore, podendo a última seção ser inferior a um metro. As árvores amostradas estão distribuídas de acordo com as seguintes classes diamétricas: [10 cm, 16 cm) com 4 árvores, [16 cm, 22 cm) com 13 árvores, [22 cm, 28 cm) com 18 árvores, [28 cm, 34 cm) com 12 árvores e [34 cm, 40 cm) com 5 árvores. A Tabela 1 apresenta as estatísticas descritivas dos dados empregados neste estudo.

Tabela 1. Estatísticas descritivas das árvores amostradas no povoamento de *Tectona grandis*.

Variáveis	Mínimo	Média	Máximo	Desvio padrão	cv (%)	K	Sk
d (cm)	14,48	25,70	38,13	5,86	22,80	2,30	0,15
h (m)	11,3	19,4	25,4	3,09	15,91	2,61	-0,19

Em que: d é o diâmetro a 1,3 m do solo, h é a altura total, $cv(\%)$ é o coeficiente de variação em percentual, K é o coeficiente de curtose ($K < 0,263$ a distribuição é leptocúrtica, $K = 0,263$ é mesocúrtica e $K > 0,263$ é platicúrtica), Sk é o coeficiente de assimetria ($-1 < Sk > 1$ a distribuição é simétrica, $Sk > 1$ assimetria positiva, $Sk < -1$ assimetria negativa).

O modelo linear do polinômio do 5º grau (Schöpfer 1966) foi a função de afilamento ajustada aos dados de diâmetros e alturas ao longo dos fustes (Eq. 1). Para isso, duas abordagens foram aplicadas, sendo as modelagens via mínimos quadrados ordinários (MQO) e por GAMLSS.

$$\frac{d_i}{d} = \beta_0 + \beta_1 \left(\frac{h_i}{h}\right) + \beta_2 \left(\frac{h_i}{h}\right)^2 + \beta_3 \left(\frac{h_i}{h}\right)^3 + \beta_4 \left(\frac{h_i}{h}\right)^4 + \beta_5 \left(\frac{h_i}{h}\right)^5 + \varepsilon_i \quad (1)$$

em que d_i é o diâmetro medido na altura h_i ao longo do fuste (cm), d é o diâmetro a 1,3 m do solo (cm), h_i é altura da seção i do fuste (m), h é a altura total (m), β_i são os parâmetros estimados, e ε_i é o componente aleatório.

No modelo linear ajustado por MQO, a média da variável resposta d_i/d foi relacionada com a razão de (h_i/h) por meio da função polinomial descrita (Eq. 1), em que o componente aleatório (ε_i) pressupõem-se como normal de média zero e variância constante (σ^2). Os parâmetros (β_i) e σ^2 foram estimados por mínimos quadrados ordinários (Draper e Smith 2014).

Na modelagem por GAMLSS, funções paramétricas ou não-paramétricas devem ser definidas para relacionar os parâmetros de locação (μ), escala (σ) e formas (ν e τ) da distribuição da variável resposta com covariáveis (Stasinopoulos et al. 2017). Genericamente, os modelos GAMLSS podem ser representados da seguinte forma (Eq. 2).

$$Y \sim D(\mu, \sigma, \nu, \tau)$$

$$\begin{aligned} \eta_1 &= g_1(\mu) = X_1\beta_1 + s_{11}(x_{11}) + \dots + s_{1j_1}(x_{1j_1}) \\ \eta_2 &= g_2(\sigma) = X_2\beta_2 + s_{21}(x_{21}) + \dots + s_{2j_2}(x_{2j_2}) \\ \eta_3 &= g_3(\nu) = X_3\beta_3 + s_{31}(x_{31}) + \dots + s_{3j_3}(x_{3j_3}) \\ \eta_4 &= g_4(\tau) = X_4\beta_4 + s_{41}(x_{41}) + \dots + s_{4j_4}(x_{4j_4}) \end{aligned} \quad (2)$$

em que $D(\mu, \sigma, \nu, \tau)$ é uma distribuição com quatro parâmetros, μ é o parâmetro de locação, σ é o parâmetro de escala, ν e τ são os parâmetros de forma; η_i é o preditor linear do parâmetro i , g_i é a função de ligação do parâmetro i , X_i é uma matriz de ordem n . J_k de efeitos fixos para k^{th} parâmetros β_k ; x_{ij} é o j^{th} vetor da variável explicativa para k^{th} parâmetros θ_k .

No presente estudo, a variável resposta d_i/d , condicional à (h_i/h) , foi modelada pela distribuição Box-Cox t (BCT) (3). O modelo usado aqui não

contém efeitos aleatórios aditivos, usados para suavização, portanto a Eq. (2) pode ser reduzida e expressa como (Eq. 3):

$$\begin{aligned} \eta_1 &= g_1(\mu) = X_1\beta_1 \cdot \\ \eta_2 &= g_2(\sigma) = X_2\beta_2 \cdot \\ \eta_3 &= g_3(\nu) = X_3\beta_3 \cdot \\ \eta_4 &= g_4(\tau) = X_4\beta_4 \cdot \end{aligned} \quad (\text{Eq. 3})$$

Desta forma o polinômio do 5º grau, apresenta a seguinte estrutura (Eq. 4):

$$\begin{aligned} \frac{d_i}{d} \left| \left(\frac{h_i}{h} \right) \right. &\sim BCT(\mu, \sigma, \nu, \tau) \\ \mu &= \beta_0 + \beta_1 \left(\frac{h_i}{h}\right) + \beta_2 \left(\frac{h_i}{h}\right)^2 + \beta_3 \left(\frac{h_i}{h}\right)^3 + \beta_4 \left(\frac{h_i}{h}\right)^4 + \beta_5 \left(\frac{h_i}{h}\right)^5 \\ \log(\sigma) &= \hat{\gamma}_0 + \hat{\gamma}_1 \left(\frac{h_i}{h}\right) + \hat{\gamma}_2 \left(\frac{h_i}{h}\right)^2 + \hat{\gamma}_3 \left(\frac{h_i}{h}\right)^3 + \hat{\gamma}_4 \left(\frac{h_i}{h}\right)^4 + \hat{\gamma}_5 \left(\frac{h_i}{h}\right)^5 \\ \nu &= \nu_0 \\ \log(\tau) &= \tau_0 \end{aligned} \quad (\text{Eq. 4})$$

Nessa distribuição, o parâmetro μ corresponde à mediana, σ expressa o coeficiente de variação, enquanto ν e τ indicam assimetria e curtose (Stasinopoulos et al. 2017).

