



Modelagem hipsométrica com uso de covariáveis em Floresta Estacional Semidecidual

Bruno Fonseca MARTINS ¹, Naila Garcia BRAGA ¹, Julyana Marques Lobão de ALMEIDA ¹,
Eduardo Vinícius da SILVA ¹, Rafaella De Angeli CURTO ^{*1}, Emanuel José Gomes de ARAÚJO ¹

¹ Instituto de Florestas, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, Brasil.

*E-mail: rafaellacurto@ufrj.br

Submetido em: 05/03/2026; Aceito em: 20/08/2026; Publicado em: 01/09/2026.

RESUMO: Objetivou-se avaliar o desempenho de modelos tradicionais e mistos no ajuste da relação hipsométrica em fragmento de Floresta Estacional Semidecidual, utilizando dados de 19 unidades amostrais localizadas no Parque Estadual de Nova Baden, em Lambari, Minas Gerais. Foram criados cenários: um a partir do banco de dados original e outro após a exclusão de pontos discrepantes, para ajustes de modelos de relação hipsométrica, que foram avaliados por meio do coeficiente de determinação ajustado, do erro padrão da estimativa, dos critérios de informação de Akaike e Bayesiano, da análise gráfica de resíduos, do teste F e da significância dos coeficientes. A melhor equação foi submetida à inclusão das covariáveis espécie, família botânica e classe de diâmetro, consideradas como efeitos aleatórios. O modelo linear de Trorey foi considerado o mais preciso em ambos os cenários. A remoção dos pontos discrepantes resultou em estimativas mais precisas e menos tendenciosas, que, quando associadas à modelagem mista, com a inclusão do fator aleatório “espécie” nos três coeficientes do modelo, propiciaram um aumento de 18% no coeficiente de determinação ajustado e uma redução de 7,2% no erro padrão da estimativa.

Palavras-chave: relação altura-diâmetro; modelos mistos; floresta nativa.

Hypsometric modeling using covariates in a semideciduous seasonal forest

ABSTRACT: The objective of this study was to evaluate the performance of traditional and mixed models in fitting the hypsometric relationship in a fragment of Semideciduous Seasonal Forest, using data from 19 sample plots located in Nova Baden State Park, in Lambari, Minas Gerais, Brazil. Two scenarios were developed: one based on the original database and another after removing outliers, for fitting hypsometric relationship models. The models were evaluated using the adjusted coefficient of determination, standard error of the estimate, Akaike Information Criterion, Bayesian Information Criterion, graphical residual analysis, F-test, and coefficient significance. The best equation was then extended to include covariates such as species, botanical family, and diameter class, considered as random effects. The Trorey linear model was considered the most accurate in both scenarios. The removal of outliers resulted in more concise and less biased estimates, which, when associated with the modeling the inclusion of the random factor "species" to an 18% increase in the adjusted coefficient of determination and a 7.2% reduction in the standard error of the estimate.

Keywords: height-diameter relationship; mixed models; native forest.

1. INTRODUÇÃO

O inventário florestal é uma atividade essencial para o manejo adequado dos recursos florestais, possibilitando a obtenção de estimativas precisas das características quantitativas dos povoamentos e das florestas (SCOLFORO; MELLO, 2006).

As florestas ineqüiâneas, mesmo sendo pequenos fragmentos e áreas de vegetação secundária jovem, apresentam estruturas e composições florísticas complexas, com alto potencial de armazenamento de carbono (POORTER et al., 2016; MARQUES; GRELE, 2021). A produtividade do ecossistema e a eficiência no uso do carbono são amplamente explicadas por características da estrutura da vegetação diretamente ligadas à complexidade estrutural, incluindo a biomassa acima do solo, o índice de

área foliar e a altura das árvores (MIGLIAVACCA et al., 2021; CONTO et al., 2024; FONSÊCA et al., 2024).

Dentre as variáveis obtidas em um inventário florestal, a altura é uma importante característica das árvores, pois está diretamente correlacionada com o volume de madeira. Em florestas naturais, a altura total das árvores tem grande importância ecológica e para fins de manejo, pois ajuda a compreender a estrutura vertical da comunidade (CURTO et al., 2018).

No inventário florestal, o diâmetro à altura de 1,30 metros do solo (DAP) é facilmente mensurado, enquanto a medição da altura das árvores em pé é onerosa devido à maior complexidade e à maior suscetibilidade a erros (SILVA et al., 2007). Diversos obstáculos são encontrados na medição da altura em florestas naturais, como a falta de acessibilidade do

local, a tortuosidade do fuste, a falta de visibilidade da base e do topo da árvore e a declividade do terreno. Por consequência, a medição da altura requer maiores investimentos, mais tempo e equipe treinada para a coleta de dados (SENA et al., 2015).

No inventário florestal é usual medir o diâmetro de todas as árvores da parcela e a altura apenas de uma parte delas. Esse conjunto de dados é utilizado para estabelecer uma relação entre a altura e o diâmetro, por meio de análise de regressão, denominada relação hipsométrica (ANDRADE; LEITE, 2011). Com base na equação obtida, a altura das demais árvores da parcela é estimada, a partir do DAP mensurado (BARTOSZECK et al., 2002).

Embora vários modelos possam adequar-se à relação hipsométrica, o uso imoderado dessa relação pode levar a erros não amostrais consideráveis, uma vez que diversos fatores influenciam a relação entre a altura e o diâmetro (BARROS et al., 2002). Na literatura florestal, diversos autores relatam a influência desses fatores, tais como a qualidade do sítio, a densidade do povoamento, a competição intraespecífica, a família botânica e as condições ambientais (CUI et al., 2022; JHA et al., 2023; FENG et al., 2024).

Neste contexto, a modelagem mista, surge como alternativa à modelagem tradicional, pois permite incorporar a influência da variabilidade de diferentes fatores nos ajustes dos modelos (OLIVEIRA et al., 2015; CYSNEIROS et al., 2020; SILVA et al., 2024; RIBEIRO et al., 2026). Autores como Cui et al. (2022) e Yang et al. (2022), demonstraram que a inclusão de fatores aleatórios em modelos resulta em ajustes mais precisos, superiores aos obtidos com modelos tradicionais que utilizam apenas o DAP como variável independente. Conforme apresentado por Pinheiros e Bates (2000), a modelagem de efeitos mistos permite que a variação interindividual seja explicada por parâmetros de efeitos fixos, comuns a toda a população, e por parâmetros de efeitos aleatórios, específicos de cada indivíduo.

Apesar dos avanços na modelagem da relação hipsométrica, os estudos que avaliam o efeito de covariáveis na Floresta Estacional Semidecidual ainda são limitados, evidenciando uma lacuna quanto à disponibilidade de equações eficientes para estimar a altura das árvores em florestas com elevada diversidade de espécies. Portanto, foi formulada a hipótese de que a inclusão de covariáveis capazes de representar características individuais das árvores e condições da floresta reduzirá a variabilidade não explicada pelo modelo e, proporcionará maior precisão na predição da altura das árvores em comparação aos modelos de efeitos fixos sem esses componentes.

