

**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

Mensuração da biomassa e construção de modelos para construção de equações de biomassa

Edgar de Souza Vismara

**Dissertação apresentada para obtenção do título de
Mestre em Recursos Florestais, com opção em
Silvicultura e Manejo Florestal**

**Piracicaba
2009**

Edgar de Souza Vismara
Engenheiro Florestal

Mensuração da biomassa e construção de modelos para construção de equações de biomassa

Orientador:
Prof. Dr. JOÃO LUIS FERREIRA BATISTA

Dissertação apresentada para obtenção do título de Mestre em Recursos Florestais, com opção em Silvicultura e Manejo florestal

Piracicaba
2009

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Vismara, Edgar de Souza

Mensuração da biomassa e construção de modelos para construção de equações de biomassa / Edgar de Souza Vismara. - - Piracicaba, 2009.
102 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura Luiz de Queiroz, 2009.
Bibliografia.

1. Biomassa 2. Florestas tropicais 3. Gravimetria 4. Seleção de modelos 5. Volumetria
I. Título

CDD 634.983
V832m

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

AGRADECIMENTOS

Ao professor João Luis Ferreira Batista, pelos exemplos, pela orientação, pela presença de espírito, pela companhia nas análises, nos desânimos referentes à dissertação e à vida.

Aos meus pais, Doris e Roberto, por sempre estarem presentes, mesmo nas ausências, e por todo o apoio. À minha avó, Saturnina e meus irmãos Fábio e Iara por me amarem apesar de tudo, e principalmente a minha esposa Lilian, que desde 1998 se faz presente, nas alegrias e tristezas, como ninguém jamais esteve.

À CAPES, pela bolsa de estudos.

À todos do CMQ, pela melhor das acolhidas, pelas conversas, cafés, risadas, pela convivência sempre agradável em todos os momentos, mesmo os nem tão agradáveis, com todos os professores e os colegas que ao longo desses três anos chegaram, partiram e algumas vezes voltaram. E em especial, meus “irmãos de orientador”, Melissa, Mauricio e Jaime, pelo companheirismo e apoio fundamental e constante.

Ao Jeferson, a quem não existem palavras suficientes para agradecer, por toda a ajuda, toda a calma, amizade e paciência.

Aos professores e funcionários do LCF, pela ajuda, pela perene boa vontade e simpatia na convivência quase diária, e pelas conversas nos cafés das 9, que deixaram saudades.

Aos amigos e colegas do curso, pela convivência e aprendizado. Aos funcionários e amigos da Vale: Jonacir, Ectore, Alan, Genil, Brás, Valdecir e Rodolfo pela imensa ajuda, com especial destaque para o Adeilton pela ajuda e fornecimento de “Seleta” nos dias mais difíceis de trabalho.

“All models are wrong, some are usefull ”
(George Box, 1979)

SUMÁRIO

RESUMO	7
ABSTRACT	8
1 INTRODUÇÃO	9
Referências	10
2 SELEÇÃO DE MODELOS EMPÍRICOS ATRAVÉS DO CRITÉRIO DE INFORMAÇÃO DE AKAIKE	12
Resumo	12
Abstract	12
2.1 Introdução	12
2.2 Abordagem tradicional	13
2.3 Abordagem das discrepâncias	17
2.3.1 Conceitos de discrepância	17
2.3.2 Distância de Kulbach-Leibler e Entropia	19
2.3.3 Critério de informação de Akaike	21
2.4 Considerações Finais	23
Referências	24
3 AVALIAÇÃO DAS TÉCNICAS DE DETERMINAÇÃO DA BIOMASSA LENHOSA DE ÁRVORES	27
Resumo	27
Abstract	27
3.1 Introdução	28
3.2 Material e Métodos	30
3.2.1 Área de estudo	30
3.2.2 Escolha das espécies	30
3.2.3 Determinação da biomassa	30
3.2.4 Os Discos de Madeira Amostrados do Tronco	31
3.2.5 Determinação do Teor de Umidade do Tronco	32
3.2.6 Determinação da Densidade da Madeira do Tronco	33
3.2.7 Modelos para Predição da Biomassa do Tronco	34
3.2.8 Ajuste e Avaliação dos Modelos de Predição	35
3.3 Resultados e Discussão	37

3.3.1 Determinação do Teor de Umidade	37
3.3.2 Determinação da Densidade da Madeira do Tronco	40
3.3.3 Determinação da Biomassa do Tronco	53
3.4 Conclusões	65
Referências	66
4 AJUSTE E SELEÇÃO DE MODELOS PREDITIVOS DA BIOMASSA AÉREA	68
Resumo	68
Abstract	68
4.1 Introdução	69
4.2 Material e Métodos	70
4.2.1 Área de estudo	70
4.2.2 Escolha das espécies	70
4.2.3 Determinação da biomassa	71
4.2.4 Determinação das variáveis preditoras	73
4.2.5 Ajuste e Seleção de Modelos	73
4.2.6 Análise dos modelos selecionados	76
4.3 Resultados e Discussão	77
4.3.1 Análise descritiva das variáveis usadas no ajuste dos modelos	77
4.3.2 Seleção dos modelos	78
4.3.3 Análise dos modelos selecionados	81
4.4 Conclusões	92
Referências	95
ANEXOS	97

RESUMO

Mensuração da Biomassa e Seleção de Modelos para a Construção de Equações de Biomassa

O interesse pela quantificação da biomassa florestal vem crescendo muito nos últimos anos, sendo este crescimento relacionado diretamente ao potencial que as florestas tem em acumular carbono atmosférico na sua biomassa. A biomassa florestal pode ser acessada diretamente, por meio de inventário, ou através de modelos empíricos de predição. A construção de modelos de predição de biomassa envolve a mensuração das variáveis e o ajuste e seleção de modelos estatísticos. A partir de uma amostra destrutiva de 200 indivíduos de dez essências florestais distintas advindos da região de Linhares, ES., foram construídos modelos de predição empíricos de biomassa aérea visando futuro uso em projetos de reflorestamento. O processo de construção dos modelos consistiu de uma análise das técnicas de obtenção dos dados e de ajuste dos modelos, bem como de uma análise dos processos de seleção destes a partir do critério de Informação de Akaike (AIC). No processo de obtenção dos dados foram testadas a técnica volumétrica e a técnica gravimétrica, a partir da coleta de cinco discos de madeira por árvore, em posições distintas no lenho. Na técnica gravimétrica, estudou-se diferentes técnicas de composição do teor de umidade dos discos para determinação da biomassa, concluindo-se como a melhor a que utiliza a média aritmética dos discos da base, meio e topo. Na técnica volumétrica, estudou-se diferentes técnicas de composição da densidade do tronco com base nas densidades básicas dos discos, concluindo-se que em termos de densidade do tronco, a média aritmética das densidades básicas dos cinco discos se mostrou como melhor técnica. Entretanto, quando se multiplica a densidade do tronco pelo volume deste para obtenção da biomassa, a utilização da densidade básica do disco do meio se mostrou superior a todas as técnicas. A utilização de uma densidade básica média da espécie para determinação da biomassa, via técnica volumétrica, se apresentou como uma abordagem inferior a qualquer técnica que utiliza informação da densidade do tronco das árvores individualmente. Por fim, sete modelos de predição de biomassa aérea de árvores considerando seus diferentes compartimentos foram ajustados, a partir das funções de Spurr e Schumacher-Hall, com e sem a inclusão da altura como variável preditora. Destes modelos, quatro eram gaussianos e três eram lognormais. Estes mesmos sete modelos foram ajustados incluindo a medida de penetração como variável preditora, totalizando quatorze modelos testados. O modelo de Schumacher-Hall se mostrou, de maneira geral, superior ao modelo de Spurr. A altura só se mostrou efetiva na explicação da biomassa das árvores quando em conjunto com a medida de penetração. Os modelos selecionados foram do grupo que incluíram a medida de penetração no lenho como variável preditora e , exceto o modelo de predição da biomassa de folhas, todos se mostraram adequados para aplicação na predição da biomassa aérea em áreas de reflorestamento.

Palavras-chave: Seleção de Modelos; Critério de Informação de Akaike; AIC; Técnica gravimétrica; Técnica volumétrica; Densidade básica; Teor de umidade; Biomassa aérea; Modelos de predição; Floresta tropical atlântica

ABSTRACT

Biomass Measurement and Models Selection for Biomass Equations

Forest biomass measurement implies a destructive procedure, thus forest inventories and biomass surveys apply indirect procedure for the determination of biomass of the different components of the forest (wood, branches, leaves, roots, etc.). The usual approach consists in taking a destructive sample for the measurement of trees' attributes and an empirical relationship is established between the biomass and other attributes that can be directly measured on standing trees, e.g., stem diameter and tree height. The biomass determination of felled trees can be achieved by two techniques: the gravimetric technique, that weights the components in the field and take a sample for the determination of water content in the laboratory; and the volumetric technique, that determines the volume of the component in the field and take a sample for the determination of the wood specific gravity (wood basic density) in the laboratory. The gravimetric technique applies to all components of the trees, while the volumetric technique is usually restricted to the stem and large branches. In this study, these two techniques are studied in a sample of 200 trees of 10 different species from the region of Linhares, ES. In each tree, 5 cross-sections of the stem were taken to investigate the best procedure for the determination of water content in gravimetric technique and for determination of the wood specific gravity in the volumetric technique. Also, Akaike Information Criterion (AIC) was used to compare different statistical models for the prediction of tree biomass. For the stem water content determination, the best procedure was the arithmetic mean of the water content from the cross-sections in the base, middle and top of the stem. In the determination of wood specific gravity, the best procedure was the arithmetic mean of all five cross-sections discs of the stem, however, for the determination of the biomass, i.e., the product of stem volume and wood specific gravity, the best procedure was the use of the middle stem cross-section disc wood specific gravity. The use of an average wood specific gravity by species showed worse results than any procedure that used information of wood specific gravity at individual tree level. Seven models, as variations of Spurr and Schumacher-Hall volume equation models, were tested for the different tree components: wood (stem and large branches), little branches, leaves and total biomass. In general, Schumacher-Hall models were better than Spurr based models, and models that included only diameter (DBH) information performed better than models with diameter and height measurements. When a measure of penetration in the wood, as a surrogate of wood density, was added to the models, the models with the three variables: diameter, height and penetration, became the best models.

Keywords: Model Selection; Akaike information criterion; AIC; Gravimetric technique; Volumetric technique; Basic density; Stem water content; Aboveground biomass; Prediction models; Atlantic rain forest

1 INTRODUÇÃO

Até recentemente, os pesquisadores no meio florestal, usavam como medida para quantificar o material lenhoso num sistema florestal, somente o volume de madeira por unidade de área. Atualmente, tem-se utilizado, também, a produção de biomassa, já que o conhecimento desta nos diversos ecossistemas tem assumido papel importante para todo um vasto conjunto de aspectos como a comercialização de produtos, estudos da produtividade do sistema florestal, estudos de fluxos de energia e de nutrientes e estudos da contribuição dos ecossistemas para o ciclo global do carbono (SOARES; LEITE; GÖRGENS, 2005).

A queima de combustíveis fósseis associado às mudanças de uso da terra tem lançado na atmosfera grandes quantidades de gás carbônico (CO_2), sendo esta a principal causa do chamado efeito estufa. No entanto, somente a partir do estabelecimento de mecanismos de flexibilização no protocolo de Quioto em 1997 na Convenção das Partes-3 (COP-3) que o seqüestro de carbono por florestas foi considerado como alternativa para compensação das emissões por países industrializados (YU, 2002).

Neste contexto, o Brasil tem grande potencial para comercialização de créditos de carbono com países industrializados, já que possui uma extensa área plantável e condições edafoclimáticas ideais para o desenvolvimento de florestas (ROCHA, 2002).

Gutiérrez e Lopera (2001) afirmam que para realizar projetos florestais de captura de CO_2 é necessário desenvolver metodologias que permitam medir e quantificar de maneira confiável e a baixo custo o carbono armazenado. Como a quantidade de carbono a ser seqüestrada depende da variável biomassa, esta, segundo Sanquetta (2002), precisa ser determinada de forma fidedigna, caso contrário não haverá consistência na quantificação do carbono fixado nos ecossistemas florestais, prejudicando as negociações nos projetos de fixação de CO_2 em florestas.

A forma de acessar essa biomassa de maneira confiável e com relativo baixo custo se dá através da construção de modelos empíricos que permitem determinar a biomassa da floresta em pé a partir de seus atributos dendrométricos, com base numa amostra destrutiva composta apenas por um pequeno número de indivíduos (MAESTRI et al., 2004).

O presente trabalho tem como objetivo, estudar o processo de construção de modelos de predição empírico de biomassa aérea para dez essências florestais distintas na região de Linhares, ES, a partir de uma amostra de 200 indivíduos. As técnicas de obtenção dos dados e de construção dos modelos, bem como os processos de seleção destes serão analisados utilizando um critério denominado critério de Informação de Akaike (AIC).

A apresentação do trabalho esta organizada em três capítulos: o primeiro consiste numa revisão crítica sobre os procedimentos de seleção de modelos usualmente utilizados na modelagem de

predição de biomassa arbórea. Neste mesmo capítulo as bases teóricas do Critério de Informação de Akaike, que é pouco difundida no meio florestal, é apresentada como alternativa aos métodos de seleção tradicionais.

O segundo capítulo, apresenta o processo de determinação da biomassa lenhosa de árvores individuais, através de duas técnicas de mensuração da biomassa aérea de árvores abatidas: a técnica volumétrica e da técnica gravimétrica. Estas técnicas são estudadas sob o ponto de vista de modelos de predição, sendo sua análise feita a partir dos conceitos de discrepância entre modelos apresentados no capítulo 1.