Para a análise da qualidade dos ajustes dos modelos para estimar os diâmetros ao longo dos fustes, foram calculadas as estatísticas do quadrado do coeficiente de correlação linear entre os valores observado e estimados (R^2), erro padrão da estimativa (SEE) e o critério de informação Akaike (AIC). A análise dos resíduos foi realizada por meio do gráfico *Worm plot* (Stasinopoulos et al. 2017), em que o eixo vertical retrata a diferença entre a localização teórica de cada observação e a distribuição empírica. Dessa forma, os dados no gráfico assumem a forma de uma linha que indica como eles diferem da distribuição assumida, sugerindo também modificações possíveis de serem realizadas no modelo (Buuren e Fredriks 2001). Além dessa análise, gráficos quantil-quantil e de resíduos padronizados foram gerados para o modelo ajustado por MQO e GAMLSS.

Adicionalmente, as estatísticas de acurácia foram calculadas por diâmetro mensurado ao longo do fuste (d_i), sendo o viés (Eq. 5), a raiz quadrada do erro médio quadrático (Eq. 6) e a soma dos resíduos percentuais (Eq. 7).

$$BIAS = \frac{\sum(d_i - \hat{d}_i)}{n} \quad (5)$$

$$REQM = \sqrt{\frac{\sum(d_i - \hat{d}_i)^2}{n-1}} \quad (6)$$

$$RP = \frac{\sum\left(\frac{d_i - \hat{d}_i}{d_i}\right)}{n} \cdot 100 \quad (7)$$

em que: BIAS é o viés, REQM é a raiz quadrada do erro médio quadrático, RP é a média dos resíduos percentuais (%), e n é a quantidade de medições no tronco somada para todas das árvores amostradas.

Uma classificação de notas foi elaborada para essas estatísticas, em que o valor 1 foi atribuído ao modelo que apresentou maior acurácia de acordo com cada estatística. Subsequentemente, valores superiores foram conferidos para os desempenhos menos acurados. Dessa forma, o método de ajuste do modelo que obteve a menor soma de notas foi considerado o mais eficiente em estimar diâmetros ao longo do fuste. De forma análoga ao diâmetro, uma classificação foi realizada para verificar em quais seções os diferentes modelos foram mais eficientes na obtenção do volume, os quais foram obtidos por meio da integração numérica.

Para verificar o efeito dos diferentes ajustes em termos práticos, foi elaborada a classificação de sortimentos para uma árvore simulada com diâmetro a 1,3 m do solo (d) de 38 cm e altura total (h) de 25 m a qual representa a árvore de maior dimensão do povoamento em estudo, a fim de obter o maior número de produtos. A classificação consistiu em produtos com comprimentos de 2,3 m, e diâmetros de ponta fina de 45,72 cm a 30,48 cm, S2 para 30,47 cm a 12,7 cm e S3 para 12,6 cm a 5,08 cm (Robinson e Hamann 2011). Após a simulação e a classificação dos produtos para os métodos de ajustes, foi verificada a diferença na quantidade de toras geradas por classes de sortimento, bem como se houve diferença na classificação do sortimento para uma mesma altura.

Após a obtenção dos produtos para os diferentes métodos de ajuste, a variável volume associada a cada tora classificada foi calculada, a fim de avaliar a diferença da mesma para os diferentes modelos. Para o cálculo do volume para as diferentes toras classificadas, foi aplicada a integral numérica (Eq. 7) para cada modelo, sendo que os limites utilizados na integral são as posições $h_{i's}$ em relação ao solo, referentes ao diâmetro inferior e superior de cada tora (Figura 1).

$$v = \frac{\pi}{40000} \int_{h_{inf}}^{h_{sup}} (d \cdot f(x))^2 dx \quad (7)$$

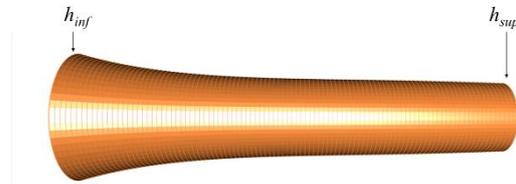


Figura 1. Esquema para demonstrar o cálculo do volume por meio da integral numérica. em que: v = volume (m^3); h_{inf} = altura inferior da tora (m) em relação ao solo; h_{sup} = altura superior da tora (m) em relação ao solo; d = Diâmetro à altura do peito; $f(x)$ = função de afilamento.

A partir dos volumes individuais calculados, foi verificado a diferença dos volumes obtidos por meio do GAMLSS em relação ao método dos MQO $\left(\frac{v_{MQO} - v_{GAMLSS}}{v_{MQO}} \cdot 100\right)$.

Além do volume para a mesma árvore simulada, foi verificado em que alturas $h_{i's}$ ocorreriam os diâmetros de 45 cm a cinco cm, em um intervalo de cinco cm, para os diferentes modelos ajustados por meio da técnica da bisettriz.

A análise dos dados foi realizada no software R (R Core team 2021), com os pacotes *GAMLSS* (Rigby e STasinopoulos 2005), *car* (Fox e Weisberg 2019) e *bbmle* (Bolker e R core team 2020).

Resultados e discussão

Os coeficientes de regressão e demais parâmetros estimados, como coeficiente de variação (σ), e curtose (τ) pelo método GAMLSS, foram estatisticamente significativos, ao nível de 5% de significância (Tabela 2). Não significância foi observada para o coeficiente de assimetria estimado (ν). As métricas estatística de coeficiente de regressão (R^2) e erro padrão da estimativa em percentual ($S_{yx}(\%)$) apresentaram valores semelhantes para dois métodos de estimação, demonstrando previsões semelhantes do perfil médio de árvores de *Tectona grandis*.