Deste modo, o objetivo deste trabalho foi avaliar o efeito da adição das covariáveis espécie, família botânica e classe de diâmetro no ajuste de modelos de relação hipsométrica em Floresta Estacional Semidecidual.

2. MATERIAL E MÉTODOS

2.1. Caracterização da área de estudo

A área de estudo é um fragmento de Floresta Estacional Semidecidual, localizado no Parque Estadual de Nova Baden, com coordenadas centrais 21°56'09.6"S e 45°19'19.8"W, no município de Lambari, no sul de Minas Gerais (Figura 1). A área está situada em uma porção do relevo conhecida como Planalto Atlântico, na Serra da Mantiqueira, com área total de 268,48 ha, sendo a área amostral de 1,9 ha.

A altitude média do local é de 1.073 m, com predominância de Latossolo Vermelho Distrófico

(MARQUES NETO, 2014). O clima é do tipo "Cwa", segundo a classificação de Köppen, caracterizado por apresentar inverno seco e verão quente, definido como temperado úmido (MARTINS et al., 2018). O índice de chuvas anual chega a superar 1.600 mm e a temperatura média anual é de 17,67 °C (GUIMARÃES et al., 2010).

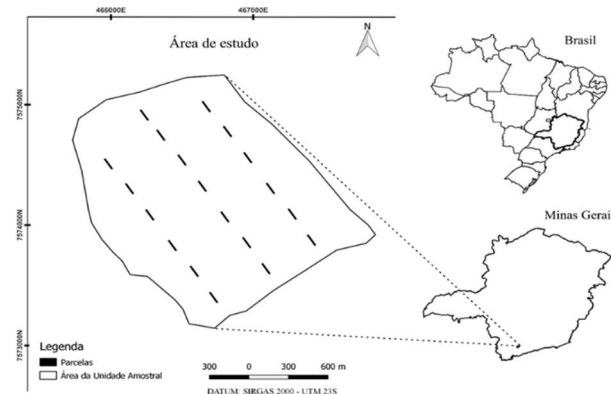


Figura 1. Localização da área de estudo e dos pontos amostrais no Parque Estadual de Nova Baden, em Lambari, Minas Gerais.

Figure 1. Location of the study area and sampling points in Parque Estadual de Nova Baden, Lambari, Minas Gerais.

2.2. Banco de dados

O banco de dados foi coletado durante o Inventário Florestal de Minas Gerais (CARVALHO et al., 2008). O processo de amostragem utilizado foi o sistemático, com 19 unidades amostrais de 1.000 m² (10 x 100 m). Em cada unidade amostral foram mensuradas as variáveis circunferência a 1,30 m do solo (CAP) e altura total (Ht) dos indivíduos arbóreos, além da identificação botânica das espécies. Posteriormente, os valores de CAP foram convertidos em diâmetro à altura de 1,30 m do solo (DAP). Para o objetivo deste trabalho foram excluídas todas as 95 árvores mortas e, para aquelas que possuíam bifurcações, foi calculado o DAP equivalente e a altura média dos fustes. O conjunto de dados foi composto por 210 espécies, pertencentes a 69 famílias botânicas.

2.3. Modelagem hipsométrica

Este estudo foi realizado considerando dois cenários (I e II). No cenário I, os modelos de relação hipsométrica foram ajustados a partir do banco de dados original, contendo informações de 2.105 árvores. No cenário II, os modelos foram ajustados com base no mesmo banco de dados, após a exclusão dos pontos discrepantes entre si, como outliers e pontos influentes. A remoção desses pontos foi realizada por meio da metodologia dos resíduos padronizados e da medida de Leverage (ROUSSEUW; LEROY, 1987). Do banco de dados original, foram removidas 70 árvores, resultando em um total de 2.035 árvores no cenário II.

Para avaliar a acurácia e precisão das equações ajustadas, o conjunto de dados foi dividido em 80% para ajuste e 20% para validação, em ambos os cenários. A seleção das árvores para ajuste e validação foi realizada aleatoriamente, proporcionalmente à frequência das classes de diâmetro à altura de 1,30m do solo (DAP). Foram utilizadas 12 classes, com amplitude de 10 cm, no primeiro cenário, e 12 classes, com amplitude de 6 cm, no segundo cenário. Em ambos os cenários, dez modelos tradicionais de relação hipsométrica foram ajustados (Tabela 1), sendo oito lineares e dois não lineares.

Tabela 1. Modelos de relação hipsométrica ajustados para um fragmento de Floresta Estacional Semidecidual, em Lambari, Minas Gerais.

Table 1. Adjusted hypsometric relationship models for a fragment of Semideciduous Seasonal Forest, in Lambari, Minas Gerais.

Nº	Modelo	Denominação
1	$Ht = \beta_0 + \beta_1 \cdot DAP$	Linear simples
2	$Ht = \beta_0 + \beta_1 \cdot DAP + \beta_2 \cdot DAP^2$	Trorey
3	$Ht = \beta_0 + \beta_1 \cdot \frac{1}{DAP}$	Assmann
4	$Ht = \beta_0 + \beta_1 \cdot \ln DAP$	Henricksen
5	$\ln Ht = \beta_0 + \beta_1 \cdot \ln DAP$	Stoffels
6	$\ln Ht = \beta_0 + \beta_1 \cdot \frac{1}{DAP}$	Curtis
7	$\frac{1}{Ht - 1,30} = \beta_0 + \beta_1 \cdot \frac{1}{DAP}$	Petterson
8	$\frac{DAP^2}{Ht} = \beta_0 + \beta_1 \cdot DAP + \beta_2 \cdot DAP^2$	Prodan
9	$Ht - 1,3 = \frac{DAP^2}{\beta_0 + \beta_1 \cdot DAP^2}$	Naslund
10	$Ht - 1,3 = \frac{DAP^2}{\beta_0 + \beta_1 \cdot DAP + \beta_2 \cdot DAP^2}$	Prodan

$\ln$ = logaritmo neperiano; Ht = altura total (m); DAP = diâmetro à 1,30 m do solo (cm); $\beta_0, \beta_1, \beta_2$ = coeficientes de regressão

Em cada cenário, o modelo selecionado foi ajustado com a inclusão das covariáveis espécie, família botânica e classe de DAP como fatores aleatórios. Os efeitos fixos modelam o valor esperado da variável altura, enquanto os efeitos aleatórios modelam sua estrutura de covariância. As covariáveis foram incorporadas aos coeficientes $\beta_0, \beta_1, \beta_2$ e na interação entre os três coeficientes, buscando avaliar a combinação de maior efeito aleatório.

Portanto, o modelo misto possui uma porção fixa e outra aleatória e, como exemplo de demonstração, apresenta-se a equação 1, que representa um modelo linear misto. Detalhes sobre a modelagem mista podem ser encontrados em Pinheiros e Bates (2000) e McCulloch e Searle (2001).

$$Y_i = X_{\beta} + Z_{\tau} + \varepsilon_i \quad (01)$$

em que: Y_i = vetor da variável altura do i -ésimo grupo de árvores; X = matriz conhecida da variável DAP (cm) do i -ésimo grupo de árvores; β = vetor de parâmetro de efeitos fixo; Z = matriz relativa aos efeitos aleatórios (espécie, família e classe de DAP); τ = vetor de parâmetros aleatórios associado ao i -ésimo grupo de árvores; ε_i = vetor de erros aleatório associado a variável Y_i .