No terceiro e último capítulo são ajustados vários modelos de predição da biomassa total e dos diferentes compartimentos de árvores individuais, construídos a partir das variáveis predictoras tradicionais e com a inclusão de uma medida penetração no lenho. Para seleção dos melhores modelos de predição da biomassa total e dos diferentes compartimentos, foi usada a abordagem de seleção de modelos empíricos proposta no capítulo 1, obtendo-se assim ao final do trabalho os modelos mais adequados aos dados coletados.

Referências

GUTIERREZ, V.H.; LOPERA, G.J. Flujo de carbono y respuesta a diferentes estrategias de manejo en plantaciones tropicales de *Pinus patula*. In: SIMPÓSIO INTERNACIONAL: MEDICION E MONITOREO DE CARBONO EM ECOSISTEMAS FORESTALES. 2001. Valdivia-Chile. **Anais...** Valdivia-Chile:, 2001. p. 22-27.

MAESTRI, R.; SANQUETTA, C.R.; MACHADO, S.A.; SCOLFORO, J.R.S.; CORTE, A.P.D. Viabilidade de um projeto florestal de *Eucalyptus grandis* considerando o seqüestro de carbono. **Floresta**, Curitiba, v. 34, n. 3, p. 347-360, dez. 2004.

ROCHA, M.T. O aquecimento global e os instrumentos de mercado para a solução do problema. In: SANQUETTA, C.R.; WATZLAWICK, L.F.; BALBINOT, R.; ZILLIOTO, M.A.B.; GOMES, F.S. (Ed.). **As florestas e o carbono**: [s.n.], 2002, cap. 1, p. 1-34.

SANQUETTA, C.R. Métodos de determinação de biomassa florestal. In: In: SANQUETTA, C.R.; WATZLAWICK, L.F.; BALBINOT, R.; ZILLIOTO, M.A.B.; GOMES, F.S. (Ed.). **As florestas e o carbono**: [s.n.], 2002, cap. 6, p. 119-140.

SOARES, C.P.B.; LEITE, H.G.; GÖRGENS, E.B. Equações para estimar o estoque de carbono no fuste de árvores individuais em plantios comerciais de eucalipto. **Revista Árvore**. Viçosa, v. 29, n. 5, p. 711-718, ago. 2005.

YU, C.M. Caracterização e tipologia dos projetos de seqüestro de carbono no Brasil. In: SANQUETTA, C.R.; WATZLAWICK, L.F.; BALBINOT, R.; ZILLIOTO, M.A.B.; GOMES, F.S. (Ed.). **As florestas e o carbono**: [s.n.], 2002, cap. 3, p. 59-87.

2 SELEÇÃO DE MODELOS EMPÍRICOS ATRAVÉS DO CRITÉRIO DE INFORMAÇÃO DE AKAIKE

Resumo

A questão dos métodos de seleção de modelos usados na área florestal é discutida, no que diz respeito a suas limitações, e o AIC é apresentado como abordagem alternativa para seleção de modelos empíricos no meio florestal. Na década de 70 Akaike propôs seu critério de seleção de modelos baseado no princípio de máxima entropia entre duas distribuições e no de máxima verossimilhança na estimativa de parâmetros. Este capítulo apresenta sucintamente as bases teóricas desse critério, bem como discute sua importância no contexto da modelagem no meio florestal.

Palavras-chave: Seleção de Modelos; Critério de Informação de Akaike; AIC

Abstract

The question of model selection methods used in forestry is discussed, with regard to its limitations, and the AIC is presented as alternative approach for the selection of empirical models in forest environment. In the 70's Akaike proposed the criterion of model selection based on the principle of maximum entropy between two distributions and of the maximum likelihood estimation of parameters. This paper presents briefly the theoretical basis of this criterion and discusses its importance in the context of forestry modeling.

Keywords: Model Selection; Akaike Information Criterion; AIC

2.1 Introdução

O processo de análise estatística consiste basicamente em encontrar um modelo apropriado, estimar os parâmetros e, por fim, determinar a ordem ou tamanho deste modelo (BOSDOGAN, 1987).

Nesse contexto, assume-se que exista um único modelo correto (ou até mesmo verdadeiro) ou pelo menos um melhor modelo que seria suficiente como base para inferir a partir dos dados (BURHAM; ANDERSON, 2004). No entanto, de acordo com Burham e Anderson (2002), a modelagem estatística de dados é um campo da estatística que busca ajustar um modelo aos dados sem conhecimento de como o modelo verdadeiro é ou deve ser.

De fato, o que ocorre normalmente nos diversos campos de aplicação da estatística, incluindo o meio florestal, é que uma vez selecionado o modelo o problema da inferência passa a ser de estimação e determinação da ordem do modelo, definida pelo número de variáveis preditoras. Desta

maneira, no processo de inferência, a incerteza na escolha do modelo não é considerada, tornando-se apenas mero componente da variância (BURHAM; ANDERSON, 2004).

Em trabalhos recentes da literatura (SAKAMOTO; ISHIGURO; KITAGAWA, 1986; BOSDOGAN, 1987; CAMERON; WINDMEIJER, 1996; BUCKLAND; BURHAM; AUGUSTIN, 1997; BOSDOGAN, 2000; BURHAM, ANDERSON, 2002, 2004) o problema da seleção de modelos vem sendo tratado de forma distinta. Tem-se buscado um critério que avalie o melhor modelo de aproximação, entre uma série de modelos candidatos com diferentes relações funcionais e com diferentes número de parâmetros, para descrever os dados.

Esse critérios, além de incluir a incerteza na seleção dos modelos, que permite a comparação de modelos não hierárquicos, ainda levam em consideração a complexidade do modelo de acordo com o princípio da parcimônia, ou seja, dentre uma série de modelos concorrentes de mesmo desempenho, os modelos mais complexos ou com maior número de parâmetros devem ser penalizados.

Essa abordagem iniciada por Akaike (1974, 1981) estabelece uma simples relação entre a distância de Kulbach-Leibler e a função de máxima log-verossimilhança de Fisher levando a uma simples e efetiva metodologia para seleção de modelos parcimoniosos para análise de dados empíricos, denominado de critério de informação de Akaike (AIC).

Este critério estima a distância relativa de Kulbach-Leibler entre dois modelos, possuindo a vantagem de permitir a comparação de modelos não-hierárquicos, considerando-os apenas concorrentes (SAKAMOTO; ISHIGURO; KITAGAWA, 1986), sendo portanto, uma ferramenta valiosa na seleção de modelos empíricos de predição no meio florestal.

Desta maneira, este capítulo se propõem discutir as limitações do uso dos critérios de seleção de modelos empíricos comumente usados no meio florestal e apresentar como o AIC poderá vir a suprir essas limitações, tornado-se ferramenta importante no processo de seleção e avaliação desses modelos.

2.2 Abordagem tradicional

Em trabalhos realizados no meio florestal (SOARES; OLIVEIRA, 2002; MAESTRI et. al., 2004; SOARES; LEITE; GÖRGENS, 2005) modelos empíricos de predição de volume ou biomassa arbórea são obtidos através de métodos estatísticos de regressão linear e não linear. Nesses trabalhos, a verificação da qualidade do ajuste e a conseqüente escolha do melhor modelo ocorrem através da abordagem clássica dos testes F do modelo e parcial, bem como através do coeficiente de determinação (R^2).

Nesses trabalhos, o teste F do modelo e o coeficiente de determinação são tratados como critérios de seleção absolutos, já que são usados para verificar se o modelo é ou não adequado para aplicação num determinado conjunto de dados. Trata-se de um equívoco pois, na verdade, esses dois critérios comparam a qualidade do ajuste de um modelo em relação a um outro modelo

mais simples que é a média amostral .

Esse equivoco fica claro a partir da observação da natureza desses critérios. O teste F do modelo, segundo (CHATERJEE; PRICE, 1991) é construído através da investigação das seguintes hipóteses:

H_0 : Todos os coeficientes da regressão são iguais à zero.

vs.

H_a : Pelo menos um coeficiente da regressão difere de zero.

Se a hipótese nula não for rejeitada, conclui-se que não há relação entre a variável resposta e as p variáveis preditoras, ou seja, o modelo testado não difere da média amostral. No caso de se rejeitar a hipótese nula conclui-se que o modelo de regressão explica melhor a variabilidade dos dados que a média amostral.

No entanto, se forem considerados dois modelos concorrentes, em que a hipótese nula tenha sido rejeitada para ambos, o teste F do modelo é de pouca utilidade como critério de decisão de qual modelo é mais adequado (DRAPER; SMITH, 1998).

O coeficiente de determinação (R^2), por sua vez, quantifica a proporção da variabilidade da variável resposta que é explicada por um modelo de aproximação qualquer (RAO, 1973). Este é definido por

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}, \quad (2.1)$$

onde $\sum(y_i - \hat{y}_i)^2$ é a soma dos desvios em relação aos valores ajustados e $\sum(y_i - \bar{y}_i)^2$ é a soma dos desvios em relação à média.

O R^2 pode assumir valores no intervalo $[0,1]$, sendo que valores próximos de 1 (um) denotam uma boa relação entre a variável resposta e as p variáveis preditoras, indicando um bom ajuste. Por outro lado, valores próximos ou iguais a 0 (zero) denotam pouca ou nenhuma relação entre a variável resposta e as p variáveis preditoras, indicando que o modelo não é superior a média amostral (DRAPER; SMITH, 1981).

Apesar do caráter relativo do coeficiente de determinação, parece lógico usa-lo como critério de seleção de modelos, já que, ao contrário do teste F do modelo, fornece uma “medida” do quanto o modelo é superior à média amostral. Sendo assim, se forem considerados dois modelos concorrentes cujos teste F do modelo rejeitassem a hipótese nula optaria-se pelo R^2 de maior valor, ou com melhor qualidade de ajuste.

No entanto, este caráter de medida de discrepância em relação à média limita muito a utilização do R^2 na seleção de modelos quando este é aplicado isoladamente (CHATERJEE; PRICE, 1991). Para corroborar com essa afirmação os autores citam o trabalho de Anscombe (1973), onde quatro diferentes conjuntos de dados, com diferentes padrões, foram gerados e ajustados a um modelo de regressão linear simples.

Os coeficientes do modelo, os coeficientes de determinação, bem como os resultados dos testes de hipóteses foram os mesmos para os exemplos (a), (b), (c) e (d) da Figura 2.1, demonstrando o problema do uso dessas estatísticas como evidência de qualidade de ajuste quando a relação entre as variáveis escapa à linearidade.

Notadamente, (a) é o único conjunto de dados em que a relação linear se aplica. Por ser o mesmo para todos os casos, se o R^2 fosse aplicado como critério para seleção entre os modelos (a), (b), (c) e (d), este não conseguiria captar a discrepância entre o modelo ajustado e o verdadeiro padrão sugerido pelo conjunto de dados.

Desta forma, a aplicação do R^2 como critério de seleção na análise de regressão não linear também é complicada. Bates e Watts (1988) afirmam que as propriedades dos modelos lineares, não

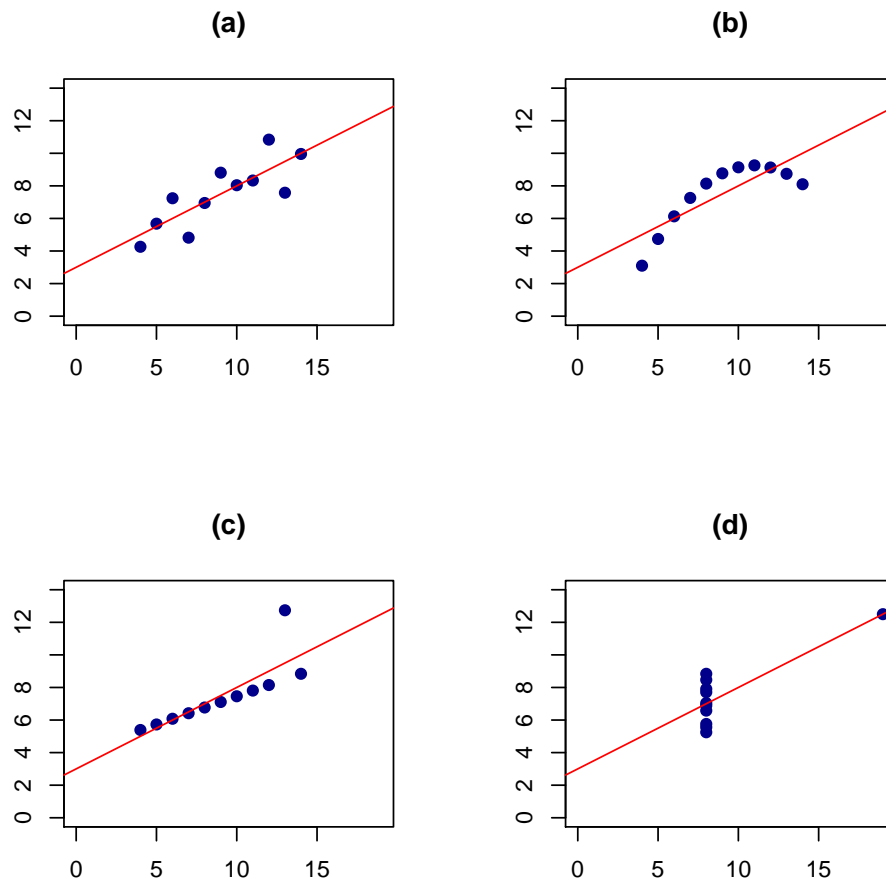


Figura 2.1 – Ajuste de quatro conjunto de dados (a, b, c e d) com padrões distintos, via regressão linear, onde foram obtidos o mesmo modelo ($y = 3 + 0,5x + e$) e o mesmo $R^2(0,666)$ (retirado de Ascombe, 1977)

são válidas para os modelos não lineares. Por exemplo, a soma dos resíduos não necessariamente é

igual à zero e a soma de quadrados do erro mais a soma de quadrados da regressão, não necessariamente é igual à soma de quadrados total. Consequentemente, o coeficiente de determinação, pode assumir valores diferentes do intervalo $[0,1]$, não sendo uma estatística descritiva importante para os modelos não lineares (CAMERON; WINDMEIJER, 1996).