Contudo, enfatiza-se que as inferências acerca dessas medidas são prejudicadas quando não há atendimento dos pressupostos da regressão linear, como a normalidade residual e a homoscedasticidade (Poole e O'Farrell 1971; Lumley et al. 2002; Yang et al. 2019). Por outro lado, o critério de informação de Akaike (AIC) expressou melhoria da qualidade das estimativas por meio do método GAMLSS.

Tabela 2. Coeficientes estimados e estatísticas do modelo do polinômio do 5º grau ajustado por MQO e GAMLSS para *Tectona grandis*.

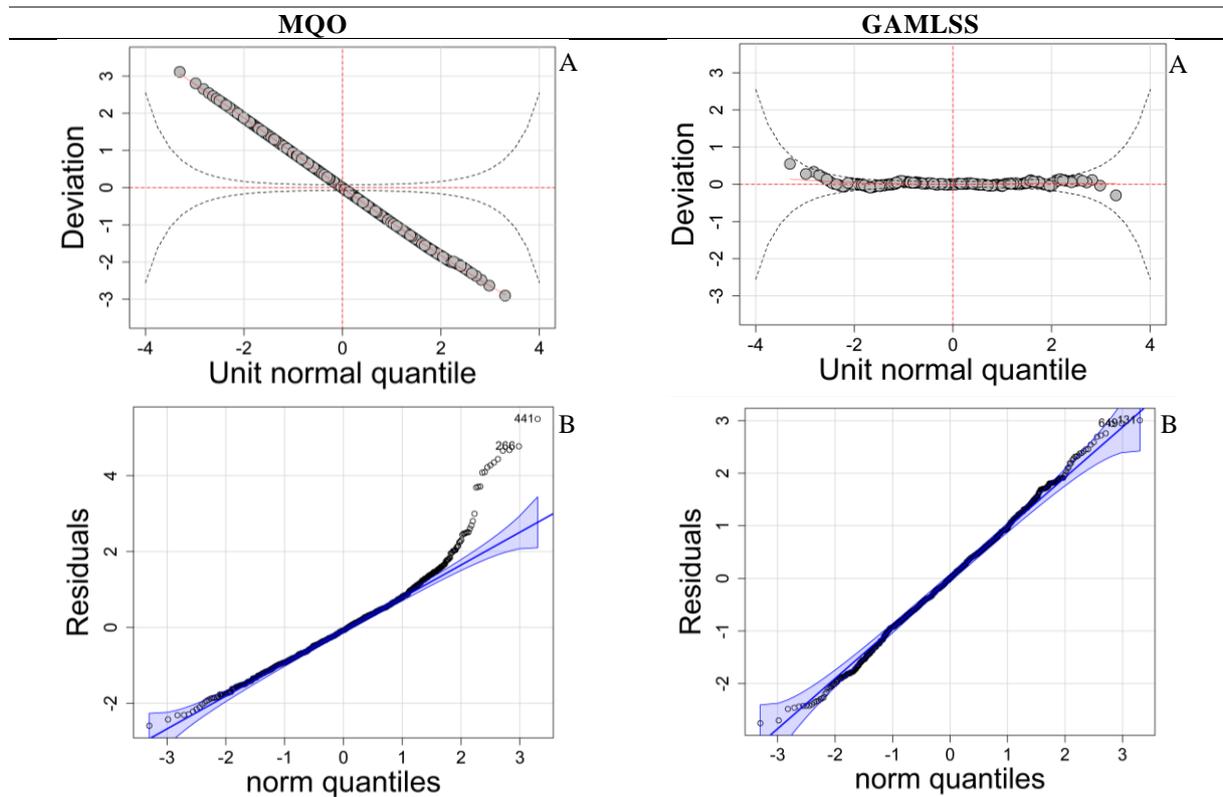
Método	Parâmetros estimados						Estatísticas		
MQO	β_0	β_1	β_2	β_3	β_4	β_5	R^2	$S_{yx}(\%)$	AIC
	1,34*	-6,51*	31,63*	-74,20*	77,26*	-29,70*			
GAMLSS	μ						0,94	10,84	-2785,2
	β_0	β_1	β_2	β_3	β_4	β_5			
	1,23*	-4,08*	16,89*	-36,95*	35,55*	-12,63*			
	$\log(\sigma)$								
	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5			
	-2,07*	-25,28*	163,03*	-406,26*	448,37*	-180,70*			
	v				$\log(\tau)$				
	-0,33 ^{ns}				1,98*				

μ é o parâmetro de locação (mediana), σ é o parâmetro de escala (coeficiente de variação baseado em percentis), v e τ são os parâmetros de forma de assimetria e curtose, respectivamente, β_i são os parâmetros estimados para o parâmetro μ , e γ_i são os parâmetros estimados para a transformação logarítmica do parâmetro σ , * significância ao nível de 5% e ^{ns} não significativo ao nível de 5%.

O componente paramétrico de Eq. 2 foi estimado por máxima verossimilhança penalizada. Por este motivo que os parâmetros ($\beta_0 - \beta_5$) estimados de GAMLSS tiveram um valor absoluto menor do que os valores ($\beta_0 - \beta_5$) estimados por MQO. A restrição ou penalização resulta num S_{yx} ligeiramente maior, mas em compensação, o método GAMLSS produz um perfil ou afilamento mais suave do que MQO.

Complementando às análises de qualidade dos ajustes, os gráficos de minhoca (*worm plots*) apresentados na Figura 2A possibilitaram retratar

a presença de expressiva tendência dos resíduos nas estimativas de diâmetros ao ajustar o modelo de afilamento pelo método MQO. Esse comportamento mostra que o modelo não se ajusta adequadamente aos dados (van Buuren e Fredriks 2001). Contudo, ao empregar o método GAMLSS, os resíduos se adequaram ao padrão das bandas de confiança retratadas pelas linhas tracejadas. Esse resultado demonstra vantagem no uso de métodos alternativos de estimação para modelagem do afilamento de árvores de *Tectona grandis*.