Os parâmetros aleatórios τ e ε_i não são correlacionados e apresentam distribuição normal com média zero e matrizes de variâncias-covariâncias D e R , respectivamente. Assumindo: $\tau \sim N(0, D)$ e $\varepsilon_i \sim N(0, R)$, em que D é a matriz de variâncias-covariâncias dos efeitos aleatórios e R é a matriz de variâncias-covariâncias residuais.

Os parâmetros aleatórios dos modelos mistos foram preditos usando o melhor preditor linear não viesado – BLUP, e os parâmetros dos efeitos fixos foram estimados usando o melhor estimador linear não viesado – BLUE (MCCULLOCH; SEARLE, 2001). Desse modo, a solução conjunta para estimar os valores de β e prever os valores de τ , passa pela resolução por um sistema de equações (Equação 2).

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + D^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\tau} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \quad (02)$$

Para a estimação e predição dos parâmetros $\hat{\beta}$ e $\hat{\tau}$ respectivamente, do modelo, é necessário estimar os parâmetros de R e D , que são desconhecidas na maioria das vezes. A estimação de D e R pode ser realizada por vários métodos, dentre os quais destaca-se o método de máxima verossimilhança restrita (REML) (MCCULLOCH; SEARLE, 2001). Isso resulta nas estimativas dos parâmetros fixos, nas variâncias e covariâncias dos efeitos aleatórios, e na variância do resíduo do modelo selecionado.

2.4. Ajuste e critérios de avaliação dos modelos

A precisão dos modelos tradicionais foi avaliada por meio das seguintes estatísticas: coeficiente de determinação ajustado (Equação 3), erro padrão da estimativa (Equação 4), teste F da análise de variância, significância dos coeficientes da regressão (β_0, β_1 e β_2), pelo teste t ($\alpha = 0,05$) e a análise gráfica dos resíduos. A normalidade dos resíduos foi avaliada pelo teste de Shapiro-Wilk, com 95% de probabilidade. Os modelos mistos foram avaliados pelas mesmas estatísticas mencionadas acima, além dos critérios de informação de Akaike (Equação 5) e Bayesiano (Equação 6).

$$R^2_{aj} = 1 - \frac{SQ_{reg}}{SQ_{total}} \cdot \left(\frac{n-1}{n-p-1} \right) \quad (03)$$

$$S_{yx} \% = \sqrt{\frac{\sum_{i=1}^n (Ht_i - \hat{Ht}_i)^2}{n-p}} \cdot 100 \quad (04)$$

$$AIC = -2 \ln(mv) + 2p \quad (05)$$

$$BIC = -2 \ln(mv) + p \ln(n) \quad (06)$$

em que: R^2_{aj} = coeficiente de determinação ajustado; n = número de observações; p = número de parâmetros do modelo (não inclui o β_0); SQ_{reg} = soma de quadrados da regressão; SQ_{total} = soma de quadrados total; $S_{yx} \%$ = erro padrão da estimativa em porcentagem; Ht_i = altura total observada; $\hat{Ht}_i$ = altura total estimada; AIC = Critério de Informação de Akaike; mv = valor da máxima verossimilhança; BIC = Critério de Informação Bayesiano.

As validações dos modelos tradicionais e mistos foram realizadas com 421 árvores no primeiro cenário e 407 árvores no segundo cenário. O processo de validação foi realizado com base na análise gráfica das alturas estimadas em função das alturas observadas, por meio de regressão linear simples.

Os modelos foram ajustados utilizando-se o Software R (R CORE TEAM, 2020), sendo que para o ajuste dos modelos não lineares e mistos, foi utilizado o pacote nlme (PINHEIRO; BATES, 2020). Os modelos tradicionais foram ajustados pelo método dos mínimos quadrados ordinários, enquanto os modelos mistos foram ajustados pelo método da máxima verossimilhança restrita, ambos pressupondo $\varepsilon \sim N(0, \sigma^2)$. Embora tenham sido ajustados dez modelos hipsométricos, apenas os quatro melhores foram apresentados nos resultados de cada cenário. Após a indicação do modelo mais adequado para a aplicação da modelagem mista, apenas a combinação dos coeficientes de maior efeito aleatório foi apresentada.

3. RESULTADOS

Para a estimativa da altura total do primeiro cenário do banco dados, observou-se que os quatro melhores modelos

ajustados foram os lineares de Trorey (modelo 2), Henricksen (modelo 4), Stoffels (modelo 5) e o modelo não linear de Prodan (modelo 10) (Tabela 2).

Na Figura 2, observa-se dispersão dos dados, viés nas estimativas e pontos discrepantes. Pela Tabela 3, observa-se que o tratamento dos pontos discrepantes resultou em melhoria no ajuste. De modo geral, houve um aumento médio de 10,4% no R^2_{aj} e uma diminuição média de 4,5% para o $S_{yx}\%$, evidenciando o aumento da precisão das estimativas pela remoção dos pontos discrepantes.

Na Figura 3 verifica-se que com a remoção dos pontos discrepantes os dados apresentaram melhor disposição gráfica dos resíduos.

Tabela 2. Parâmetros estatísticos dos modelos tradicionais de relação hipsométrica ajustados ao cenário I do banco de dados.

Modelo	β_0	β_1	β_2	R^2_{aj}	$S_{yx}\%$	F
2	4,5857*	0,4808*	-0,0033*	0,552	27,855	1036,00*
4	-2,9040*	5,4042*		0,540	28,216	1976,19*
5	1,0182*	0,5022*		0,541	28,183	2083,11*
10	-2,0105*	1,1588*	0,0320*	0,555	27,740	1041,86*

* = significância a 5%.

* = significance at the 5% level.

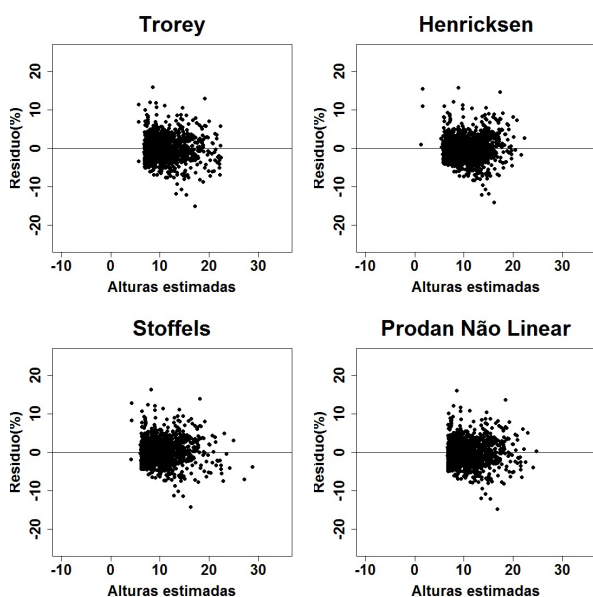


Figura 2. Distribuição dos resíduos dos modelos de relação hipsométrica ajustados ao cenário I do banco de dados.