Outra limitação ao uso do R^2 na seleção de modelos se refere à análise de regressão múltipla. Neste tipo de análise nem todas as variáveis preditoras são necessariamente efetivas para predição da variável resposta, mas a inclusão de novas variáveis geralmente reduz a soma de quadrados dos resíduos, consequentemente aumentando o valor do R^2 (SAKAMOTO; ISHIGURO; KITAGAWA, 1986).

O coeficiente de determinação ajustado (eq. 2.2), é uma tentativa de tentar corrigir esse problema ajustando o numerador e o denominador da eq. 2.1, através dos respectivos graus de liberdade (DRAPER; SMITH, 1998):

$$R_a^2 = \left(\frac{n-1}{n-k-1} \right) 1 - R^2 \quad (2.2)$$

onde R^2 é um argumento já definido, n é o tamanho da amostra e k o número de parâmetros do modelo.

Contrariamente ao coeficiente de determinação, o coeficiente de determinação ajustado pode diminuir em valor se a contribuição da variável adicional na explicação da variação total, for inferior ao impacto que essa adição acarreta nos graus de liberdade (DRAPER; SMITH, 1998).

O coeficiente de determinação ajustado não possui a mesma interpretação do coeficiente de determinação, no que diz respeito a proporção da variabilidade da variável resposta explicada pelo modelo (CHATERJEE; PRICE, 1991). No entanto, o R_a^2 é também geralmente usado para julgar a qualidade do ajuste de modelos de regressão múltipla possuindo, porém, as mesmas limitações do R^2 , no que se refere a sua aplicação como critério único de seleção de modelos lineares e não lineares.

Por fim, outro critério de seleção de modelos usado no meio florestal, aplicado a regressão múltipla, é o teste F parcial. Este teste é, segundo Rao (1973), usado para comparar dois modelos concorrentes investigando as seguintes hipóteses:

H_0 : Um subconjunto dos coeficientes é igual à zero.

vs.

H_a : Nenhum dos coeficientes da regressão é igual a zero.

Para Draper e Smith (1998), o teste F parcial, diferentemente do teste F do modelo, compara modelos hierárquicos, ou seja, o modelo completo com p coeficientes contra um modelo reduzido composto por um subconjunto desses coeficientes. Trata-se de um critério relativo, pois procura investigar se o modelo reduzido é tão adequado quanto o modelo completo.

Apesar de mais adequado como critério de seleção de modelos, o teste F parcial, compara somente modelos hierárquicos não possibilitando a comparação de modelos com diferentes relações funcionais (DRAPER; SMITH, 1998). Além disso, quando o tamanho da amostra é suficientemente grande, o que geralmente ocorre em problemas de modelagem florestal, o teste F parcial tende a escolher modelos mais complexos e rejeitar modelos parcimoniosos por apresentarem significativa falta de ajuste (KUHA, 2004). Sakamoto, Ishiguro e Kitagawa (1986) afirmam que um modelo com um número desnecessariamente grande de variáveis preditoras pode ser instável e gerar uma idéia falsa de super ajuste.

A partir das limitações da abordagem clássica torna-se interessante considerar abordagens distintas de seleção de modelos para o meio florestal. Uma possível alternativa é o da abordagem das discrepâncias entre dois modelos. Esta abordagem, além da utilidade e flexibilidade, permite uma nova visão no contexto da modelagem (FORSTER, 2000). Ela é particularmente útil por fornecer critérios de seleção relativos, parcimoniosos e aplicáveis a diferentes relações funcionais, ampliando assim, a liberdade do pesquisador e evitando os problemas de super ajuste.

Para se entender essa abordagem primeiro faz-se necessário definir alguns conceitos que auxiliam na compreensão do processo de modelagem de dados empíricos. O conceito de discrepância entre modelos como medida de falta de ajuste, bem como os elementos que compõe essa discrepância serão apresentados a seguir.

2.3 Abordagem das discrepâncias

2.3.1 Conceitos de discrepância

Em modelagem de dados estatísticos busca-se, a partir de um conjunto de dados tomados de uma população, descrever e ou inferir sobre um determinado fenômeno. A descrição completa desse fenômeno é muito difícil, já que as inferências são realizadas a partir de alguns parâmetros da população, arbitrariamente selecionados e estimados segundo um determinado esforço amostral (FORSTER, 2000).

Linhart e Zuchini (1986) denominam de modelo operacional $f(x)$, o modelo mais próximo do fenômeno a ser descrito pelo pesquisador, sendo que apenas em casos excepcionais se tem informação disponível para especificar completamente o modelo operacional. Para os mesmos autores, como regra, em modelagem pode-se apenas obter um modelo relativamente próximo a $f(x)$ a partir do conjunto de dados.

Zuchinni (2000) afirma que, é necessário primeiramente, especificar uma família de modelos de aproximação $g(x|\theta)$ com $\theta \in \Theta$, cujos membros individuais são indentificados pelo vetor de parâmetros $\theta = (\theta_1, \theta_2, \dots, \theta_p)^T$. O modelo ajustado $g(x|\hat{\theta})$, um membro da família de modelos $g(x|\theta)$ com $\theta \in \Theta$, é então obtido através de algum método de estimação dos parâmetros (mínimos

quadrados, máxima verossimilhança, por exemplo).

O próximo passo é selecionar o modelo ajustado $g(x|\hat{\theta})$ mais próximo do modelo operacional $f(x)$ e independente da estratégia utilizada para esta seleção existe como regra um número de aspectos em que o modelo ajustado e o modelo operacional diferem. Cada um desses aspectos de falta de ajuste pode ser visto como alguma discrepância Δ e esta, para que se realize a escolha do melhor modelo, deve ser minimizada (BUCKLAND; BURHAM; AUGUSTIN, 1997). Os mesmos autores dividem a discrepância total em dois componentes: discrepância na aproximação e discrepância na estimação.

A discrepância entre o modelo operacional $f(x)$ e a família de modelos $g(x|\theta)$ é definida por:

$$\Delta(\theta) = \Delta[f(x), g(x|\theta)]. \quad (2.3)$$

Forster (2000) define que a discrepância na aproximação entre a família de modelos $g(x|\theta)$ e o modelo operacional $f(x)$ é dada por $\Delta(\theta_0)$, onde θ_0 é o valor que minimiza a eq. (2.3), tornando $g(x|\theta_0)$ o melhor modelo de aproximação para a família $g(x|\theta)$. Assume-se que θ_0 exista e que é único e isso claramente ocorre se $f(x) \in g(x|\theta)$, com $\theta \in \Theta$.

Desta forma, a discrepância na aproximação não depende de forma alguma do tamanho da amostra e do método de estimação utilizado, mas somente do espaço paramétrico considerado (LINHART; ZUCHINNI, 1986).

Em contrapartida, Linhart e Zuchini (1986) afirmam que a discrepância devido à estimação se refere à discrepância entre o modelo ajustado $g(x|\hat{\theta})$ e o melhor modelo de aproximação $g(x|\theta_0)$ entre os membros a família de modelos de aproximação $g(x|\theta)$, como segue:

$$\Delta[g(x|\theta_0), g(x|\hat{\theta})]. \quad (2.4)$$

A discrepância na estimação, ao contrário da discrepância na aproximação, depende fortemente do tamanho da amostra e dos métodos de estimação utilizados (FORSTER, 2000).

No processo de seleção de modelos, quanto maior for o espaço paramétrico considerado de um modelo, menor será sua discrepância na aproximação em relação ao modelo operacional. Em outras palavras, se apenas a discrepância na aproximação for utilizada como quesito para escolha de modelos concorrentes esta privilegiará sempre os modelos de maior complexidade (FORSTER, 2000). Por outro lado, segundo o mesmo autor, modelos mais complexos tendem à apresentar maior incerteza na estimativa dos seus parâmetros, ou seja, tendem à apresentar maior discrepância na estimação.

O problema de seleção de modelos é encontrar, dentro deste contexto, um caminho para conciliar essas duas propriedades opostas e encontrar o modelo de melhor performance geral (ZUCHINNI, 2000).

A discrepância geral $\Delta(\theta_0)$ ou a discrepância entre o modelo operacional $f(x)$ e o modelo ajustado $g(x|\hat{\theta})$ é definida em (ZUCHINNI, 2000) por:

$$\Delta(\theta_0) = \Delta[f(x), g(x|\hat{\theta})]. \quad (2.5)$$

Burham e Anderson (2002) afirmam que, na prática, não é possível calcular nenhuma das discrepâncias acima, pois não se conhece completamente o modelo operacional $f(x)$. Ou seja, para uma determinada amostra não é possível calcular diretamente a discrepância geral e nem mesmo seus dois componentes.

No entanto, segundo Zuchinni (2000), pode-se obter alguma ajuda na seleção de modelos através do cálculo do seu valor esperado para uma determinada amostra. Esta é chamada de função de discrepância geral esperada, definida por:

$$E\Delta(\theta_0) = \Delta[f(x), g(x|\hat{\theta})]. \quad (2.6)$$

Porém, também não é possível calcular a discrepância geral estimada, representada pela eq. 2.6, sem o conhecimento do modelo operacional $f(x)$. Uma saída para este problema é estimar essa discrepância a partir de evidências fornecidas pelo conjunto de dados, obtendo-se um estimador da discrepância geral esperada. Este estimador é, de acordo com Zuchinni (2000), um critério de seleção de modelos baseado na perspectiva das discrepâncias.

Uma função de discrepância geral esperada largamente utilizada é a chamada distância entre dois modelos de Kulbach-Leibler. Esta medida de discrepância, baseada no princípio de entropia máxima de Boltzman (veja Burham e Anderson (2002) para mais detalhes) forneceu uma importante base teórica para criação de diversos critérios de seleção de modelos nas últimas décadas, incluindo o Critério de Informação de Akaike.

2.3.2 Distância de Kulbach-Leibler e Entropia

A interpretação estatística de entropia, uma medida do aumento de energia de qualquer sistema isolado termodinamicamente, foi desenvolvida por Boltzmann no final do século XIX (AKAIKE, 1985). Essa foi definida em Boltzman (1877 apud AKAIKE, 1981) como uma medida da entropia de uma distribuição $g(x)$ em relação a outra $f(x)$:

$$B(f, g) = - \int f(x) \log \left[\frac{f(x)}{g(x)} \right] dx. \quad (2.7)$$

A entropia é critério natural do ajuste de $g(x)$ em relação a $f(x)$ e quanto maior a entropia melhor é aproximação entre os modelos, sendo este chamado de princípio da máxima entropia (AKAIKE, 1981).

Desta forma, dois modelos estão mais próximos quanto maior a entropia entre eles ou quanto menor a entropia negativa ou “neg-entropia” (BOSDOGAN, 2000).

$$I(f, g) = -B(f, g) \quad (2.8)$$

A entropia negativa é também chamada de distância de Kulbach-Leibler ($I(f, g)$) e por consequência a minimização desta distância é equivalente a maximização da entropia de Boltzman (BOSDOGAN, 2000).

No processo de modelagem precisamos escolher o melhor modelo de aproximação $g(x|\theta_0)$ em relação ao modelo operacional $f(x)$, que gerou os dados. Sendo assim e considerando que se conhece o modelo operacional, pode-se medir a distância entre este e diversas famílias de aproximação propostas $g(x|\theta)$, através do cálculo da distância de Kulbach-Leibler (LINHART; ZUCHINNI, 1986).

Neste caso, considerando-se o modelo operacional $f(x)$ como dado (fixo) e somente $g(x|\theta)$ variando num espaço de modelos θ , procura-se uma família de aproximação que minimize $I(f, g)$. Isto pode ser feito aplicando a eq. 2.9 às diferentes famílias de aproximação $g(x|\theta)$ e ao modelo operacional $f(x)$

$$I(f, g) = \int f(x) \log \left[\frac{f(x)}{g(x|\theta)} \right] dx. \quad (2.9)$$

A distância de Kulbach-Leibler, segundo Burham e Anderson (2004) pode ser vista como medida de discrepância da família $g(x|\theta)$ ao se aproximar de $f(x)$, sendo, no entanto, necessário se conhecer completamente $f(x)$ e ainda conhecer os parâmetros de $g(x|\theta)$ para ser calculada.

Sendo assim, a distância de Kulbach-Leibler pode ser conceitualizada como uma distância direta entre dois modelos e, segundo Wada e Kashiwagi (1990), é a mais fundamental de todas as medidas de informação pela simplicidade e propriedades aditivas, tornando-a, em conjunto com a teoria de máxima verossimilhança, uma base racional para seleção de modelos.