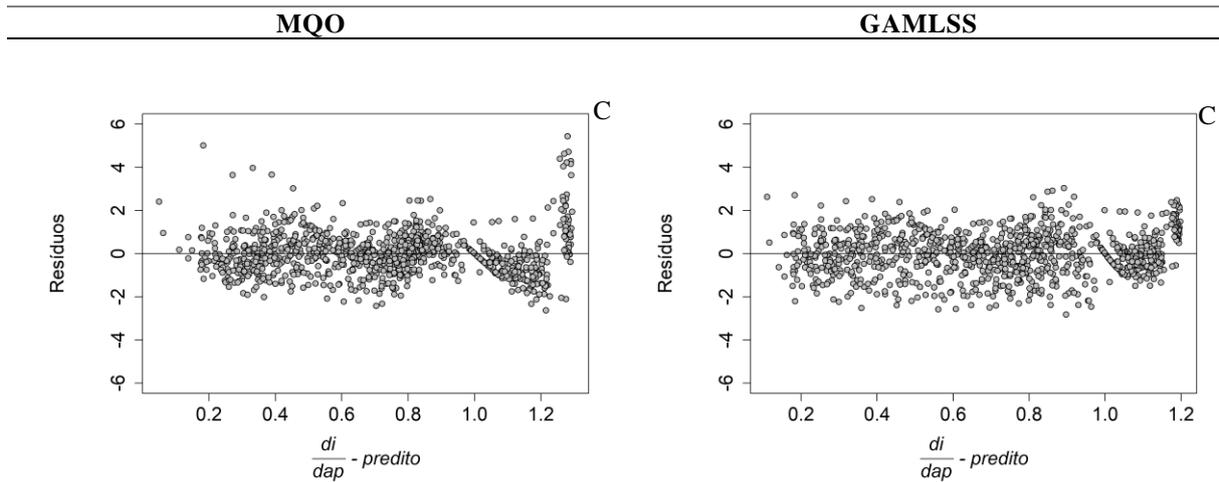


Figura 2. Diagnóstico dos resíduos do modelo do polinômio do 5º grau ajustado por MQO e GAMLSS para *Tectona grandis*. A - Worm plot, B – gráfico quantil-quantil, C - gráfico de resíduos padronizados.

Os gráficos quantil-quantil mostraram a presença de fuga dos resíduos na cauda direita pela modelagem MQO (Figura 2B), mostrando assimetria na distribuição residual, bem como ausência de normalidade (Marden 2004; Loy et al. 2016). Por outro lado, a modelagem GAMLSS acarretou na adequação dos resíduos à distribuição normal teórica, evidenciando os ganhos a serem alcançados quanto a inferência estatística do modelo ajustado ao perfil médio de árvores de *Tectona grandis*.

Adicionalmente, os gráficos de resíduos padronizados evidenciaram a melhoria das estimativas pelo método GAMLSS, ao adequá-los à dispersão aleatória ao longo do eixo dos diâmetros estimados (Figura 2C). Observou-se, também, a redução da presença de valores extremos, os quais foram fortemente atenuados

ao aplicar a modelagem GAMLSS, comparativamente ao método MQO.

A modelagem GAMLSS possibilitou a flexibilização das distribuições dos diâmetros ao longo do fuste de *Tectona grandis* (Figura 3). Verifica-se que as distribuições tendem ao comportamento simétrico nas primeiras seções de medição diâmetros pelo perfil, no entanto, com o avanço dos valores da variável independente da razão de (h_i/h) , a distribuição torna-se assimétrica positiva. Esse efeito pode ser resultante da presença de árvores com diferentes portes na base de dados, o que acarreta aumento da variabilidade, especialmente nas porções superiores do fuste. Corroborar-se essa hipótese ao considerar que fatores, como competição e dominância, tendem a influenciar na forma das árvores (Jacobs et al. 2020).

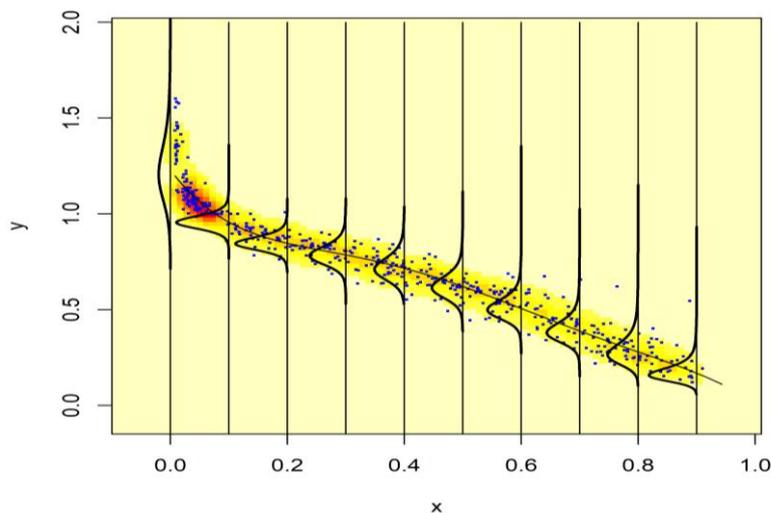


Figura 3. Distribuição condicional ajustada (BCT) da variável dependente y (d_i/d) e a mudança para os diferentes valores da variável independente x (h_i/h).