Figure 2. Distribution of the residuals of the adjusted hypsometric relationship models for scenario I of the database.

Tabela 3. Parâmetros estatísticos dos modelos de relação hipsométrica ajustados ao cenário II do banco de dados.

Modelo	β_0	β_1	β_2	R^2_{aj}	$S_{yx}\%$	F
2	3,7240*	0,5809*	-0,0051*	0,654	23,404	1540,40*
4	-3,8760*	5,7993*		0,650	23,530	3030,20*
5	0,9151*	0,5473*		0,644	23,754	3259,03*
10	-0,7035*	0,9949*	0,0346*	0,657	23,310	1560,55*

* = significância a 5%.

* = significance at the 5% level.

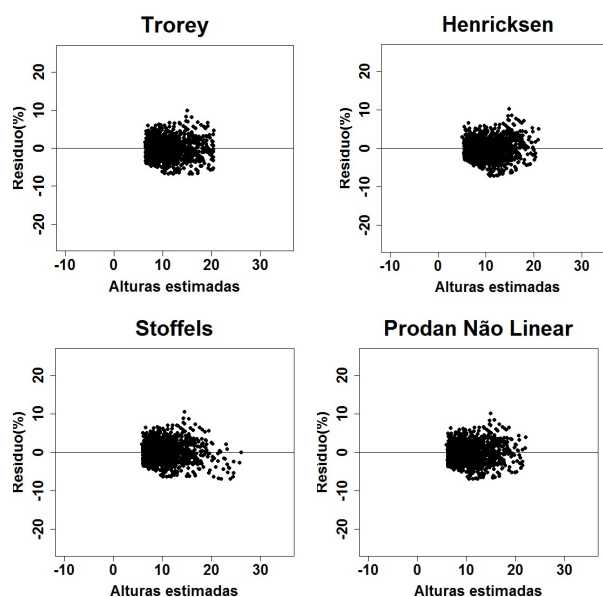


Figura 3. Distribuição dos resíduos dos modelos de relação hipsométrica ajustados ao cenário II do banco de dados.

Figure 3. Distribution of the residuals of the adjusted hypsometric relationship models for scenario II of the database.

Na Tabela 4 estão apresentadas as estatísticas de ajuste dos modelos ajustados com a inclusão das covariáveis espécie, família botânica e classe de DAP, para o primeiro cenário, considerando apenas a combinação entre os coeficientes que apresentou maior efeito aleatório (interação entre os coeficientes β_0, β_1 e β_2).

A inclusão da covariável "espécie" no modelo selecionado resultou em aumento de precisão, comprovado pelo incremento de 11,00% no R^2_{aj} e na redução de 3,61% no $S_{yx}\%$. Por outro lado, ao considerar a família botânica como um efeito aleatório, constatou-se um ganho menor de precisão. No entanto, a qualidade do ajuste foi reforçada pela diminuição dos valores dos critérios AIC e BIC em relação ao modelo selecionado, o que indica maior explicação da variância total. Não houve melhoria na precisão dos ajustes com a inclusão da covariável classe de DAP no modelo, tendo os critérios de AIC e BIC apresentado aumento, indicando perda de informação na estimativa.

Ao analisar a distribuição dos resíduos do modelo original e dos modelos modificados pela inclusão das covariáveis (Figura 4), observa-se leve heterocedasticidade. Este problema é evidenciado pela variação decrescente, em forma de funil, dos valores plotados. No entanto, a inclusão das covariáveis espécie e família botânica resultou em uma distribuição de resíduos mais centralizada e equilibrada, com redução da tendenciosidade. Por outro lado, a inclusão do fator aleatório de classe de DAP não contribuiu para explicar a variância total, nem melhorou a estimativa da altura nem a distribuição dos resíduos.

O modelo modificado, com a inclusão da covariável espécie na interação entre os três coeficientes, foi o mais preciso, apresentando a melhor e mais equilibrada distribuição residual, e foi considerado o mais adequado para o estabelecimento da relação hipsométrica no primeiro cenário do banco de dados.

Como constatado anteriormente, o tratamento dos pontos discrepantes resultou em melhoria na precisão dos ajustes e em ganhos percentuais subsequentes quando combinado à inclusão das covariáveis espécie e família

botânica (Tabela 5). A inclusão da covariável classe de DAP também não melhorou a precisão do ajuste no segundo cenário. Este fato é atestado pela semelhança entre os valores

de R^2_{aj} e $S_{yx}\%$ com o modelo original e com os maiores valores de AIC e BIC.

Tabela 4. Parâmetros estatísticos do modelo de relação hipsométrica selecionado e dos modelos ajustados com a inclusão das covariáveis espécie, família botânica e classe de DAP, no primeiro cenário do banco de dados, por meio da modelagem mista.

Table 4. Statistical parameters of the selected hypsometric relationship model and of the models fitted with the inclusion of the covariates species, botanical family, and DBH class, for the first scenario of the database, through mixed modeling.

Modelo	Covariável	β_0	β_1	β_2	R^2_{aj}	$S_{yx}\%$	AIC	BIC
Trorey	-	4,5857*	0,4808*	0,0033*	0,552	27,855	8311,47	8311,47
Trorey.b012	Espécie	4,8050*	0,4753*	-0,0040*	0,662	24,247	8196,26	8250,54
Trorey.b012	Família	4,4411*	0,5121*	-0,0046*	0,620	25,717	8251,94	8306,21
Trorey.b012	Classe DAP	5,3250*	0,4518*	-0,0032*	0,560	27,647	8342,65	8396,92

* = significância a 5%; * = significance at 5%.

Tabela 5. Parâmetros estatísticos do modelo de relação hipsométrica selecionado e dos modelos ajustados com a inclusão das covariáveis espécie, família botânica e classe de DAP, para o segundo cenário do banco de dados, por meio da modelagem mista.

Table 5. Statistical parameters of the selected hypsometric relationship model and of the models fitted with the inclusion of the covariates species, botanical family, and DBH class, for the second scenario of the database, by means of mixed modeling.

Modelo	Covariável	β_0	β_1	β_2	R^2_{aj}	$S_{yx}\%$	AIC	BIC
Trorey	-	3,7240*	0,5809*	0,0051*	0,654	23,404	7468,04	7489,62
Trorey.b012	Espécie	3,8225*	0,5803*	-0,0055*	0,732	20,627	7382,17	7436,11
Trorey.b012	Família	3,6013*	0,6014*	-0,0059*	0,706	21,583	7402,10	7456,04
Trorey.b012	Classe DAP	3,3040*	0,6130*	-0,0056*	0,661	23,178	7492,83	7546,77

* = significância a 5%; * = significance at 5%.