Como dito até agora, para se calcular a distância de Kulbach-Leibler entre $f(x)$ e $g(x)$ faz-se necessário conhecer ambos os modelos bem como seus parâmetros. Porém, usando uma medida de distância relativa, pode-se comparar vários modelos de aproximação $g(x|\theta)$ sem a necessidade de conhecer o modelo operacional. Isto ocorre, pois, a eq. 2.9 pode ser também apresentada da seguinte forma:

$$I(f, g) = \int f(x) \log[f(x)] dx - \int f(x) \log[g(x|\theta)] dx. \quad (2.10)$$

Cada um dos dois termos da eq. 2.10, segundo Burham e Anderson (2000), se refere ao valor esperado em relação ao modelo operacional $f(x)$, podendo ser reescrita como a diferença de duas esperanças em relação a $f(x)$, como segue:

$$I(f, g) = E_f[\log(f(x))] - E_f[\log(g(x|\theta))]. \quad (2.11)$$

O primeiro termo é uma constante que depende do modelo operacional $f(x)$. Chamando esse termo constante de C , tem-se:

$$I(f, g) = C - E_f[\log(g(x|\theta))] \quad (2.12)$$

ou ainda

$$I(f, g) - C = -E_f[\log(g(x|\theta))]. \quad (2.13)$$

Sendo assim, tratando o termo desconhecido como uma constante, é possível se calcular a distância relativa entre várias famílias de aproximação candidatas $g(x|\theta)$ e o modelo operacional $f(x)$ (BOSDOGAN, 1987).

O termo $I(f, g) - C$ é a distância relativa entre $f(x)$ e $g(x|\theta)$, tornando $-E_f[\log(g(x|\theta))]$ a quantidade de interesse para selecionar o melhor modelo de aproximação $g(x|\theta_0)$ entre várias famílias de aproximação $g(x|\theta)$ candidatas. Desta forma, pode-se postular várias famílias de aproximação $g(x|\theta)$ e selecionar a melhor entre elas.

Calcular, no entanto, a distância relativa entre o modelo operacional $f(x)$ e as várias candidatas $g(x|\theta)$, é calcular somente a discrepância na aproximação. Mas o que se busca é um modelo que minimize a discrepância geral, já que na prática não se tem informação sobre o parâmetro θ da família de aproximação, sendo necessário estima-lo a partir do conjunto de dados. Neste caso, precisa-se estimar a distância entre $g(x|\hat{\theta})$ e $f(x)$ para que se escolha o modelo com a menor distância relativa estimada, ou modelo com menor discrepância geral esperada estimada. Akaike (1981) encontrou uma maneira de obter a estimativa da discrepância geral esperada de Kulbach-Leibler, baseada na função de log-verossimilhança no seu ponto máximo (BURHAM; ANDERSON, 2002).

2.3.3 Critério de informação de Akaike

Akaike (1974, 1981) afirma que aplicar a função de discrepância geral esperada de Kulbach-Leibler como critério de seleção de modelos é aplicar essa dupla esperança:

$$E_y E_x [\log(g(x|\hat{\theta}(y)))] \quad (2.14)$$

onde o termo interno é equivalente a $E_f[\log(g(x|\theta))]$ (distância de Kulbach-Leibler), com o parâmetro θ sendo substituído pela estimativa de máxima verossimilhança de θ , baseada na família de distribuição $g(x|\theta)$ e no conjunto de dados y .

Segundo Burham e Anderson (2004), apesar de y denotar o conjunto de dados, x e y devem ser conceitualizadas como sendo amostras aleatórias tomadas do mesmo modelo operacional $f(x)$ e, portanto, com as duas esperanças tomadas em relação a este mesmo modelo.

Akaike (1974, 1981) relacionou, então a distância de Kulbach-Leibler e teoria da verossimilhança, demonstrando que a estimativa de máxima log-verossimilhança era um estimador enviesado de $E_y E_x [\log(g(x|\hat{\theta}(y)))]$.

Bosdogan (2000); Burham e Anderson (2004) afirmam que, apesar de $\log(L(\hat{\theta}|\text{dados}))$ ser um estimador enviesado de $E_y E_x[\log(g(x|\hat{\theta}(y)))]$, este viés é aproximadamente igual ao número de parâmetros p da família de aproximação $g(x|\theta)$. Sendo esse um resultado assintótico de fundamental importância (BURHAM; ANDERSON, 2004).

Sendo assim, um estimador não enviesado de $E_y E_x[\log(g(x|\hat{\theta}(y)))]$ para grandes amostras, como se tem geralmente na meio florestal, torna-se:

$$\log(L(\hat{\theta}|\text{dados})) - p. \quad (2.15)$$

Sendo esse resultado, então equivalente à:

$$\log(L(\hat{\theta}|\text{dados})) - p = C - \hat{E}_\theta[I(f, \hat{g})], \quad (2.16)$$

onde $\hat{g} = g(x|\hat{\theta})$.

Este estimador da distância relativa de Kulbach-Leibler, segundo Bosdogan (2000), torna possível combinar estimação e seleção de modelos numa estrutura de otimização unificada. Em outras palavras, pode-se através de um unico critério estimar a discrepância geral esperada de uma família de aproximação a partir de um conjunto de dados.

Desta maneira, Akaike (1973 apud SAKAMOTO; ISHIGURO; KITAGAWA, 1986) encontrou um estimador da distância relativa esperada de Kulbach-Leibler baseada na função de máxima log-verossimilhança, corrigida pelo viés assintótico p

$$\log(L(\hat{\theta}|\text{dados})) - p = \hat{E}[I(f, g)] \textit{ relativa}. \quad (2.17)$$

O termo p é o termo de correção do viés da estimativa e é importante ressaltar que esse termo não é de forma alguma arbitrário e confere ao AIC propriedades assintóticas desejadas (veja Burham e Anderson (2004) para mais detalhes)

Sendo assim, Akaike (1973) (1973 apud SAKAMOTO; ISHIGURO; KITAGAWA, 1986) definiu “an information criterion” (AIC) multiplicando os dois termos desse resultado por -2 , ficando este conhecido como Akaike Information Criterion, ou simplesmente AIC.

$$AIC = -2\log(L(\hat{\theta}|\text{dados})) + 2p \quad (2.18)$$

O motivo da multiplicação desses termos por -2 é importante, mas completamente técnico e não será alvo desta revisão, sendo que para maiores detalhes veja Bosdogan (1987) e Akaike (1985).

Sakamoto; Ishiguro e kitagawa (1986) afirmam que para o modelo homocedástico e com erros gaussianos, comumente usados no meio florestal, o AIC torna-se:

$$AIC = -2\log(L((\hat{\sigma}^2)) + 2p \quad (2.19)$$

onde n é o tamanho da amostra, p é o número de parâmetros (incluindo os coeficientes do modelo e $\hat{\sigma}^2$) e $\hat{\sigma}^2 = \frac{\sum(\hat{\epsilon}^2)}{n}$.

Desta forma, o valor de AIC é simples de ser obtido para os casos de estimação por mínimos quadrados, como é o caso do ajuste de modelos de regressão, e para os casos de análises baseadas na estimativas de verossimilhança de uma forma geral (BURHAM; ANDERSON, 2004).

Assumindo uma série de famílias de aproximação *a priori* selecionadas e fundamentadas em alguma teoria científica, o AIC pode ser obtido para cada família, classificando-as da melhor para a pior com base no conjunto de dados disponível. As melhores famílias de aproximação serão aquelas com menores valores de AIC, ou com equivalentes menores estimativas de discrepância geral esperada, em relação ao modelo operacional (BOSDOGAN, 2000).

Se parâmetros conhecidos forem adicionados a uma família de aproximação candidata, o valor de AIC diminuirá. Mas como não se conhece, em geral, os parâmetros de uma determinada família, sendo estes apenas estimados a partir dos dados, o acréscimo de parâmetros irá aumentar a incerteza na estimação, e por consequência, aumentar o valor de AIC (KUHA, 2004).

Neste sentido o AIC é um critério de seleção que privilegiará famílias de aproximação parcimoniosas mais próximas ao modelo operacional. No entanto é preciso ressaltar que este critério irá apenas classificar as famílias propostas, escolhendo uma entre elas, mesmo que todas estejam bem distantes do modelo operacional (BUCKLAND; BURHAM; AUGUSTIN, 1997).

2.4 Considerações Finais

Alguns autores (FORSTER, 2000; KUHA, 2004), interpretam o AIC também como um otimizador do balanço entre entrada e saída de parâmetros no processo de escolha de modelos parcimoniosos. No entanto, apesar dessa interpretação não ser incorreta, o presente trabalho prefere a interpretação do AIC como estimador da discrepância geral esperada, de acordo com a perspectiva de Burham e Anderson (2004) e Linhart e Zuchinni (1986).

Na segunda perspectiva fica claro que o AIC, apesar da sua simplicidade quanto a forma e quanto a aplicação, é baseado em conceitos profundos e extremamente consolidados da teoria da informação (Distância de Kulbach-Leibler) e da teoria estatística (paradigma da verossimilhança).

Outro conceito importante no processo de modelagem, inserido nessa perspectiva, é o da intangibilidade do modelo operacional. Em outras palavras, a realidade completa nunca pode ser descrita a partir de um modelo, sendo esta apenas circunscrita através de uma série de modelos de aproximação concorrentes.

Isto implica que o processo de modelagem não é simplesmente um processo de estimação, mas também um processo de aproximação de uma realidade que não pode ser descrita totalmente pelo modelo.

No meio florestal, como já visto, se usam critérios de seleção de modelos que não permitem

comparar modelos de diferentes relações funcionais, simplificando o processo de modelagem à estimação e à definição da ordem (número de parâmetros) do modelo. Além disso, não existe nenhum critério que aplicado isoladamente classifique os modelos, quanto a sua falta de ajuste, em relação ao modelo operacional.

A hipótese deste trabalho é que o AIC pode, se interpretado e usado corretamente sob o ponto de vista das discrepâncias, vir a ser uma excelente ferramenta de seleção de modelos empíricos de predição no meio florestal.

Referências

AKAIKE, H. A New Look at Statistical Model Identification. **IEEE Transactions on Automatic Control**, Tokio, v. 19, n. 6, p. 717-723, Dec. 1974.

AKAIKE, H. Likelihood of a Model and Information Criteria. **Journal of Econometrics**, Amsterdam, v. 16, n.1, p. 3-14, May. 1981.

ANSCOMBE, F.J. Graphs in Statistical Analysis. **The American Statistician**, Alexandria, v. 27, n. 1, p. 17-21, Feb. 1973.

BATES, D.M.; WATTS, D.G. Nonlinear regression analysis and its applications. 2nd.ed. New York : Wiley, 1988. 371 p.

BOSDOGAN, H. Model selection and akaike information criterion (AIC): The general theory and its analitical extensions. **Psychometrika**, New York, v. 52, n. 3, p. 345-370, Sept. 1987.

BOSDOGAN, H. Akaike information criterion and recent developments in information complexity. **Journal of Mathematical Psychology**, Oxford, v. 44, n. 1, p. 62-91, Mar. 2000.

BUCKLAND, S.T.; BURNHAM, K.P.; AUGUSTIN, N.H. Model selection: A integral part of inference. **Biometrics**, London, v. 53, n. 2, p. 603-618, Jun. 1997.

BURNHAM, K.P.; ANDERSON, D.R. Model selection and multimodel inference: Practical information theoretic approach. 2nd.ed. New York: Springer, 2002. 488 p.

BURNHAM, K.P.; ANDERSON, D.R. Multimodel inference: Understanding AIC e BIC in model selection, **Sociological Methods Reseach**. London, v. 33, n. 2, p. 261-304, Nov. 2004.

CAMERON, A.C.; WINDMEIJER, F.A.G. An R-squared measure of goodness of fit for some common nonlinear regression models. **Journal of Econometrics**, Amsterdam, v. 77, n. 2, p. 329-342, Apr. 1997.

CAMPOS, J.C.C.; CAMPOS, A.L.A.S.; LEITE, H.G. Decisão silvicultural empregando um sistema de predição do crescimento e da produção. **Revista Árvore**, Viçosa, v. 12, n. 2, p. 100-110, mai. 1988.

CAMPOS, J.C.C.; SOARES, C.P.B.; LEITE, H.G.; CAMPOS, M.B. Estimação de diferentes volumes comerciais utilizando um modelo do tipo povoamento total. **Revista Árvore**, Viçosa, v. 25, n. 2, p. 223-230, set. 2001.

CHATTERJEE, S.; PRICE, B. Regression analysis by example. 2nd.ed. New York: Wiley, 1977. 228 p.

DRAPER, N.R.; SMITH, H. Applied regression analysis. 3rd.ed. New York: Wiley, 1998. 706 p.

FORSTER, M.R. Key concepts in model selection: Performance and generalizability. **Journal of Mathematical Psychology**, Oxford, v. 44, n. 1, p. 205-231, Mar. 2000.

KUHA, J. AIC e BIC: Comparisons of assumptions and performance. **Sociological Methods Research**, London, v. 33, n. 2, p. 188-228, Nov. 2004.

LEITE, H.G.; NOGUEIRA, G.S.; MOREIRA, A.M.; LIMA, J.E. Um modelo de crescimento e produção para *Pinus taeda* L. na Região Sul do Brasil. **Revista Árvore**, Viçosa, v. 25, n. 1, p. 105-112, mai. 2001.

LINHART, H., ZUCHINNI, W. Model selection. 2nd.ed. New York: Wiley, 1986. 301 p.

MAESTRI, R.; SANQUETTA, C.R.; MACHADO, S.A.; SCOLFORO, J.R.S.; CORTE, A.P.D. Viabilidade de um projeto florestal de *Eucalyptus grandis* considerando o sequestro de carbono. **Floresta**, Curitiba, v. 34, n. 3, p. 347-360, dez. 2004.

RAO, C. Linear statistical inference and its applications. 2nd.ed. New York: Wiley, 1973. 625 p.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. Akaike information criterion statistics. Tokio: KTK Scientific Publisher/D. Riedel, 1986. 290 p.

SOARES, C.P.B.; OLIVEIRA, M.L.R. Equações para estimar o carbono na parte aérea de árvores de eucalipto em Viçosa, Minas Gerais. **Revista Árvore**, Viçosa, v. 26, n. 5, p. 533-539, set. 2002.