Por meio das estatísticas de acurácia aplicadas aos diâmetros estimados (d_i) e ao

volume (v_i) ao longo do fuste de *Tectona grandis* pelos dois métodos de estimação dos parâmetros

da função de afilamento (Tabela 3 e 4), foram observados tendência de superestimação do d_i e do v_i nas diferentes alturas (h_i) dado os valores negativos de Bias. Com isso, maiores vieses negativos foram notados nas últimas seções do fuste, o que demonstra a maior dificuldade das funções de afilamento não-segmentadas em explicar o ápice. Isso se deve a simplificação da forma do tronco ao considerar apenas um sólido geométrico, uma vez que somente uma equação é

utilizada para explicar o perfil, comparativamente aos modelos segmentados (Souza et al. 2016; Bernardi et al. 2021). Por outro lado, maiores valores positivos Bias foram encontrados na estimação dos diâmetros da base. Esses vieses positivos estão relacionados com a forma da base do fuste das árvores de *Tectona grandis*, as quais tendem a apresentar uniformidades maiores como foi relatado por Rocha et al. (2022) e Moya et al. (2020).

Tabela 3. Estatísticas de acurácia do modelo do polinômio do 5º grau ajustado por MQO e GAMLSS para estimativa dos diâmetros em diferentes alturas de *Tectona grandis*.

h_i (m)	MQO				GAMLSS				Result
	Bias (cm)	REQM (cm)	RP (%)	Score	Bias (cm)	REQM (cm)	RP (%)	Score	
0,2	2,78	4,17	7,51	3	5,04	6,03	14,01	6	MQO
0,5	-1,30	2,35	-4,55	6	0,21	1,85	0,61	3	GAMLSS
0,7	-1,53	2,15	-5,34	6	-0,44	1,32	-1,48	3	GAMLSS
1,0	-1,03	1,71	-3,37	6	-0,46	1,19	-1,35	3	GAMLSS
1,3	-0,68	1,24	-2,10	6	-0,51	0,94	-1,58	3	GAMLSS
2,3	0,43	1,36	2,47	6	-0,07	1,30	0,26	3	GAMLSS
3,3	0,88	1,61	4,35	6	0,33	1,46	1,95	3	GAMLSS
4,3	0,49	1,15	2,59	5	0,15	1,17	1,13	4	GAMLSS
5,3	0,23	1,15	1,23	5	0,14	1,22	0,94	4	GAMLSS
6,3	-0,13	1,33	-0,57	5	-0,06	1,36	-0,09	4	GAMLSS
7,3	-0,67	1,56	-3,57	6	-0,51	1,48	-2,76	3	GAMLSS
8,3	-0,85	2,10	-4,04	6	-0,71	1,98	-3,32	3	GAMLSS
9,3	-0,34	1,78	-1,15	6	-0,27	1,70	-0,87	3	GAMLSS
10,3	-0,15	1,91	0,13	5	-0,16	1,86	-0,06	4	GAMLSS
11,3	0,01	1,94	1,31	4	-0,02	1,96	1,18	5	MQO
12,3	0,02	1,95	0,82	3	0,02	1,99	1,28	5	MQO
13,3	-0,07	1,81	-1,40	6	-0,02	1,79	-0,47	3	GAMLSS
14,3	-0,15	2,10	-6,31	6	-0,02	1,98	-3,67	3	GAMLSS
15,3	-0,41	1,52	-8,36	6	-0,19	1,39	-4,56	3	GAMLSS
16,3	-0,21	1,77	-6,55	6	0,03	1,68	-3,06	3	GAMLSS
17,3	-0,41	1,64	-8,82	6	-0,15	1,57	-5,77	3	GAMLSS
18,3	-0,31	1,92	-6,65	6	0,10	1,87	-1,76	3	GAMLSS
19,3	-0,64	1,89	-8,64	6	-0,34	1,56	-5,55	3	GAMLSS
20,3	-1,06	1,95	-18,71	6	-0,81	1,62	-14,90	3	GAMLSS
21,3	-1,82	2,40	-41,35	5	-1,81	2,22	-42,23	4	GAMLSS

BIAS é o viés, REQM é a raiz quadrada do erro médio quadrático, RP é o resíduo percentual.

Tabela 4. Estatísticas de acurácia do modelo do polinômio do 5º grau ajustado por MQO e GAMLSS para estimativa dos volumes em diferentes alturas de *Tectona grandis*.

v_i -Section (m)		MQO				GAMLSS				Result
		Bias (m³)	REQM (m³)	RP (%)	Score	Bias (m³)	REQM (m³)	RP (%)	Score	
0,2	0,5	1,54E-03	4,21E-03	5,10	3	4,42E-03	6,19E-03	16,09	6	MQO
0,5	0,7	-1,39E-03	2,41E-03	-9,97	6	-1,10E-04	1,65E-03	-0,76	3	GAMLSS
0,7	1	-1,87E-03	2,99E-03	-8,81	6	-6,92E-04	1,86E-03	-2,77	3	GAMLSS
1	1,3	-1,29E-03	2,20E-03	-5,46	6	-7,63E-04	1,49E-03	-2,99	3	GAMLSS
1,3	2,3	-5,04E-04	4,53E-03	1,56	5	-1,27E-03	4,35E-03	-0,52	4	GAMLSS
2,3	3,3	2,20E-03	5,30E-03	7,20	6	1,44E-04	5,26E-03	2,43	3	GAMLSS
3,3	4,3	2,19E-03	4,60E-03	7,08	5	4,94E-04	4,66E-03	3,19	4	GAMLSS
4,3	5,3	9,96E-04	3,79E-03	3,55	5	1,31E-04	4,16E-03	1,84	4	GAMLSS
5,3	6,3	-5,88E-05	4,02E-03	0,45	3	-1,24E-04	4,33E-03	0,69	6	MQO
6,3	7,3	-1,59E-03	4,46E-03	-3,97	6	-1,22E-03	4,38E-03	-2,56	3	GAMLSS
7,3	8,3	-1,13E-03	1,31E-02	60,08	5	-6,63E-04	1,28E-02	61,09	4	GAMLSS
8,3	9,3	-2,30E-03	5,98E-03	-6,44	6	-1,94E-03	5,55E-03	-5,30	3	GAMLSS
9,3	10,3	-1,29E-03	4,97E-03	-2,43	6	-1,14E-03	4,66E-03	-2,36	3	GAMLSS
10,3	11,3	-7,98E-04	4,68E-03	0,45	5	-8,30E-04	4,55E-03	0,02	4	GAMLSS
11,3	12,3	-4,40E-04	4,30E-03	1,24	3	-5,30E-04	4,34E-03	1,31	6	MQO
12,3	13,3	-3,06E-04	3,64E-03	-3,75	4	-3,60E-04	3,70E-03	-2,51	5	MQO
13,3	14,3	-1,07E-04	3,21E-03	-11,35	6	-6,44E-05	3,18E-03	-6,80	3	GAMLSS
14,3	15,3	-3,00E-04	2,86E-03	-19,60	6	-1,30E-04	2,74E-03	-11,92	3	GAMLSS
15,3	16,3	-2,03E-04	2,51E-03	-13,86	6	5,00E-05	2,43E-03	-6,32	3	GAMLSS
16,3	17,3	-1,70E-04	2,60E-03	-18,36	5	1,91E-04	2,58E-03	-9,86	4	GAMLSS
17,3	18,3	-3,87E-04	2,55E-03	-17,78	6	1,73E-04	2,50E-03	-8,07	3	GAMLSS
18,3	19,3	-6,64E-04	2,49E-03	-15,17	6	-1,00E-04	2,22E-03	-4,10	3	GAMLSS
19,3	20,3	-1,39E-03	2,55E-03	-37,92	6	-7,89E-04	2,08E-03	-23,00	3	GAMLSS
20,3	21,3	-1,90E-03	2,93E-03	-83,56	6	-1,45E-03	2,27E-03	-67,17	3	GAMLSS