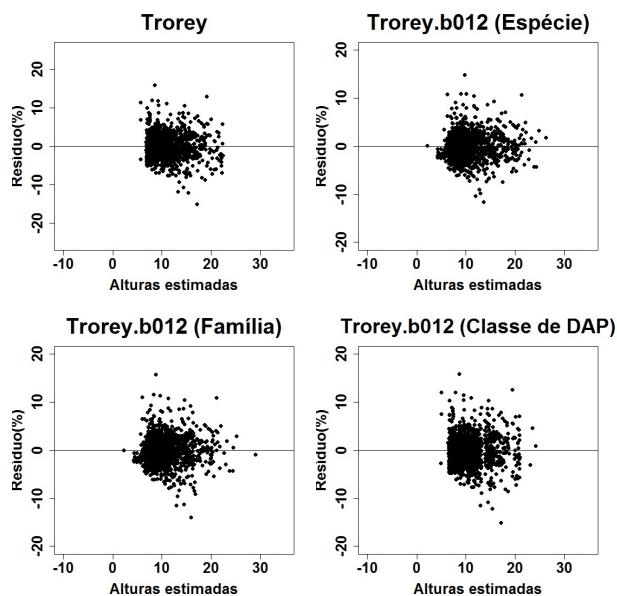


Figura 4. Distribuição dos resíduos do modelo de relação hipsométrica selecionado e dos modelos ajustados com a inclusão das covariáveis espécie, família botânica e classe de DAP no primeiro cenário do banco de dados, por meio da modelagem mista. Figure 4. Distribution of the residuals of the selected hypsometric relationship model and of the models fitted with the inclusion of the covariates species, botanical family, and DBH class, for the first scenario of the database, through mixed modeling.

Na Figura 5, observa-se melhoria visível na dispersão de resíduos, sobretudo quanto à homogeneidade da variância e à redução dos pontos dispersos, com exceção do modelo misto com a classe de DAP como fator aleatório, que não apresentou melhorias.

Logo, a combinação do procedimento de tratamento dos pontos discrepantes associado à modelagem mista, com a inclusão do fator aleatório “espécie” na interação entre os três coeficientes,

resultou em um aumento de 18,0% no R^2_{aj} e diminuição de 7,2% no $S_{yx}\%$, em relação ao modelo de Trorey selecionado para o primeiro cenário do banco de dados, o que representa a estratégia de melhor desempenho. No procedimento de validação (Figura 6), espera-se que os pontos dos dados se aproximem o máximo possível da linha de regressão ajustada, indicando maior explicação da variância pelo modelo de regressão.

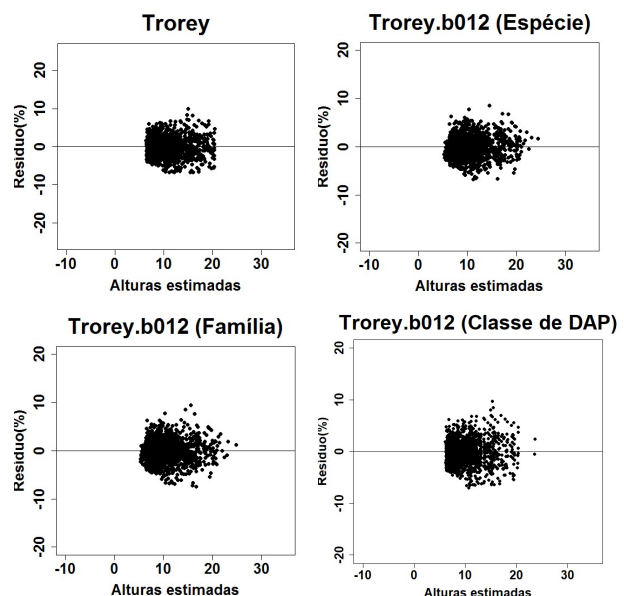


Figura 5. Distribuição dos resíduos do modelo de relação hipsométrica selecionado e dos modelos ajustados com a inclusão das covariáveis espécie, família botânica e classe de DAP, no segundo cenário do banco de dados, por meio da modelagem mista. Figure 5. Distribution of the residuals of the selected hypsometric relationship model and of the models fitted with the inclusion of the covariates species, botanical family, and DBH class, for the second scenario of the database, through mixed modeling.

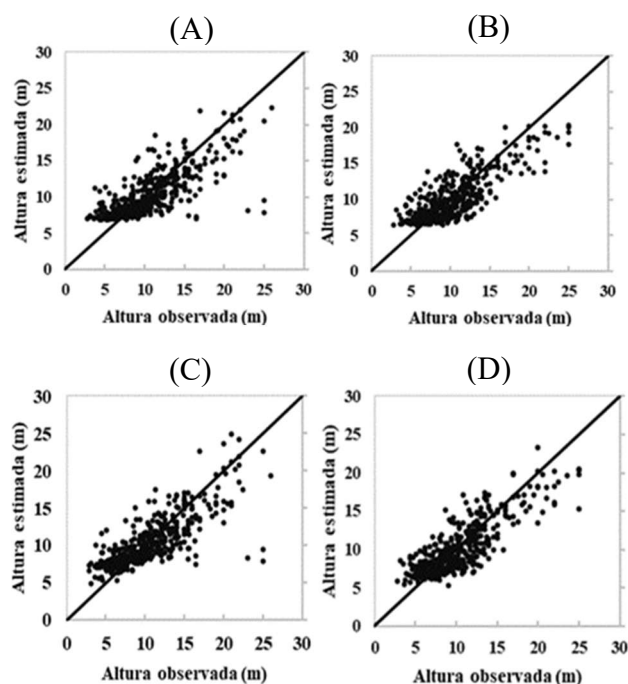


Figura 6. Gráfico de validação dos modelos de relação hipsométrica selecionados, sendo modelo tradicional para (A) o primeiro cenário e (B) segundo cenário e, modelagem mista com a covariável espécie (C) para o primeiro cenário (D) para o segundo cenário.

Figure 6. Validation plot of the selected hypsometric relationship models, traditional model for (A) the first scenario and (B) the second scenario; and mixed-effects modeling with species as a covariate for (C) the first scenario and (D) the second scenario.

6

A validação do modelo tradicional de Trorey, ajustado para o primeiro cenário (Figura 6A), mostrou-se eficaz, com pequenas e não viesadas diferenças entre os valores observados e os preditos pelo modelo. O ajuste do modelo de Trorey para o segundo cenário (Figura 6B) foi ainda melhor, como consequência da diminuição da variabilidade dos dados e aumento da precisão do ajuste. A modelagem mista, ao adicionar a covariável espécie melhorou significativamente a precisão do ajuste, aproximando os dados da linha de regressão ajustada (Figura 6C). Tal conclusão já era esperada, devido ao modelo misto incorporar a variabilidade do fator espécie na estimativa de seus parâmetros, haja vista que o fator espécie apresenta alta correlação com o desenvolvimento e crescimento do indivíduo. A validação mostrou-se ainda mais eficaz ao utilizar o modelo misto de Trorey ajustado para o segundo cenário (Figura 6D), resultando em alturas estimadas mais próximas das observadas em campo, como consequência da diminuição da dispersão dos dados, decorrente da remoção de outliers e pontos influentes, e da inclusão da variabilidade do fator espécie nos três parâmetros do modelo.