SOARES, C.P.B.; LEITE, H.G.; GÖRGENS, E.B. Equações para estimar o estoque de carbono no fuste de árvores individuais e em plantios comerciais de eucalipto. **Revista Árvore**, Viçosa, v. 29, n. 5, p. 711-718, set. 2005.

WADA, Y.; KASHIWAGI, N. Selecting statistical models with information statistics, **Journal of Dairy Science**. Palo Alto, v.73, n. 2, p. 3575-3582, Jun. 1990.

ZUCHINNI, W. An introduction to model selection, **Journal of Mathematical Psychology**, Oxford, v. 44, n. 1, p. 41-61, Mar. 2000.

3 AVALIAÇÃO DAS TÉCNICAS DE DETERMINAÇÃO DA BIOMASSA LENHOSA DE ÁRVORES

Resumo

O objetivo deste capítulo é apresentar e discutir as duas técnicas disponíveis para determinação da biomassa do tronco de árvores nativas da Floresta Tropical Atlântica. Para tanto, foram coletadas 200 árvores de dez espécies nativas da floresta atlântica para determinação da sua biomassa através da técnica gravimétrica e da técnica volumétrica. Em cada árvore foram tomados cinco discos referentes a cinco posições da altura comercial (até a primeira bifurcação): 0% (base), 25%, 50% (meio), 75% e 100% (topo). Na técnica gravimétrica, estudou-se diferentes técnicas de composição do teor de umidade dos discos para determinação da biomassa, concluindo-se como a melhor a que utiliza a média aritmética dos discos da base, meio e topo. Na técnica volumétrica, estudou-se diferentes técnicas de composição da densidade do tronco com base nas densidades básicas dos discos. Em termos de densidade do tronco, a média aritmética das densidades básicas dos cinco discos se mostrou como melhor técnica. Entretanto, quando se multiplica a densidade do tronco pelo volume deste para obtenção da biomassa, a utilização da densidade básica do disco do meio se mostrou superior a todas as técnicas. A utilização de uma densidade básica média da espécie para determinação da biomassa, via técnica volumétrica, se apresentou como uma abordagem inferior a qualquer técnica que utiliza informação da densidade do tronco das árvores individualmente.

Palavras-chave: Técnica gravimétrica; Técnica volumétrica; Densidade básica; Teor de umidade; Discrepância geral; Discrepância de aproximação; Discrepância de estimação

Abstract

Two Different techniques for determination of stem biomass are studied: the gravimetric and the volumetric technique. The gravimetric technique obtains biomass by measuring the green weight of the stem and discounting its water content. The volumetric technique obtains biomass by the product of stem volume and wood density of the stem. Both techniques were studied in ten tree species from Atlantic Forest in the State of Espírito Santo, Brazil, based on destructive sampling of 200 trees. In each tree, 5 cross-section discs were sampled in relative positions to commercial height: 0% (base), 25%, 50% (middle), 75% e 100% (top). For the determination of water content of the stem, the arithmetic mean of the discs of base, middle and top showed the best results. For stem density determination, the arithmetic mean of the the basic density of the five discs showed the best estimation. But when stem density is multiplied by the volume of the stem to predict stem biomass, the use of the middle disc basic density as surrogate of stem density generated the best predictions. The use of a average wood basic density for each species generated stem biomass predictions worse than any method of prection that uses stem density determination at the individual tree basis.

Keywords: Gravimetric technique; Volumetric technique; Basic density; Stem water content; General discrepancy; Approximation discrepancy; Estimation discrepancy

3.1 Introdução

Segundo Wendling (1998), o termo biomassa florestal refere-se a quantidade constituída por organismos no ecossistema florestal em termos de massa. Já, Teixeira (2003), define a biomassa como a quantidade de material vegetal contida por unidade de área numa floresta e expressa em unidade de massa.

Na literatura, o acesso à biomassa florestal ocorre, geralmente, por meio de dois métodos distintos. Vieira et al. (2008) chamam de métodos diretos os que envolvem determinações diretas de biomassa e de métodos indiretos os métodos que envolvem algum tipo de estimativa da biomassa a partir de relações alométricas. Métodos diretos são, segundo os mesmos autores, aplicáveis à áreas de pequena extensão enquanto que os métodos indiretos são aplicáveis à grandes áreas quando a relação empírica obtida é abrangente.

Os métodos diretos podem ser divididos dois tipos, o método da árvore individual e o método da parcela. No primeiro caso, árvores-amostra são eleitas para que determinações diretas de biomassa sejam realizadas. Na segunda, uma determinada unidade de área é pré definida e toda a biomassa ali contida é determinada (SANQUETTA; BALBINOT, 2004).

Os dados produzidos no processo de determinação da biomassa pelo método da árvore individual são usados, geralmente no ajuste de relações alométricas que irão gerar as estimativas de biomassa em grandes áreas pelos métodos indiretos (SANQUETTA, 2002).

Os métodos indiretos, desta forma, são baseados nos dados produzidos no processo de determinação da biomassa em pequenas áreas. Sendo assim, a precisão destas estimativas depende, dentre outros fatores, da precisão do método direto usado para determinação da biomassa.

Essa determinação, segundo Sanqueta (2002), pode ser aplicada à todos os componentes do ecossistema florestal, como o estrato arbóreo, arbustivo, serapilheira e o solo, ou aplicado à apenas um desses componentes, conforme o objetivo do trabalho.

No entanto, as determinações de biomassa mais comuns se referem a biomassa do estrato arbóreo acima do solo, por ser este o compartimento de fácil acesso e o que mais contribui com a biomassa total no ecossistema florestal (SANQUETTA; BALBINOT, 2004).

Dentro do compartimento estrato arbóreo acima do solo, o método direto de determinação da biomassa pode ser aplicado na árvore com um todo ou através da separação dos seus diferentes componentes (fuste, galhos e folhas). O componente fuste, seguramente, compõe a fração mais representativa da biomassa na árvore e da biomassa florestal, sendo que sua determinação, por isso, deve ser feita da forma mais precisa possível (SANQUETTA, 2002).

A determinação da biomassa do fuste de árvores é usualmente realizada através de técnicas gravimétricas e volumétricas. A técnica gravimétrica consiste da pesagem de secções do fuste da árvore no campo e coleta de amostras deste fuste para posterior determinação do seu teor de

umidade médio em laboratório. De posse do teor de umidade médio, este é aplicado ao fuste como um todo e a biomassa é, então determinada (HIGUSHI; CARVALHO JR., 1994).

A técnica volumétrica é aplicada, segundo Sanquetta (2002), para facilitar os trabalhos de campo. Esta consiste em determinar o volume da árvore através de algum método de cubagem rigorosa, coletando-se discos de madeira para determinação da densidade básica média da árvore e consequente conversão do seu volume em biomassa.

As duas técnicas caracterizam o método direto e podem ser aplicadas de maneira isolada ou em conjunto (ARAUJO; HIGUSHI; CARVALHO JR., 1996), de acordo com os objetivos do trabalho. Geralmente a técnica gravimétrica é aplicada em árvores de pequeno porte onde os fustes são fáceis de manusear e a volumetria é aplicada às árvores de grande porte, cujos fustes são, mesmo seccionados, de difícil manuseio (HIGUSHI; CARVALHO JR., 1994).

É comum na literatura encontrar trabalhos onde estimativas de densidades médias por espécies, obtidas empiricamente ou retiradas da literatura, são usadas associadas a dados de volumetria, para determinação da biomassa arbórea acima do solo (FEARNSIDE, 1997; BARBOSA; FEARNSIDE, 2004; NOGUEIRA; NELSON; FEARNSIDE, 2004). Esses trabalhos, realizados na região amazônica, justificam sua metodologia pelo elevado custo de se estabelecer equações empíricas de biomassa e pela pré existência de dados de inventário volumétrico na área em questão.

No entanto, sabe-se que existe muita variação na densidade entre indivíduos de uma mesma espécie. Essas variações tem origem, segundo Vital (1984), na idade dos indivíduos, na variação genotípica dentro da espécie, na localização geográfica, nos aspectos climáticos e no aspectos de manejo. Neste contexto, métodos de determinação da biomassa de árvores individuais, baseadas na utilização da densidade média da espécie, podem levar a estimativas pouco confiáveis (NOGUEIRA; NELSON; FEARNSIDE, 2004).

A técnica volumétrica possui duas fontes de erros associadas, sendo uma a incerteza relacionada à determinação do volume, já que envolve o processo de medição e de totalização através de uma fórmula de aproximação, e a outra relacionada a determinação da densidade básica média, em nível de árvore, ou em nível de espécie.

A técnica gravimétrica, por suposição, possui menos fontes de incerteza, já que consiste na medição direta da massa verde da árvore no campo, sendo a que a única fonte de incerteza associada a esta técnica se refere a estimativa do teor de umidade para posterior conversão em biomassa.

Neste trabalho, a técnica gravimétrica e a técnica volumétrica foram aplicadas conjuntamente na determinação da biomassa do caule de 200 árvores de essências nativas da floresta tropical atlântica, de diferentes tamanhos e idades, provenientes de reflorestamento e localizadas em Linhares, ES.

Estas técnicas foram aplicadas com o objetivo de estudar as duas fontes de incertezas associadas a técnica volumétrica na determinação da biomassa pelo método direto. Este estudo foi feito através da verificação do impacto do método de obtenção da densidade média por árvore ou por espécie

e da determinação do volume na determinação da biomassa de árvores individuais, partindo do pressuposto de que a técnica gravimétrica é mais precisa, servindo portanto, de parâmetro para comparação.

3.2 Material e Métodos

3.2.1 Área de estudo

O estudo foi realizado em área pertencente à Companhia Vale do Rio Doce, recuperada com essências nativas e localizada ao norte do Estado do Espírito Santo, predominantemente, no município de Linhares com uma pequena abrangência nos de Sooretama e de Jaguaré. Geograficamente situa-se entre os paralelos 19° 06'-19° 18' de latitude sul e entre os meridianos 39° 45'-40° 19' de longitude oeste, contando com uma área de 22.000 hectares.

3.2.2 Escolha das espécies

Os indivíduos arbóreos usados no trabalho foram selecionadas, segundo um processo de amostragem estratificada, sendo cada classe de diâmetro definida com um estrato. Ela foi uniforme nos estratos, ou seja, cada classe de diâmetro foi contemplada com dois indivíduos.

A escolha das espécies para realização deste estudo seguiu dois critérios que juntos buscavam uma certa generalidade quanto a aplicação deste em áreas distintas àquela onde o trabalho foi realizado. Primeiramente buscou-se espécies de ampla ocorrência geográfica e entre essas foram escolhidas espécies que possuíam, segundo a literatura, densidades básicas bem distintas. Desta forma foram escolhidas dez espécies a saber: *Caesalpinia ferrea*, *Pteropcarpus violaceus*, *Zeyhera tuberculosa*, *Schizolobium parayba*, *Pterygota brasiliensis*, *Cordia trichotoma*, *Astronium graveolens*, *Joannesia princeps*, *Cariniana legalis* e *Tabebuia serritifolia*.

Desta maneira o presente trabalho contou com 200 indivíduos de dez espécies distintas divididas em dez classes de diâmetro, sendo cada classe contemplada com dois indivíduos.

3.2.3 Determinação da biomassa

A determinação da biomassa do fuste das árvores se deu através de duas técnicas comumente utilizadas para obtenção da biomassa acima do solo de árvores individuais: a técnica gravimétrica e a técnica volumétrica. A descrição de como as duas técnicas foram aplicadas no presente trabalho é apresentada a seguir.

Técnica gravimétrica

Primeiramente a massa verde do tronco foi medida com o auxílio de uma célula de carga com capacidade máxima de 200 kg e precisão de centenas de gramas. Discos de madeira com casca de

aproximadamente 5 cm de espessura foram retirados do tronco da árvore para determinação do teor de umidade do tronco. A biomassa é obtida pelo produto:

$$b_{Gi} = m_i (1 - u_i/100) \quad (3.1)$$

onde: i é o índice que se refere às árvores; b_{Gi} é a biomassa do tronco determinada pela técnica gravimétrica (kg); m_i é a massa verde do tronco (kg); e u_i é o teor de umidade do tronco (%).

Técnica volumétrica

O fuste das árvores foram cubados rigorosamente pelo método de Smalian, com diâmetros tomados a cada metro, para obtenção de seu volume sólido. Os discos extraídos para determinação do teor de umidade serviram também como amostras de madeira para a determinação da densidade básica da madeira do tronco. A biomassa do tronco é obtida pelo produto das duas grandezas:

$$b_{Vi} = v_i d_{Bi} \quad (3.2)$$

onde: i é o índice que se refere às árvores; b_{Vi} é a biomassa do tronco determinada pela técnica volumétrica (kg); v_i é o volume do tronco obtido por cubagem (m^3); d_{Bi} é a densidade básica do tronco ($kg\ m^{-3}$).

Considera-se que a técnica gravimétrica produz melhor determinação da biomassa do tronco, uma vez que a medida da massa verde e a determinação do teor de umidade podem ser realizadas com grande precisão e acurácia. O método volumétrico, é de operacionalização mais fácil que o método gravimétrico, particularmente no caso de grandes árvores, mas está sujeito a um número maior de fontes de erro. A determinação do volume do tronco por cubagem rigorosa possui duas pressuposições, que podem ser problemáticas para árvores nativas: (1) a secção transversal do tronco é circular, (2) a forma do tronco entre duas medidas do diâmetro do tronco pode ser aproximada por um parabolóide. Já a determinação da densidade básica dos discos, envolve uma operação de laboratório mais complexa que a determinação do teor de umidade.