BIAS é o viés, REQM é a raiz quadrada do erro médio quadrático, RP é o resíduo percentual.

De forma complementar, ao avaliar a estatística de raiz quadrada do erro médio quadrático (REQM), as estimativas de diâmetros foram menos acuradas na base e no ápice do fuste de *Tectona grandis* pelos dois métodos de estimação (Tabela 3 e 4). De modo geral, as estimativas pelo método GAMLSS foram mais acuradas, comparativamente ao MQO. Contudo, maiores ganhos na acurácia pelo método MQO foram notados apenas em uma estimativa de uma das porções da base.

Os resultados das estatísticas de acurácia foram particularmente semelhantes ao empregar a estatística de resíduo percentual (RP), cujas subestimativas de d_i foram mais evidentes à 0,2 de altura, ao passo que superestimativas foram observadas à 20,3 e 21,3 de h_i . Dessa forma, ao considerar os *scores* das estatísticas de acurácia (Tabela 3), o método de estimação GAMLSS proporcionou, de modo geral, melhores estimativas dos diâmetros ao longo do fuste.

Ao verificar a aplicação prática dos diferentes métodos, no sortimento (Figura 4), nota-se que o GAMLSS classificou uma tora a mais no sortimento três (S3) em relação ao modelo linear clássico no ápice da árvore (Figura 4 A). Esta diferença na

classificação pode ser explicada, que ao observar a obtenção das alturas em que os diâmetros comerciais ocorrem (Figura 4 B), o GAMLSS, representado pela coloração vermelha, na maior parte das situações obteve os mesmos diâmetros em alturas inferiores em relação ao linear clássico. Para a espécie *Tectona grandis* cujo plantios tem por objetivo multiprodutos, as primeiras toras são as que possuem o maior valor comercial, desta forma a predição mais acurada atribuída ao GAMLSS por meio das análises realizadas, faz com que os volumes dos sortimentos, principalmente aqueles com maiores diâmetros, sejam realizadas de forma mais acurada, assim a diferença dos produtos obtidos em campo para o modelo ajustado seja menor.

Em relação as estimativas dos volume individuais das toras (Figura 4 A), nota-se que nas primeiras, a diferença no volume obtido (valores em vermelho), pode chegar entre 5% à 6%, e em relação das alturas de ocorrência dos diâmetros estipulados (Figura 4 B), é verificado diferenças entre 4 cm à 44 cm, o que pode acarretar, em outros cenários estipulados para os sortimentos, assim alteração na classificação dos produtos, uma vez que o mesmo diâmetro pode ocorrer até 44 cm de diferença de um método em relação ao outro.