4. DISCUSSÃO

Os valores dos coeficientes de determinação ajustado ($R^2_{aj.}$) e dos erros ($S_{yx}\%$) foram considerados aceitáveis devido à fraca correlação entre o diâmetro e a altura em florestas inequiduais (ARAÚJO et al., 2012). Essa relação é influenciada pela distribuição diamétrica exponencial negativa em formato de "J" invertido, na qual há maior concentração de indivíduos em classes de diâmetro inferiores e alta variabilidade de alturas dentro de uma mesma classe de diâmetro (ASSMANN, 1970; MARANGON et al., 2008).

Este fato foi verificado, uma vez que mais de 80% dos indivíduos estão agrupados nas duas primeiras classes de diâmetro. Resultados semelhantes foram observados por Marangon et al. (2008) em um fragmento de Floresta Estacional Semidecidual em Viçosa, Minas Gerais, onde a maioria dos indivíduos arbóreos concentrou-se nas três primeiras classes de diâmetro e altura.

Na Figura 2, observa-se dispersão dos dados, viés nas estimativas e a presença de pontos discrepantes. Esses resultados são atribuídos à capacidade limitada dos modelos hipsométricos em explicar a relação biológica em florestas nativas. Isso ocorre devido à alta variabilidade das alturas dentro de uma mesma classe de diâmetro e à ampla diversidade de espécies, que apresentam diferentes parâmetros de crescimento e estágios sucessionais (CURTO et al., 2014).

Sendo assim, o modelo linear de Trorey (modelo 2) foi escolhido para o objetivo em questão devido as estatísticas de $R^2_{aj.}$ e $S_{yx}\%$ mais precisas e com a melhor dispersão dos resíduos. Embora a linearidade não represente bem o crescimento biológico, esta escolha é válida, uma vez, de acordo com Sena et al. (2015), a adição de um termo quadrático melhora o desempenho do modelo ao representar o efeito não linear, tornando-o mais adequado para representar o crescimento em altura.

Na Figura 3 verifica-se que, após a remoção dos pontos discrepantes, os resíduos apresentaram uma distribuição gráfica mais homogênea. Isto se explica pelo fato de que os pontos discrepantes atuam como alavancas, afetando a estimativa dos parâmetros do modelo. Estudos com dados de inventários florestais demonstram que a inclusão ou remoção de observações discrepantes pode alterar de forma relevante as estimativas de biomassa e sua variância, reforçando a necessidade de critérios consistentes para a identificação desses registros (KNOTT et al., 2023). Entretanto, em florestas nativas, valores extremos não devem ser automaticamente interpretados como erros, uma vez que podem refletir a elevada heterogeneidade estrutural desses ecossistemas, incluindo a ocorrência de indivíduos de grande porte ou de espécies com características distintas das predominantes na comunidade. Dessa forma, a remoção dos pontos discrepantes foi realizada com base em critérios estatísticos e na avaliação da consistência dos dados, de modo a minimizar a influência de observações potencialmente distorcidas, sem comprometer a representatividade ecológica da amostra.

O modelo linear de Trorey e o não linear de Prodan foram os que apresentaram as estatísticas de ajuste mais precisas. Porém, optou-se pelo modelo linear de Trorey devido à maior facilidade de ajuste e à menor necessidade de recursos computacionais, em comparação com a estrutura não linear.

Estudos sobre a inclusão de covariantes na modelagem hipsométrica mista estão avançando, mas ainda são limitados, especialmente em florestas inequiduais. Embora existam alguns trabalhos na literatura, a maioria está focada em povoamentos plantados, como os estudos de Mendonça et al. (2015) em plantios de *Eucalyptus* spp; e de Oliveira et al. (2015) em povoamento não manejado de *Eucalyptus* spp. em idade avançada; Sena et al. (2015), para duas espécies de *Pinus* tropicais, entre outros. No estudo de Mendonça et al. (2015), observou-se melhoria nos valores de $S_{yx}\%$ quando a variabilidade da idade e do material genético foram consideradas como covariáveis no ajuste do modelo. Isso

ocorre porque essas variáveis têm impacto direto na resposta das árvores quanto à altura total.

A ausência de ganho de precisão com a classe de DAP sugere que essa categorização não constitui uma fonte adicional de variabilidade suficientemente consistente para justificar sua inclusão como efeito aleatório. Isso pode se dever à concentração da maioria dos dados nas duas primeiras classes de DAP, o que dificulta a diferenciação das alturas estimadas devido à variabilidade introduzida por cada classe. Sendo assim, embora o DAP seja uma das principais variáveis associadas à altura das árvores, sua categorização em classes representa apenas parcialmente a variação estrutural dos indivíduos. Além disso, em florestas naturais, árvores pertencentes a uma mesma classe de DAP podem apresentar diferenças substanciais em altura associados à competição, à estrutura do povoamento, às condições ambientais e às características funcionais das espécies, que podem explicar variações adicionais na altura que não são capturadas pelo DAP isoladamente (CUI et al., 2022; YANG et al., 2022; FENG et al., 2024; ZHENG et al., 2026).

Quanto à validação, é importante ressaltar que a estratégia adotada considerou, conforme descrito por Ayer; Bhatta (2026), a estrutura dos dados, o tamanho da amostra e o objetivo da modelagem. Nesse sentido, o tamanho amostral disponível no presente estudo permitiu manter um conjunto de validação suficientemente representativo, sem comprometer o número de observações destinadas ao ajuste, o que torna a estratégia adotada adequada para a avaliação do desempenho preditivo dos modelos.

Em relação à melhoria substancial da precisão dos modelos com a inclusão da espécie como efeito aleatório, esta pode ser interpretada não apenas como uma consequência estatística da heterogeneidade dos dados, mas também como reflexo de diferenças ecológicas e funcionais entre as espécies.

A relação altura-diâmetro pode variar substancialmente entre espécies, e essa variação está associada à qualidade do sítio, à competição, à densidade do povoamento e às condições ambientais (CUI et al., 2022; JHA et al., 2023; FENG et al., 2024). Assim, árvores pertencentes à mesma classe de DAP podem ocupar posições distintas no dossel e, consequentemente, estar submetidas a diferentes níveis de competição por luz e espaço. A incorporação da espécie como efeito aleatório permite, portanto, acomodar parte dessa variação interespecífica não explicada pelas variáveis dendrométricas utilizadas no modelo, proporcionando maior capacidade preditiva dos modelos (CUI et al., 2022; JHA et al., 2023). Outros fatores, como índices de competição, são considerados variáveis importantes a serem testadas na estimativa de altura, embora nem sempre seja certo que a precisão dos modelos possa ser melhorada pela inclusão dos efeitos da competição em vizinhança próxima (CUI et al., 2022; YANG et al., 2022).