3.2.4 Os Discos de Madeira Amostrados do Tronco

De cada árvore estudada, foram retirados do tronco discos de madeira com casca de aproximadamente 5 cm de espessura em 5 posições relativas referentes à altura comercial da árvore: 0% (disco da base), 25%, 50% (disco do meio), 75% e 100% (disco do topo). Entende-se por altura comercial a altura até a primeira bifurcação do tronco. Cada disco teve sua massa verde medida no campo com balança analítica com precisão de 1 (um) grama e foi acondicionado em saco plástico para transporte.

Os discos foram utilizados tanto na determinação teor de umidade quanto da densidade básica do tronco. A densidade básica de cada disco foi determinada pelo método da balança hidrostática,

seguindo as normas da ABCP M14/70 e a metodologia de Vital (1984). Para determinação do teor de umidade, a massa seca dos discos foi medida através de secagem em estufa de circulação forçada de ar a 103°C até peso constante. O teor de umidade foi determinado pela razão da massa seca pela massa verde do disco.

3.2.5 Determinação do Teor de Umidade do Tronco

Técnicas de Combinação dos Teor de Umidade dos Discos

Para determinar o teor de umidade do tronco, contou-se com os cinco discos amostrados de forma que foi possível avaliar diferentes técnicas de combinação dos valores de teor de umidade para obter um teor de umidade geral do tronco. A melhor medida que se pode obter é o uso dos cinco discos, nesse caso utilizou-se a média aritmética do teor de umidade dos cinco discos do tronco e assumiu-se como o teor de umidade do tronco. Outras combinações de discos foram estudadas para se verificar se o número de discos amostrados poderia ser reduzido sem comprometer a precisão e acurácia com que o teor de umidade é determinado. Novamente a combinação dos teores de umidades dos discos foi realizada através da média aritmética. A técnicas estudadas são apresentadas na Tabela 3.1.

Tabela 3.1 – Técnicas de combinação do teor de umidade dos discos amostrados para obtenção do teor de umidade do tronco

Variável	Descrição	Posições Relativas
u	Teor de Umidade do Tronco	0%-25%-50%-75%-100%
u_b	valor do disco da base	0%
u_t	valor do disco da topo	100%
u_m	valor do disco do meio	50%
u_{bt}	média dos discos da base e do topo	0%-100%
u_{bmt}	média dos discos da base, do meio e do topo	0%-50%-100%

Modelos de Predição do Teor de Umidade do Tronco

Cada uma das técnicas de combinação dos teores de umidade dos discos gera uma predição do teor de umidade do tronco. A forma *operacional* de gerar essa predição é simplesmente tomando o valor da combinação do teor de umidade dos discos como sendo o teor de umidade do tronco:

$$\hat{u}_i = u_{ki}$$

onde \hat{u} é a predição do teor de umidade do tronco e u_k é a combinação dos teores de umidade dos discos segundo as técnicas apresentadas na tabela 3.1.

Assumindo uma relação linear entre o teor de umidade do tronco (u) e as combinações de teores de umidades dos discos (u_k), cada combinação teria um *modelo operacional de predição* da seguinte forma:

$$u_i = (0) + (1) u_{ki} + \varepsilon_i \quad (3.3)$$

onde ε_i é um erro aleatório com distribuição Gaussiana de média nula e variância constante. Os valores entre parênteses no modelo são para enfatizar que, no modelo operacional, o coeficiente do intercepto é nulo (0) e o coeficiente de inclinação é unitário (1), pois a combinação dos teores de umidade dos discos é tomada como sendo o teor de umidade do tronco.

O modelo operacional é o modelo que será utilizado para determinação do teor de umidade do tronco. Mas nesse estudo, pode-se analisar quais das combinações dos teores de umidade dos discos tem a *melhor aproximação linear* do teor de umidade do tronco. Essa melhor aproximação linear é obtida ao se permitir que os coeficientes do modelo possam assumir outros valores, gerando o *modelo de aproximação*:

$$u_i = \beta_0 + \beta_1 u_{ki} + \varepsilon_i \quad (3.4)$$

onde β_0 e β_1 são os coeficientes que conferem a *melhor aproximação linear* entre as variáveis, e ε_i é um erro aleatório com distribuição Gaussiana de média nula e variância constante.

Note que o modelo de aproximação não pode e não deve ser utilizado operacionalmente, mas ele permite medir qual das combinações possui a *melhor aproximação* para o teor de umidade do tronco. A melhor aproximação não implica necessariamente na *melhor predição*. Note também que o modelo de aproximação *sempre* terá melhor ajuste às observações, uma vez que seus coeficientes podem ser livremente estimados para se obter a *melhor relação linear* entre as variáveis.

3.2.6 Determinação da Densidade da Madeira do Tronco

Técnicas de Combinação das Densidades Básicas dos Discos

Na técnica volumétrica de determinação da biomassa, o volume do tronco deve ser multiplicado pela densidade da madeira do tronco. Diferentes combinações das densidades básicas dos cinco discos amostrais podem ser utilizadas como forma de obter a densidade da madeira do tronco. A Tabela 3.2 apresenta as técnicas de combinação analisadas nesse estudo.

A qualidade dessas técnicas de determinação da densidade do tronco foi aferida analisando a sua relação com a densidade obtida pela razão da biomassa gravimétrica e o volume do tronco. Essa razão não é necessariamente a melhor medida de *densidade básica da árvore*, mas é a melhor medida de densidade a ser multiplicada pelo volume para se obter a biomassa do tronco. Essa razão foi designada simplesmente por *densidade do tronco*, sendo tomada como a medida correta da densidade.

Tabela 3.2 – Técnicas de combinação da densidade básica dos discos amostrais para determinação da densidade da madeira do tronco

Variável	Descrição	Posições Relativas
d_{Ba5}	Média aritmética dos cinco discos	0%-25%-50%-75%-100%
d_{Bb}	Valor do disco da base	0%
d_{Bm}	Valor do disco do meio	50%
d_{Bt}	Valor do disco do topo	100%
d_{Bbt}	Média aritmética dos valores da base e do topo	0%-100%
d_{Bbmt}	Média aritmética dos valores da base, meio e topo	0%-50%-100%
d_{Bpb}	Média ponderada dos valores da base e do topo	0%-100%
d_{Bpbmt}	Média ponderada dos valores da base, meio e topo	0%-50%-100%
d_{Bp5}	Média ponderada dos cinco discos	0%-25%-50%-75%-100%

Modelos de Predição da Densidade do Tronco

De forma análoga aos modelos de predição de teor de umidade, foi utilizado o seguinte modelo operacional para predição da densidade do tronco das árvores:

$$d_{BTi} = (0) + (1) d_{Bki} + \varepsilon_i \quad (3.5)$$

onde i é o índice referente às árvores, d_{BTi} é a densidade do tronco d_{Bki} é a densidade obtida por uma das técnicas de combinação da densidade básica do tronco apresentadas na tabela 3.2, e ε_i é um erro aleatório com distribuição Gaussiana, de média nula e variância constante.

Já o modelo de aproximação, que representa a *melhor aproximação linear* entre as variáveis foi definido como

$$d_{BTi} = \beta_0 + \beta_1 d_{Bki} + \varepsilon_i \quad (3.6)$$

onde β_0 e β_1 são os coeficientes estimados de modo a estabelecer a melhor relação linear entre as variáveis, e os demais termos como definidos acima.

3.2.7 Modelos para Predição da Biomassa do Tronco

Na determinação da biomassa do tronco das árvores pela técnica volumétrica, cada um das técnicas de combinação da densidade básica dos discos resulta numa medida diferente da densidade da madeira do tronco que, ao ser multiplicada pelo volume do tronco, gera uma predição diferente da biomassa. A medida de biomassa resultante é influenciada pelos erros de mensuração do volume e pelos erros de determinação da densidade do tronco, de forma que o comportamento das predições da biomassa será bastante complexo. É importante que a combinação de volume e densidade do tronco resultem numa boa predição da biomassa, o que não pode ser garantido apenas por uma boa predição da densidade.

Na literatura (BROWN; GILLESPIE; LUGO, 1989; BROWN, 1997) é comum que a biomassa do tronco seja obtida pelo produto do volume do tronco pela *densidade básica da espécie*. Para avaliar essa abordagem uma nova técnica de predição da biomassa foi acrescentada, onde a medida da densidade do tronco foi tomada como a densidade média das árvores da espécie (d_{Bsp}).

Para avaliar a predição da biomassa pelo método volumétrico, em função das diversas técnicas de combinação da densidade básica dos discos amostrais, utilizou-se os mesmos modelos operacionais e de aproximação utilizados no estudo do teor de umidade e da densidade do tronco. O modelo operacional para biomassa foi

$$b_{Gi} = (0) + (1) (v_i d_{Bki}) + \varepsilon_i \quad (3.7)$$

onde i é o índice referente às árvores, b_{Gi} biomassa pela técnica gravimétrica (kg); v_i volume do tronco (m^3); d_{Bki} é a densidade obtida por uma das técnicas de combinação da densidade básica do tronco apresentadas na tabela 3.2, e ε_i é um erro aleatório com distribuição Gaussiana, de média nula e variância heteroscedástica. Já o modelo de aproximação foi o seguinte:

$$b_{Gi} = \beta_0 + \beta_1 (v_i d_{Bki}) + \varepsilon_i \quad (3.8)$$

onde β_0 e β_1 são os coeficientes estimados de modo a estabelecer a melhor relação linear entre as variáveis, e os demais termos como definidos acima.

Os modelos de determinação da biomassa do tronco assumem erros aleatórios heteroscedásticos por duas razões. Primeiramente, a pressuposição de homocedasticidade é irreal, uma vez que a biomassa de árvores individuais sempre apresenta um comportamento heteroscedástico, com a variância aumentando com o tamanho da árvore. Em segundo lugar, a transformação da variável resposta (biomassa) tornaria muito complexa a análise e interpretação dos modelos, uma vez que a modelagem seria numa escala distinta da original em que a biomassa foi medida. Para ajustar os modelos, assumiu-se que a variância do erro aleatório segue uma função de potência da variável preditora ($v_i d_{Bki}$).

3.2.8 Ajuste e Avaliação dos Modelos de Predição

Todos os modelos foram ajustados pelo Método da Máxima Verossimilhança (MOOD; GRAYBILL; BOES; 1974) utilizando um algoritmo de otimização numérica em todos os casos. Nos casos em as estimativas de máxima verossimilhança são coincidentes com as estimativas de quadrados mínimos, ou seja, nos modelos com erro Gaussiano homoscedástico, as estimativas obtidas pelo algoritmo de otimização numérica foram idênticas às estimativas de quadrados mínimos. O ajuste foi implementado no pacote “stats4” do software estatístico “R”, utilizando a função de ajuste “mle” (*maximum likelihood estimation*), que utiliza por default a função “optim” para otimização numérica.

Akaike (1973 apud BURNHAM e ANDERSON 2002) encontrou, através da relação entre a distância entre dois modelos de Kulbach-Leibler e a função de máxima log-verossimilhança de Fisher, uma simples e efetiva metodologia para seleção de modelos para análise de dados empíricos, denominado de critério de informação de Akaike (AIC). Este critério estima a discrepância geral relativa entre um modelo de aproximação e o modelo real (hipotético) que gerou os dados. Desta forma, quanto menor valor do AIC de um modelo, menor sua discrepância geral estimada e, por consequência, menor sua distância relativa para o modelo real (hipotético). Segundo Sakamoto Sakamoto; Ishiguro e Kitagawa (1986) a fórmula para o cálculo do AIC para o modelo homocedástico e com erros gaussianos é:

$$AIC = -2\log(L(\hat{\sigma}^2)) + 2p \quad (3.9)$$

onde n é o tamanho da amostra, p é o número de parâmetros (incluindo os coeficientes do modelo e $\hat{\sigma}^2$) e $\hat{\sigma}^2 = \frac{\sum(\hat{\epsilon}^2)}{n}$.

Nesse estudo, tomou-se o AIC do modelo operacional (AIC_0) como sendo uma medida da *discrepância geral*, ou seja uma medida da qualidade geral da predição. Diferenças de AIC superiores a 2 são consideradas, segundo Burnham e Anderson (1984), diferenças relevantes e esse valor limite foi utilizado na avaliação de modelos concorrentes. Já o AIC do modelo de aproximação (AIC_a) foi tomado como uma medida de *discrepância de aproximação*, enquanto que a diferença desses dois valores de AIC foi tomada como uma medida da *discrepância de estimação* (AIC_e). Dessa forma, a relação entre as medidas de discrepância permanece válida:

$$AIC_0 = AIC_a + AIC_e \quad (3.10)$$

Além do AIC, erro padrão da estimativa foi usado como critério de análise e seleção dos modelos operacionais. Esse é definido em Draper e Smith (1998) como:

$$s_{\hat{y}} = \frac{\sqrt{\sum(y_i - \hat{y}_i)^2}}{n} \quad (3.11)$$

onde $s_{\hat{y}}$ é o erro padrão da estimativa, y_i é o valor observado, \hat{y}_i é valor predito e n é o tamanho da amostra. Neste critério quanto seu menor, melhor será a qualidade da predição do modelo.

No caso dos modelos de biomassa, os erros foram modelados como heteroscedásticos, de forma que a estimativa acima do erro padrão não pode ser implementada, pois o erro padrão muda com o valor da biomassa a ser predito. Assim, os valores de erro padrão da estimativa reportados para esses modelos se referem a valor estimado do erro padrão segundo a função que modela a variância dos erros do modelo, tomando-se como ponto de estimação o valor médio da biomassa na amostra (150 kg).

Além da utilização do AIC e do erro padrão da estimativa, foi realizada uma análise gráfica do comportamento dos modelos gerados, no que diz respeito a presença ou ausência de erro sistemático de predição em função da estimação de coeficientes dos modelos ($\hat{\beta}_0$ e $\hat{\beta}_1$). Através dessa análise gráfica foi possível revelar tendência de sub ou super estimação dos diferentes modelos quando aplicados na predição.