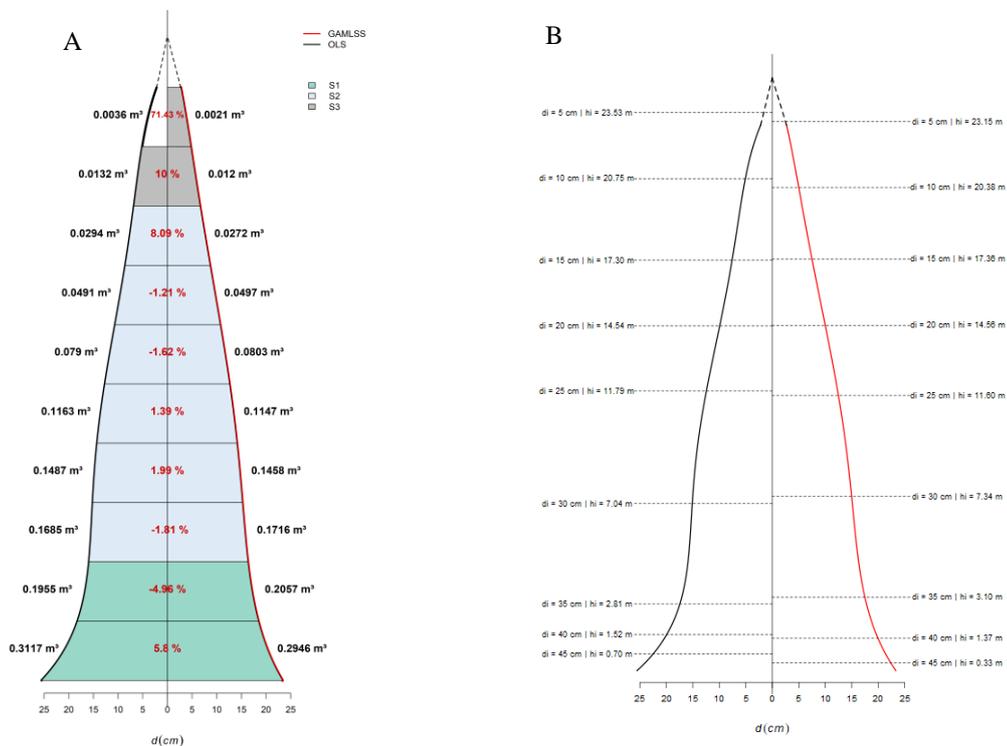


Figura 4. Simulação do uso do modelo do polinômio do 5º grau ajustado por MQO e GAMLSS para estimativa de sortimentos de *Tectona grandis*. A – Classificação dos produtos e volumes individuais das toras para cada método de ajuste; B – Alturas estimadas pelo método da bisetrix para os diâmetros especificados, para os diferentes métodos de ajuste.

Conclusão

O método GAMLSS para estimativa do afilamento de *Tectona grandis* foi mais eficiente nas estimativas dos diâmetros ao longo do fuste, proporcionando ganhos de inferência estatística.

Na avaliação do efeito prático para os diferentes ajustes notou-se diferentes resultados obtidos para as diferentes metodologias, para o volume, alturas estimadas e para a classificação dos produtos diante os diâmetros de ponta fina estabelecidos.

Referências

Alvares CA, Stape JL, Sentelhas PC, de Moraes Gonçalves JL, Sparovek G (2013) Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6): 711–728. doi: 10.1127/0941-2948/2013/0507

Bernardi LK, Thiersh MFBM., Arteaga AJM., de Almeida A.A.A., de Pádua FA, Thiersch CR (2021) Diferentes modelos para o afilamento do tronco de *Eucalyptus* sp. para o cenário florestal brasileiro. *Ciência Florestal*, 31(3):1364–1382. doi: 10.5902/1980509840376

Bolker B, R Development Core Team (2020) *bbmle: Tools for General Maximum Likelihood Estimation*. R package version 1.0.23.1. <https://CRAN.R-project.org/package=bbmle>

Buuren SV, Fredriks M (2001) Worm plot: a simple diagnostic device for modelling growth reference curves. *Statist. Med*, 20:1259-1277.

Cao QV, Wang J (2015) Evaluation of Methods for Calibrating a Tree Taper Equation. *Forest Science*, 61(2): 213-219. doi: 10.5849/forsci.14-008.

Carvalho SPC, de Mendonça AR, de Lima MP, Calegario N (2010) Different strategies to estimate the commercial volume of *Anadenanthera colubrina* (Vell.) Brenan. *Cerne*, 16(3): 399–406. doi: 10.1590/S0104-77602010000300016.

Chagas SF, Evangelista WV, Silva JC, Macedo A M, Carvalho L (2014) Propriedades da madeira de teca de diferentes idades e oriundas de desbaste. *Ciência da madeira*, 5 (2): 138-150. doi: 10.12953/2177-6830.v05n02a08

Dobner Junior M, Nicoletti MF, Novack Junior NS, Nunes GC (2019) Growth of *Pinus glabra* in southern Brazil. *Floresta*, 49(2): 277–286. 2019. doi: 10.5380/ufv.v49i2.57531.

Draper NR, Smith H. *Applied Regression Analysis*. Wiley-Interscience. 3° ed. 2014. 736 p.

Favalessa CMC, Ubialli JA, Caldeira SF, Drescher R, Acosta FC (2012) Equações de sortimentos para *Tectona grandis* na região centro – sul de Mato Grosso. *Pesquisa Florestal Brasileira*, 32(72): 389–399. doi: 10.4336/2012.pfb.32.72.389.

Fox J, Weisberg S (2019) *An {R} Companion to Applied Regression*. 3° Ed. Thousand Oaks CA: Sage. 608 p. URL: <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>

IBÁ – Indústria Brasileira de Árvores. (2019) *Relatório 2019*. IBÁ.

Jacobs M, Rais A, Pretzsch H (2020) Analysis of stand density effects on the stem form of Norway spruce trees and volume miscalculation by traditional form factor equations using terrestrial laser scanning (TLS). *Canadian Journal of Forest Research*, 50(1):51–64. doi: doi.org/10.1139/cjfr-2019-012.

Kneib T (2013) Beyond mean regression. *Statistical Modelling*, 13(4), 275-203. doi: doi.org/10.1177/1471082X13494159

Koirala A, Montes CR, Bullock BP, Wagle BH (2021) Developing taper equations for planted teak (*Tectona grandis* L.f.) trees of central lowland Nepal. *Trees, Forests and People*, 5: 1–9. doi: 10.1016/j.tfp.2021.100103

Lansanova RL, Machado AS, Orso GA, Pelissari AL, Figueiredo Filho A, Silva FA (2020) Calibration of a mixed-effect stem taper model for *Tectona grandis*. *Journal of Tropical Forest Science*, 32(4): 341-348. doi: 10.26525/jtfs2020.32.4.341.

Lansanova LR, Machado AS, Garrett ATA, Bonete IP, Pelissari AL, Figueiredo Filho A, Silva FA, Ciarnoschi LD (2019) Mixed-effect non-linear modelling for diameter estimation along the stem of *Tectona grandis* in mid-western Brazil. *Southern Forests*, 81(2): 167–173. doi: 10.2989/20702620.2018.1531279.

Loy A, Follett L, Hofmann H. Variations of Q – Q Plots: The Power of Our Eyes! *The American Statistician*, 70(2): 202–214. doi: doi.org/10.48550/arXiv.1503.02098

Lumley T, Diehr P, Emerson S, Chen L (2022) The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health*, v. 23, n. 1, p. 151–169, 2002.