Estudos sobre a modelagem mista em florestas nativas aplicados à área de mensuração e manejo florestal, especialmente quanto às aplicações na modelagem hipsométrica, são recentes no Brasil (CYSNEIROS et al., 2020; SILVA et al., 2024; RIBEIRO et al., 2026), e apresentam implicações práticas, visto que previsões precisas da altura podem aumentar a precisão do inventário florestal (RIBEIRO et al., 2026), otimizando as estimativas de volume e biomassa.

Assim sendo, a partir dos resultados, nota-se a grande importância de estudar a aplicação de covariáveis na relação

hipsométrica, com a expectativa de obter maior precisão nas estimativas dos parâmetros, uma vez que a relação altura/diâmetro é influenciada por inúmeros fatores ambientais e biológicos.

5. CONCLUSÕES

O modelo de Trorey foi recomendado devido à melhor qualidade dos ajustes, o que fornece estimativas consistentes da altura das árvores no fragmento de Floresta Estacional Semidecidual.

A identificação e remoção de pontos discrepantes revelaram-se um procedimento crucial para aprimorar a precisão dos ajustes, possibilitando a obtenção de relações hipsométricas mais confiáveis e representativas da variação da altura.

A modelagem mista, ao incorporar a variabilidade dos fatores aleatórios de espécie e de família botânica na interação entre os três coeficientes, mostrou-se uma abordagem eficaz para aumentar a precisão das estimativas, sobretudo quando se aplicou o tratamento de pontos discrepantes em conjunto com a modelagem mista.

6. REFERÊNCIAS

- ANDRADE, V. C. L. de; LEITE, H. G. Hipsometric relationship modeling using data sampled in tree scaling and inventory plots. **Revista Árvore**, v. 35, n. 1, p. 157-164, 2011. <https://doi.org/10.1590/S0100-67622011000100019>
- ARAÚJO, E. J. G. de; PELISSARI, A. L.; DAVID, H. C.; SCOLFORO, J. R. S.; PÉLICO NETTO, S.; MORAIS, V. A. Relação hipsométrica para candeia (*Eremanthus erythropappus*) com diferentes espaçamentos de plantio em Minas Gerais, Brasil. **Pesquisa Florestal Brasileira**, v. 32, n. 71, p. 257-268, 2012. <https://doi.org/10.4336/2012.pfb.32.71.257>
- ASSMANN, E. **The principles of forest yield: studies in the organic production, structure, increment and yield of forest stands**. Braunschweig: Pergamon Press, 1970. 520p.
- AYER, S.; BHATTA, K. P. A practical workflow for applying validation strategies in tree height-diameter modelling. **MethodsX**, v. 16, e103982, 2026. <https://doi.org/10.1016/j.mex.2026.103982>
- BARROS, D. A. de; MACHADO, S. do A.; ACERBI JÚNIOR, F. W.; SCOLFORO, J. R. S. Comportamento de modelos hipsométricos tradicionais e genéricos para plantações de *Pinus oocarpa* em diferentes tratamentos. **Pesquisa Florestal Brasileira**, n. 45, p. 3-28, 2002.
- BARTOSZECK, A. C. de P. e S.; MACHADO, S. do A.; FIGUEIREDO FILHO, A.; OLIVEIRA, E. B. de. Modelagem da relação hipsométrica para bracingais da região metropolitana de Curitiba-PR. **Floresta**, v. 32, n. 2, p. 189-204, 2002.
- CARVALHO, L. M. T.; ACERBI JÚNIOR, F. W.; SCOLFORO, J. R. S.; CAVALCANTI, H. C. Monitoramento da Flora Nativa 2005-2007. In: **Inventário Florestal de Minas Gerais**. Lavras: Editora Lavras. v. 1, 2008. 357p.
- CONTO, T.; ARMSTON, J.; DUBAYAH, R. Characterizing the structural complexity of the Earth's forests with spaceborne lidar. **Nature**