A comparação das técnicas também se deu através de gráficos de dispersão do teor de umidade do tronco contra a combinação dos teores de umidade dos discos. Nesses gráficos, foram adicionadas a linha do modelo operacional e a linha do modelo de aproximação correspondente, para que fosse possível observar os aspectos discrepantes entre a relação sugerida pelo modelo de aproximação e a relação fixa do modelo operacional.

3.3 Resultados e Discussão

3.3.1 Determinação do Teor de Umidade

Os modelos para todas as técnicas de combinação do teor de umidade dos discos não apresentaram problemas quanto a heterocedasticidade e nem quanto à normalidade dos erros. Os modelos operacionais apresentaram um desvio sistemático em relação os modelos de aproximação em todas as técnicas de combinações dos teores de umidade dos discos (Figuras 3.1, 3.2, 3.3, 3.4, 3.5). Esse desvio sistemático diferiu em grau nas diferentes técnicas, mas sempre apontou para uma tendência de sub-estimativa quando o teor de umidade era baixo e de super-estimativas quando o teor de umidade era elevado.

Nota-se que ao se usar apenas o disco da base ou o disco do topo para determinar o teor de umidade do fuste, o desvio de do modelo operacional (linha tracejada) é bem marcante (Figuras 3.1 e 3.3). O desvio se torna mais ameno quando se utiliza o disco da porção média do fuste e também se usa os discos da base e do topo combinados (Figura 3.2 e 3.4). O desvio se torna ainda mais tênue e pouco marcante, ao se utilizar a combinação de três discos (base, meio e topo) na determinação do teor de umidade do fuste (Figura 3.5).

Os critérios de avaliação dos modelos das cinco técnicas de combinação do teor de umidade dos discos são apresentados na tabela 3.3. A partir da observação dos valores da discrepância geral entre o modelo operacional e o modelo de aproximação é possível afirmar que a melhor técnica é a que utiliza a combinação dos discos da base, meio e topo, pois apresenta o menor valor de discrepância geral ($AIC_0 = -1225,951$). Além disso, os valores de discrepância de aproximação, discrepância na estimação e o valor do erro padrão da estimativa dessa técnica são inferiores às das demais, também indicam sua melhor adequação para a determinação do teor de umidade do fuste no conjunto de dados em questão.

Quanto às demais técnicas, pode-se observar que os valores das discrepâncias de estimação

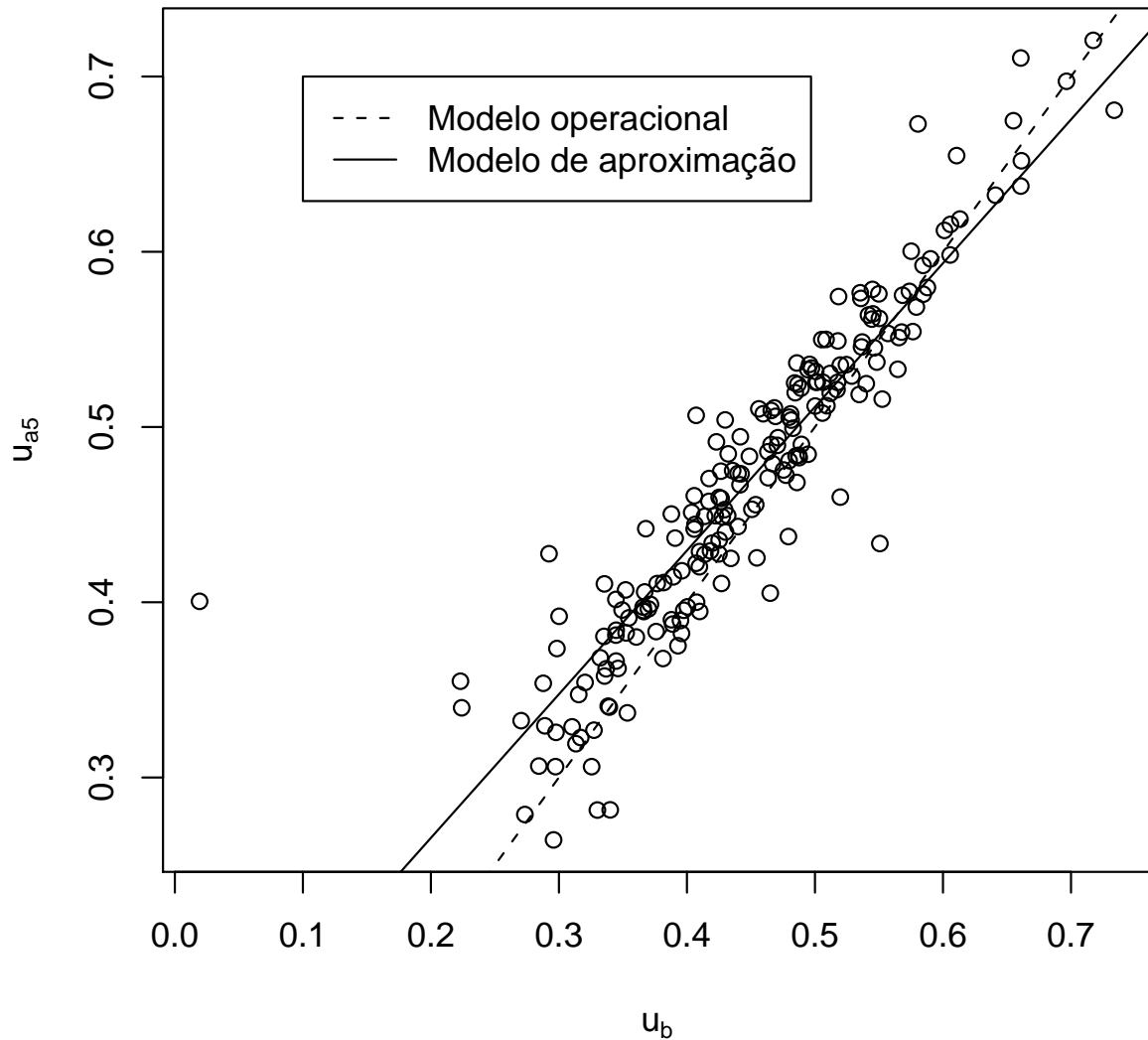


Figura 3.1 – Gráfico de dispersão do teor de umidade do tronco (u) e o teor de umidade do disco da base (u_b), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

classificaram os modelo de acordo com o que foi observado na análise gráfica. Sendo assim, as técnicas que apresentaram desvios mais marcantes graficamente, apresentaram valores maiores de discrepância de estimação (disco da base e disco do topo). Enquanto que as técnicas que obtiveram desvios mais amenos apresentam discrepância de estimação inferiores (disco do meio e combinação disco da base e do topo).

Com relação à discrepância de aproximação, é possível notar também, uma correspondência gráfica. Ao se observar as Figuras 3.1, 3.2, 3.3, 3.4 e 3.5, é possível notar que a dispersão dos

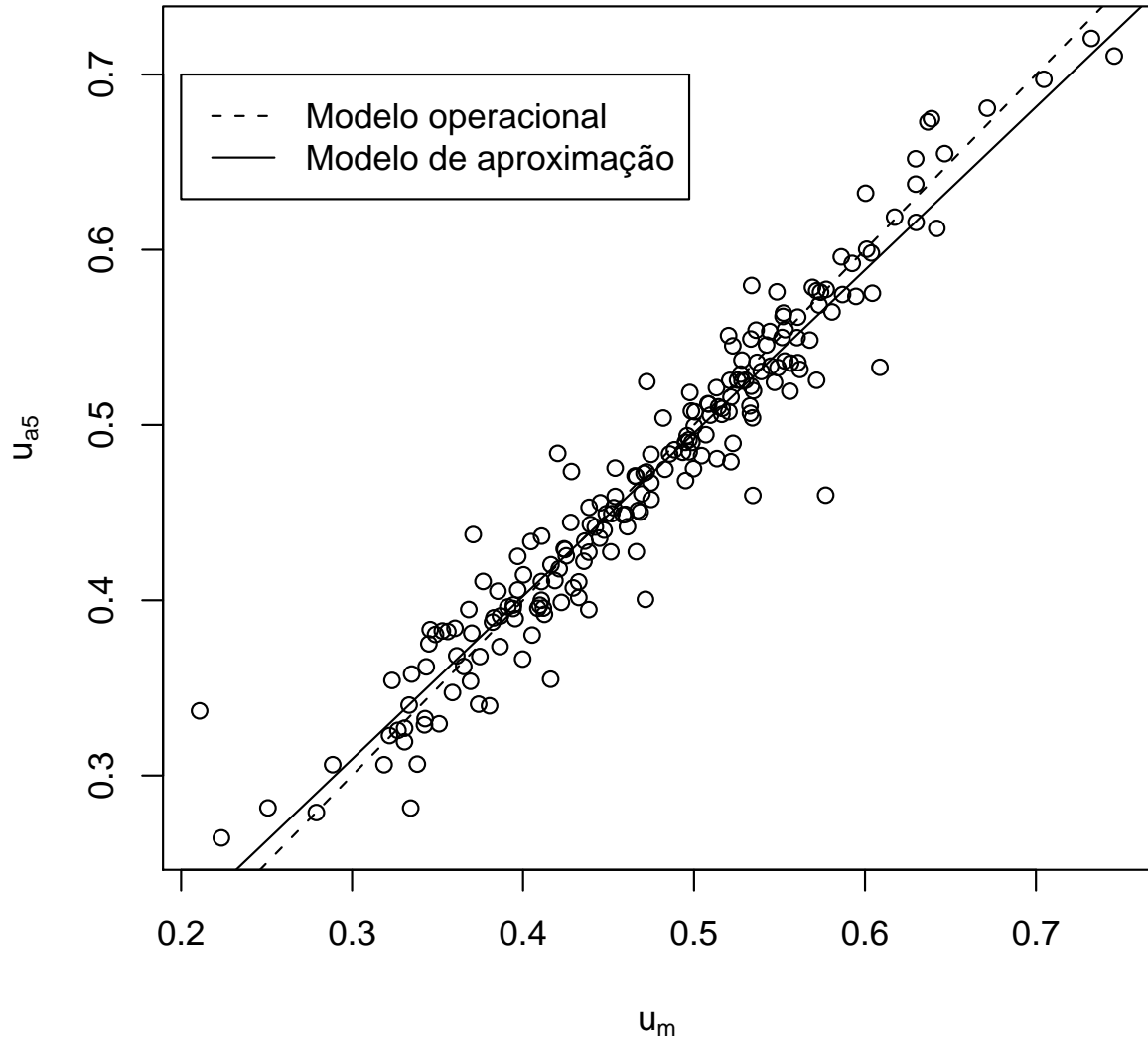


Figura 3.2 – Gráfico de dispersão do teor de umidade do tronco (u) e o teor de umidade do disco do meio (u_m), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

dados é menor quanto menor a discrepância na aproximação na Tabela 3.3. Para os tratamentos em questão essas estimativas foram bem distintas e demonstraram que quanto maior o número de discos, melhor a aproximação do modelo em relação teor de umidade com base nos cinco discos amostrados.

A tendência do maior número de discos produzir previsões mais precisas do teor de umidade do tronco é ainda mais acentuada nos valores de discrepância geral, que é bem distinta entre as diferentes técnicas de combinação dos teores de umidade dos discos, sempre decrescendo com o

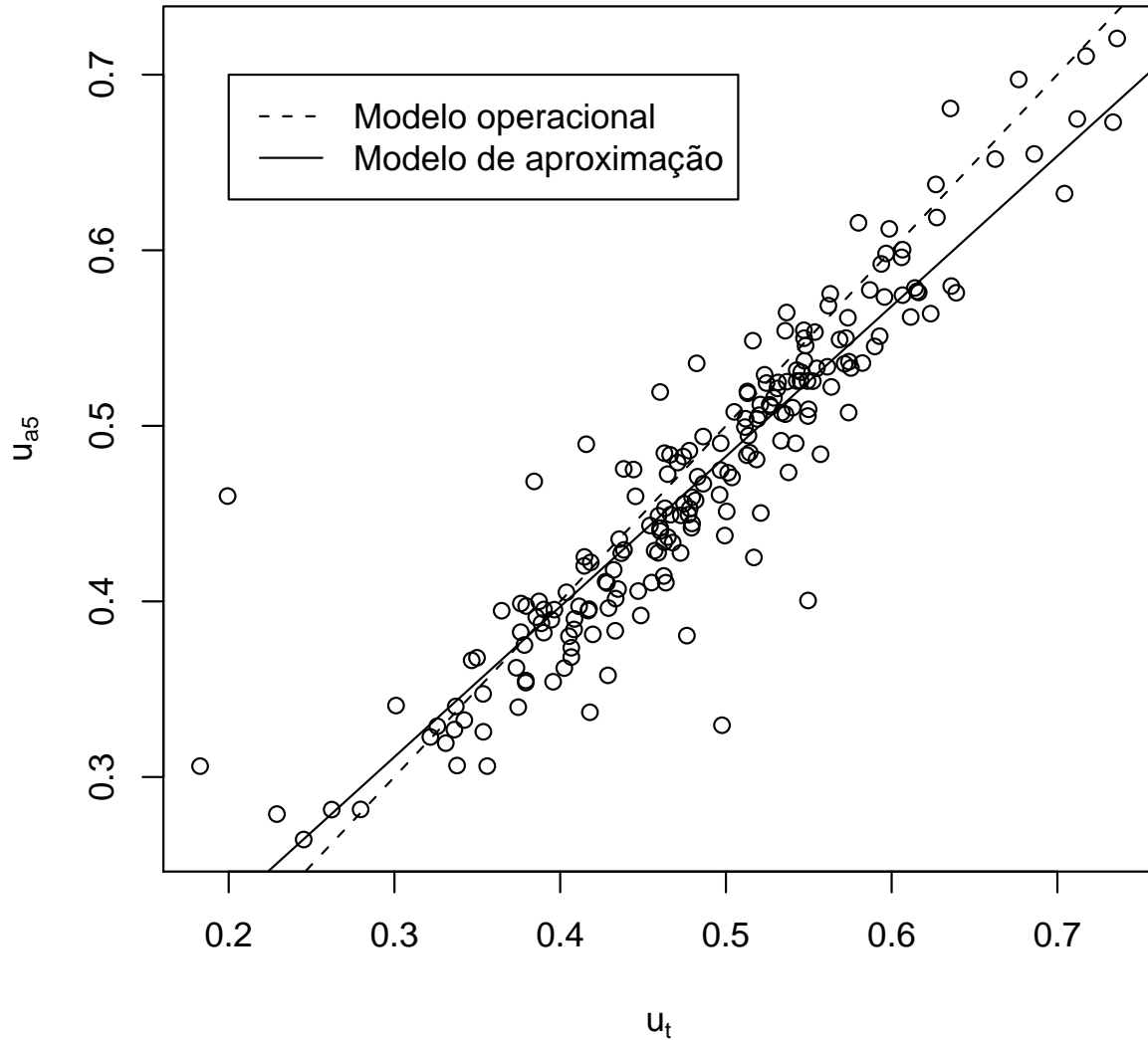


Figura 3.3 – Gráfico de dispersão do teor de umidade do tronco (u) e o teor de umidade do disco do topo (u_t), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

aumento no número de discos utilizados (Tabela 3.3).