Marden JI (2004) Positions and QQ Plots. *Statistical Science*, 19(4): 606-614. doi: 10.1214/088342304000000512

- Martins ER, Binoti MLMS, Leite HG, Binoti DHB, Dutra GC (2016) Configuração de redes neurais artificiais para estimação do afilamento do fuste de árvores de eucalipto. *Revista Brasileira de Ciências Agrárias*, 11(1):33-38. doi: 10.5039/agraria.v11i1a5354.
- Mctague JP, Weiskittel A (2021) Evolution, history, and use of stem taper equations: a review of their development, application, and implementation. *Can. J. For. Res.*, 51:210-235. doi: 10.1139/cjfr-2020-0326.
- Moya R., Gaitán-Álvarez J, Ortiz-Malavassi E, Berrocal A, Fernández-Sólis D (2020) Equations for predicting heartwood merchantable volume and tradable sawlog in *Tectona grandis*. *Journal of Tropical Forest Science*, 32(4):379-390. doi: doi.org/10.26525/jtfs2020.32.4.379
- Nicoletti MF, Carvalho SPC, Machado SA, Filho AF, Oliveira GS (2019) Partial volume prediction through nonlinear mixed modeling. *Floresta e Ambiente*, 26 (4):1–10. doi: 10.1590/2179-8087.032917.
- Poole MA, O'Farrell PN (1971) The Assumptions of the Linear Regression Model. *Transactions of the Institute of British Geographers*, 52:145.
- Oliveira TA, XAVIER JÚNIOR SFA, FÁRIA GA, Lopes BG, Barbosa EM, Peixoto APB (2020) An Application of Generalized Additive Models of Location, Scale, and Shape (GAMLSS) to estimate the Eucalyptus Height. *Ciência e Natura*, 42: 1–10. doi: 10.5902/2179460X41710.
- Özçelik R, Alkan O (2020) Fitting and calibrating a mixed-effects segmented taper model. *Cerne*, 26(4): 464-473. doi: 10.1590/01047760202026032737.
- R Core Team. R (2019) *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rigby RA, Stasinopoulos DM (2004) Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Statistics in Medicine*, 23: 3053–3076. doi: 10.1002/sim.1861.
- Rigby RA, Stasinopoulos DM (2006) Using the Box-Cox *t* distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, 6: 209–229. doi: 10.1191/1471082X06st122oa.
- Rigby RA, Stasinopoulos MD (2005) Generalized additive models for location, scale and shape. *Appl. Statist.*, 54(3): 507–554. doi: 10.1111/j.1467-9876.2005.00510.x.
- Robinson AP, Hamann JD (2011) *Forest Analytics with R*. Springer: New York. 341 p.
- Robinson AP, Lane SE, Thérien G (2011) Fitting forestry models using generalized additive models: a taper model example. *Can. J. For. Res.*, 41(10): 1909-1916. 2011. doi: 10.1139/x11-095
- Rocha KJ, Finger CAG, Favalessa CMC, Caldeira SF, Fleig FD (2022) Form and Volume of the Stem of *Tectona grandis* L.f. in the Central-Western Region of Brazil. *Forests*, 13(11):1818. doi: doi.org/10.3390/f13111818.
- Stasinopoulos MD, Rigby RA, De Bastiani F (2018) GAMLSS: A distributional regression approach. *Statistical Modelling*, 18(3-4): 248-273. doi:10.1177/1471082X18759144
- Sakici OE, Misir N, Yavuz H, Misir M (2008) Stem taper functions for *Abies nordmanniana* subsp. *bornmulleriana* in Turkey. *Scandinavian Journal of Forest Research*, 23(6): 522-533. doi: doi.org/10.1080/02827580802552453.
- Schikowski AB, Corte APD, Sanquetta CR (2015) Estudo da forma do fuste utilizando redes neurais artificiais e funções de afilamento. *Pesq. flor. bras.*, 35(82): 119-127. doi: 10.4336/2015.pfb.35.82.867.
- Schröder T, Costa EA, Valério AF, Lisboa GS (2015) Taper equations for *Pinus elliottii* Engelm. in southern Parana, Brazil. *Forest Science*, 61(2): 311-319. doi: 10.5849/forsci.14-054.
- Scolforo HF, Mctague JP, Raimundo MR, Weiskittel A, Carrero O, Scolforo JRS (2018) Comparison of taper functions applied to eucalypts of varying genetics in Brazil: Application and evaluation of the penalized mixed spline approach. *Canadian Journal of Forest Research*, 48(5): 568-580. doi: 10.1139/cjfr-2017-0366.
- Souza RR, Nogueira GS, Júnior LSM, Oliveira MLR, Abrahão CP, Leite HG (2016) Forma de fuste de árvores de eucalipto em plantios adensados. *Scientia Forestalis*, 44(109):33-40. doi:dx.doi.org/10.18671/scifor.v44n109.03.
- Stasinopoulos MD, Rigby RA, Heller GZ, Voudouris V, de Bastiani F (2017) *Flexible Regression and Smoothing Using GAMLSS in R*. 1° ed. 572 p.
- Voudouris V, Gilchrist R, Rigby R, Sedgwick J, Stasinopoulos D (2012) Modelling skewness and kurtosis with the BCPE density in GAMLSS.

Journal of Applied Statistics, 39(6): 1279-1293. doi: doi.org/10.1080/02664763.2011.644530

van Buuren S, Fredriks M (2001) Worm plot: a simple diagnostic device for modelling growth reference curves. *Statistics in Medicine*, 20: 1259-1277.

Vendruscolo DGS, Drescher R, Carvalho SPC, Souza HS, Chaves AGS (2016) Forma do fuste de árvores de *Tectona grandis* em diferentes espaçamentos. *Advances in Forestry Science*, 3(3): 51-54.

Yang K, Tu J, Chen T (2019) Homoscedasticity: an overlooked critical assumption for linear regression. *General Psychiatry*, 32(5), p. e100148. doi: 10.1136/gpsych-2019-100148

Zhang S, Sun J, Duan A, Zhang J. Variable-Exponent Taper Equation Based on Multilevel Nonlinear Mixed Effect for Chinese Fir in China. *Forests*, 12(2):126. doi: 10.3390/f12020126.