- Communications**, v. 15, e8116, 2024. <https://doi.org/10.1038/s41467-024-52468-2>
- CUI, K.; WU, X.; ZHANG, C.; ZHAO, X.; von GADOW, K. Estimating height-diameter relations for structure groups in the natural forests of Northeastern China. **Forest Ecology and Management**, v. 519, e120298, 2022. <https://doi.org/10.1016/j.foreco.2022.120298>
- CURTO, R. de A.; LOUREIRO, G. H.; MÔRA, R.; MIRANDA, R. O. V.; PÉLLICO NETTO, S.; SILVA, G. F. Relações hipsométricas em Floresta Estacional Semidecidual. **Revista de Ciências Agrárias**, v. 57, n. 1, p. 57-66, 2014. <http://dx.doi.org/10.4322/rca.2013.066>
- CURTO, R. de A.; WINK, C.; ARAÚJO, E. J. G. de; KOHLER, S. V. Modelos de relação hipsométrica por classe de aproveitamento em floresta de transição Cerrado-Amazônia no Mato Grosso. **Enciclopédia Biosfera**, v. 15, n. 28, p. 687-700, 2018.
- CYSNEIROS, V. C.; PELISSARI, A. L.; GAUI, T. D.; FIORENTIN, L. D.; CARVALHO, D. C. de; SILVEIRA FILHO, T. B.; MACHADO, S. do A. Modeling of tree height-diameter relationships in the Atlantic Forest: effect of forest type on tree allometry. **Canadian Journal of Forest Research**, v. 50, n. 12, p. 1289-1298, 2020. <https://doi.org/10.1139/cjfr-2020-0060>
- FENG, Y.; CHAI, Y. QIN, Y. LI, G. HE, L., ZHANG, Z. WU, L., MENG, G. Plant functional traits and tree size inequality improved individual tree height prediction of mid-montane humid evergreen broad-leaved forests in southwest China. **Forest Ecology and Management**, v. 551, e121526, 2024. <https://doi.org/10.1016/j.foreco.2023.121526>
- FONSECA, N. C.; CUNHA, J. S. A.; ALBUQUERQUE, E. R. G. M. de; LINS-E-SILVA, A. C. B. Carbon stock in aboveground biomass and necromass in the Atlantic Forest: an analysis of data published between 2000 and 2021. **Anais da Academia Brasileira de Ciências**, v. 96, n. 1, e20220761, 2024. <https://doi.org/10.1590/0001-3765202420220761>
- GUIMARÃES, D. P.; REIS, R. J.; LANDAU, E. C. **Índices pluviométricos em Minas Gerais**. 30. ed. Sete Lagoas: Embrapa Milho e Sorgo, 2010. 88p.
- JHA, S.; YANG, S.-I.; BRANDEIS, T. J.; KUEGLER, O.; MARCANO-VEGA, H. Evaluation of regression methods and competition indices in characterizing height-diameter relationships for temperate and pantropical tree species. **Frontiers in Forests and Global Change**, v. 6, e1282297, 2023. <https://doi.org/10.3389/ffgc.2023.1282297>
- KNOTT, J. A.; LIKNES, G.C.; GIEBINK, C. L.; OH, S.; DOMKE, G. M.; MCROBERTS, R. E.; QUIRINO, V. F.; WALTERS, B. F. Effects of outliers on remote sensing-assisted forest biomass estimation: A case study from the United States national forest inventory. **Methods in Ecology and Evolution**, v. 14, p. 1587-1602, 2023. <https://doi.org/10.1111/2041-210X.14084>
- MARANGON, L. C.; FELICIANO, A. L. P.; BRANDÃO, C. F. L. S.; ALVES JÚNIOR, F. T. Relações florísticas, estrutura diamétrica e hipsométrica de um fragmento de floresta estacional semidecidual em Viçosa (MG). **Floresta**, v. 38, n. 4, p. 699-709, 2008. <https://doi.org/10.5380/rev.v38i4.13166>
- MARQUES, M. C. M.; GRELLE, C. E. V. (Org). The Atlantic Forest. 1 ed., Springer International Publishing, 2021. 517p.
- MARQUES NETO, R. O meio físico no Parque Estadual Nova Baden (Lambari, MG) e entorno: implicações inerentes ao relevo na gestão de unidades de conservação. **Estudos Geográficos**, v. 12, n. 1, p. 116-131, 2014.
- MARTINS, F. B.; GONZAGA, G.; SANTOS, D. F. dos; REBOITA, M. S. Classificação climática de Köppen e de Thornthwaite para Minas Gerais: cenário atual e projeções futuras. **Revista Brasileira de Climatologia**, p. 129-156, 2018. <https://doi.org/10.5380/abclima.v1i0.60896>
- MCCULLOCH, C. E.; SEARLE, S. R. **Generalized, linear, and mixed models**. New York: John Wiley & Sons, 2001. 432p.
- MENDONÇA, A. R.; CARVALHO, S. P. C.; CALEGARIO, N. Modelos hipsométricos generalizados mistos na predição da altura de *Eucalyptus* sp. **Cerne**, v. 21, n. 1, p. 107-115, 2015. <https://doi.org/10.1590/01047760201521011191>
- MIGLIAVACCA, M.; MUSAVI, T.; MAHECHA, M. D.; NELSON, J. A.; KNAUER, J.; BALDOCCHI, D. D.; PEREZ-PRIEGO, O.; CHRISTIANSEN, R.; PETERS, J.; ANDERSON, K.; et al. The three major axes of terrestrial ecosystem function. **Nature**, v. 598, p. 468-472, 2021. <https://doi.org/10.1038/s41586-021-03939-9>
- OLIVEIRA, G. M. V.; MELLO, J. M. de; ALTOÉ, T. F.; SCALON, J. D.; SCOLFORO, J. R. S.; PIRES, J. V. Equações hipsométricas para *Eucalyptus* spp. não manejado em idade avançada com técnicas de inclusão de covariantes. **Cerne**, v. 21, n. 3, p. 483-492, 2015. <https://doi.org/10.1590/01047760201521031740>
- PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer - Verlag, 2000. 528p.
- POORTER, L.; BONGERS, F.; AIDE, T. M.; ZAMBRANO, A. M. A.; BALVANERA, P.; BECKNELL, J. M.; BOUKILI, V.; BRANCALION, P. H. S.; BROADBENT, E. N.; CHAZDON, R. L.; et al. Biomass resilience of Neotropical secondary forests. **Nature**, v. 530, e7589, p. 211-214, 2016. <https://doi.org/10.1038/nature16512>
- RIBEIRO, R. B. da S.; REIS, L. P.; WOYCIKIEVICZ, A. P. F.; MELLO, M. N. de; OLIVEIRA, A. H. M.; DIAS, C. T. dos S.; MARTORANO, L. G. Modeling commercial height in Amazonian Forests: accuracy of mixed-effects regression versus random forest. **Forests**, v. 17, n. 1, e30, 2026. <https://doi.org/10.3390/f17010030>
- ROUSSEEUW, P. J.; LEROY, A. M. **Robust regression and outlier detection**. New York: John Wiley e Sons, 1987. 352p.
- SCOLFORO, J. R. S.; MELLO, J. M. D. **Inventário Florestal**. Lavras, MG: Ed. UFLA/FAEPE, 2006. 561p.
- SENA, A. L. M.; SILVA NETO, A. J. da; OLIVEIRA, G. M. V.; CALEGARIO, N. Modelos lineares e não lineares com uso de covariantes para relação hipsométrica de duas espécies de Pinus tropicais. **Ciência Florestal**, v. 25, n. 4, p. 969-980, 2015. <https://doi.org/10.5902/1980509820651>
- SILVA, G. F. da; XAVIER, A. C.; RODRIGUES, F. L.; PETERNELLI, L. A. Análise da influência de diferentes tamanhos e composições de amostras no ajuste de uma relação hipsométrica para *Eucalyptus grandis*. **Revista**

- Árvore**, v. 31, n. 4, p. 685-694, 2007.
<https://doi.org/10.1590/S0100-67622007000400013>
- SILVA, A. V. dos S.; SOUZA, R. G. T. de; SOUZA, R. L. F. de; SANTOS, R. M.; LIMA, R. B. de; ABREU, J. C. de. Use of machine learning, fixed and mixed models for volume estimation in floodplain forest in the Amazon estuary. **Floresta**, v. 54, e83115, 2024.
<https://doi.org/10.5380/rf.v54.83115>
- YANG, S.-I.; BRANDEIS, T. J.; HELMER, E. H.; OATHAM, M. P.; HEARTSILL-SCALLEY, T.; MARCANO-VEGA, H. Characterizing height-diameter relationships for Caribbean trees using mixed-effects random forest algorithm. **Forest Ecology and Management**, v. 524, e120507, 2022.
<https://doi.org/10.1016/j.foreco.2022.120507>
- ZHENG, Y.; YUE, Y.; GAO, R.; SHARMA, R. P. Developing mixed-effects height-diameter model using stand and environmental factors for mixed forests in northern China. **iScience**, v. 29, n. 1, e114446, 2026.
<https://doi.org/10.1016/j.isci.2025.114446>

Contribuições dos autores: B.F.M.: metodologia, investigação ou coleta de dados, análise estatística; redação (original), redação (revisão e edição); N.G.B.: validação, redação (revisão e edição); J.M.L.A.: validação, redação (revisão e edição); E.V.S.: metodologia, investigação ou coleta de dados, redação (revisão e edição); R.A.C.: metodologia, investigação ou coleta de dados, redação (revisão e edição); E.J.G.A.: conceituação, metodologia, investigação ou coleta de dados, administração ou supervisão, redação (original), redação (revisão e edição). Todos os autores leram e concordaram com a versão publicada do manuscrito.

Disponibilidade de dados: Os dados desta pesquisa poderão ser obtidos por e-mail, mediante solicitação ao autor correspondente.

Conflito de interesses: Os autores declaram não haver conflitos de interesses.



Copyright: © 2026 by the authors. This article is an Open-Access article distributed under the terms and conditions of the Creative Commons **Attribution-NonCommercial (CC BY-NC)** license (<https://creativecommons.org/licenses/by/4.0/>).