3.3.2 Determinação da Densidade da Madeira do Tronco

Os modelos referentes à determinação da densidade da madeira do tronco, a partir das diversas técnicas de combinação das densidades básicas dos discos amostrados, não apresentaram problemas quanto a heterocedasticidade e nem quanto á normalidade dos erros. Entretanto, todas as nove técnicas de combinação apresentaram de uma forma geral, desvio sistemático no modelo opera-

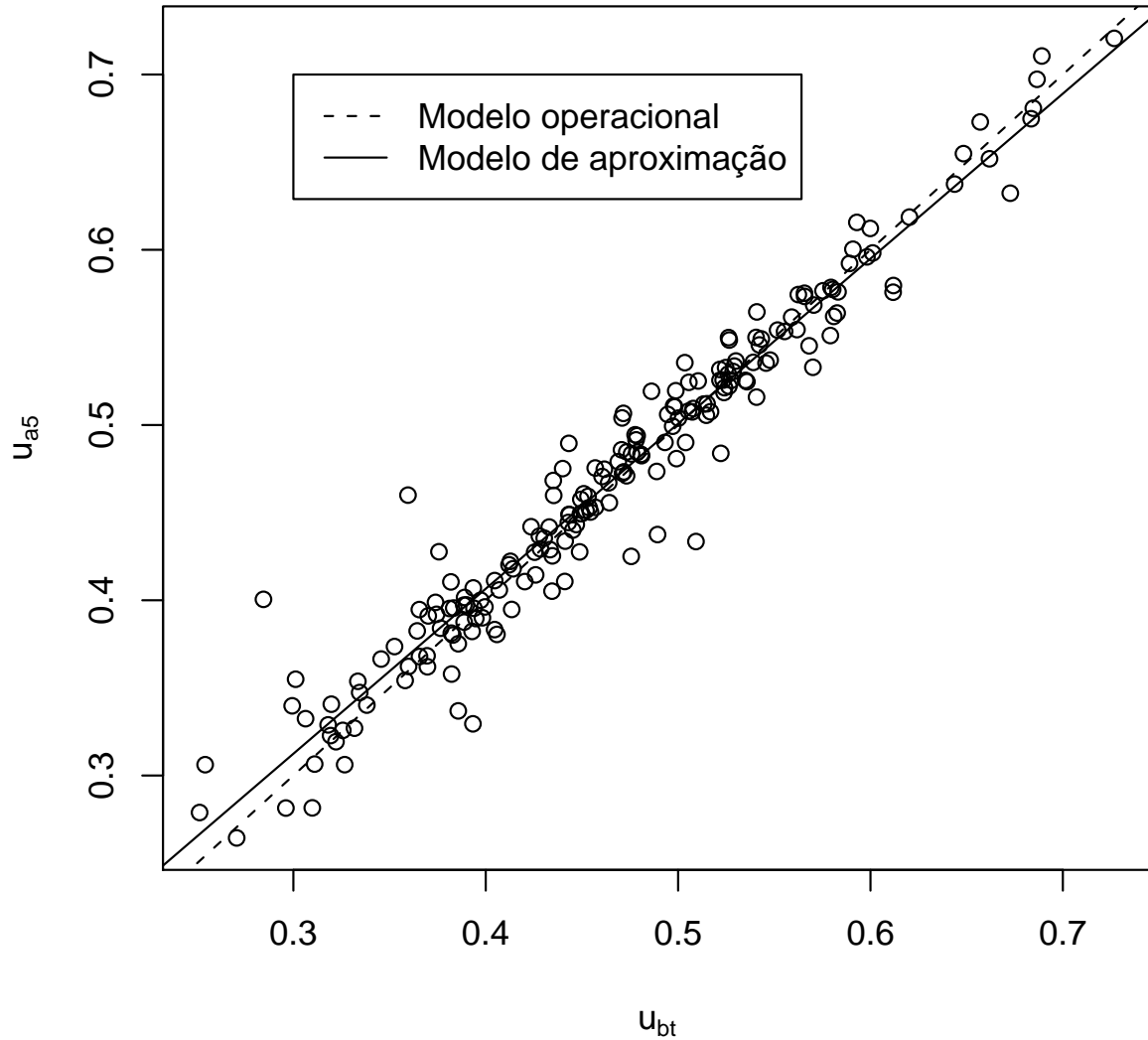


Figura 3.4 – Gráfico de dispersão do teor de umidade do tronco (u) e a média dos teores de umidade dos discos da base e do topo (u_{bt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

cional que cresce com o aumento da densidade. A discrepância geral dos modelos diferiu em grau nas diferentes técnicas, mas sempre apontou para uma tendência de super-estimativas para qualquer valor de densidade do tronco (Figuras 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13 e 3.14). A única exceção foi a técnica que utiliza o disco do topo (Figura 3.8), em que se observa menor desvio gráfico entre o modelo operacional e o modelo de aproximação. Diferentemente dos demais, essa técnica possui tendência de sub-estimação dos valores pequenos de densidade e de super-estimação

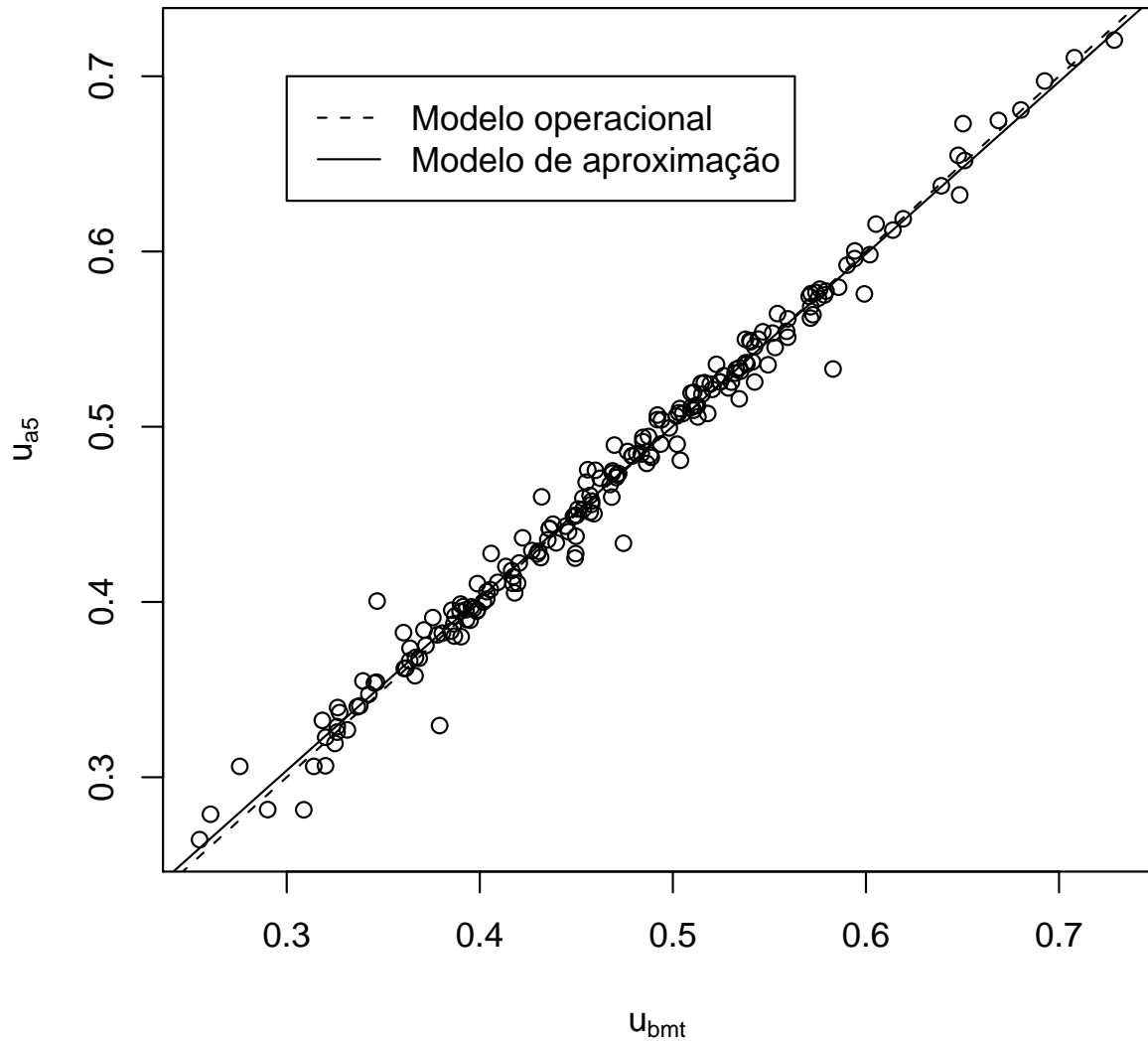


Figura 3.5 – Gráfico de dispersão do teor de umidade do tronco (u) e a média dos teores de umidade dos discos da base, do meio e do topo (u_{bmt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

Tabela 3.3 – Valores das discrepâncias e do erro padrão da estimativa para as técnicas de combinação do teor de umidade dos discos amostrados para determinação do teor de umidade do tronco

Variável	Discos Utilizados	Discrepâncias			Erro Padrão da Estimativa
		Geral	de Aproximação	de Estimação	
u_b	base	-663,526	-749,803	86,277	0,0366
u_t	topo	-854,040	-908,452	54,412	0,0357
u_m	meio	-747,026	-759,019	11,993	0,0246
u_{bt}	base e topo	-961,646	-974,098	12,452	0,0209
u_{bmt}	base, meio e topo	-1124,433	-1225,951	1,518	0,0111

dos valores maiores.

Os critérios de avaliação das técnicas de combinação das densidades básicas dos discos para determinação da densidade do tronco são apresentadas na tabela 3.4.

Tabela 3.4 – Valores das discrepâncias e do erro padrão da estimativa para as técnicas de combinação das densidades básicas dos discos amostrados para determinação da densidade do tronco

Variável	Discos	Média	Discrepâncias			Erro Padrão da Estimativa
			Geral	de Aproximação	de Estimação	
d_{Ba5}	todos	aritmética	-532,1344	-612,1588	80,0215	0,06264
d_{Bb}	base	—	-377,5645	-585,4483	207,8848	0,09255
d_{Bm}	meio	—	-502,4075	-590,3547	87,9472	0,06360
d_{Bt}	topo	—	-529,6587	-541,5009	11,8422	0,06303
d_{Bbt}	base e topo	aritmética	-498,1786	-594,7873	96,6087	0,06825
d_{Bbmt}	base, meio e topo	aritmética	-520,9801	-606,8153	85,8352	0,06443
d_{Bpbt}	base e topo	ponderada	-438,664	-599,2415	160,5775	0,07932
d_{Bpbmt}	base, meio e topo	ponderada	-475,4383	-613,4345	137,9962	0,07228
d_{Bp5}	todos	aritmética	-499,3778	-619,5832	118,2056	0,06804

A técnica de combinação com menor estimativa de discrepância geral é a que determina a densidade do tronco pela média aritmética das densidades básica dos cinco discos (d_{Ba5} ; $-532,1344$). Imediatamente abaixo dessa técnica está a que utiliza simplesmente a densidade básica do disco do topo (d_{Bt} ; $-529,6587$), mas como a diferença das discrepâncias gerais entre as duas técnicas é superior a 2, o modelo com base nos cinco discos deve ser considerado superior (BURHAM, ANDERSON, 2002). Mas a utilização da densidade do disco do topo resultou na menor discrepância de estimativa (11,8422), sendo essa discrepância muito menor que todas as demais técnicas estudadas. Essa técnica só não foi melhor que a técnica da média aritmética dos cinco discos porque apresentou a maior discrepância de aproximação ($-759,0193$), sendo essa discrepância a mais elevada dentre as técnicas que utilizaram apenas um disco para determinação da densidade do tronco.

A discrepância de aproximação tende, a ser menor quanto maior o número de discos combi-

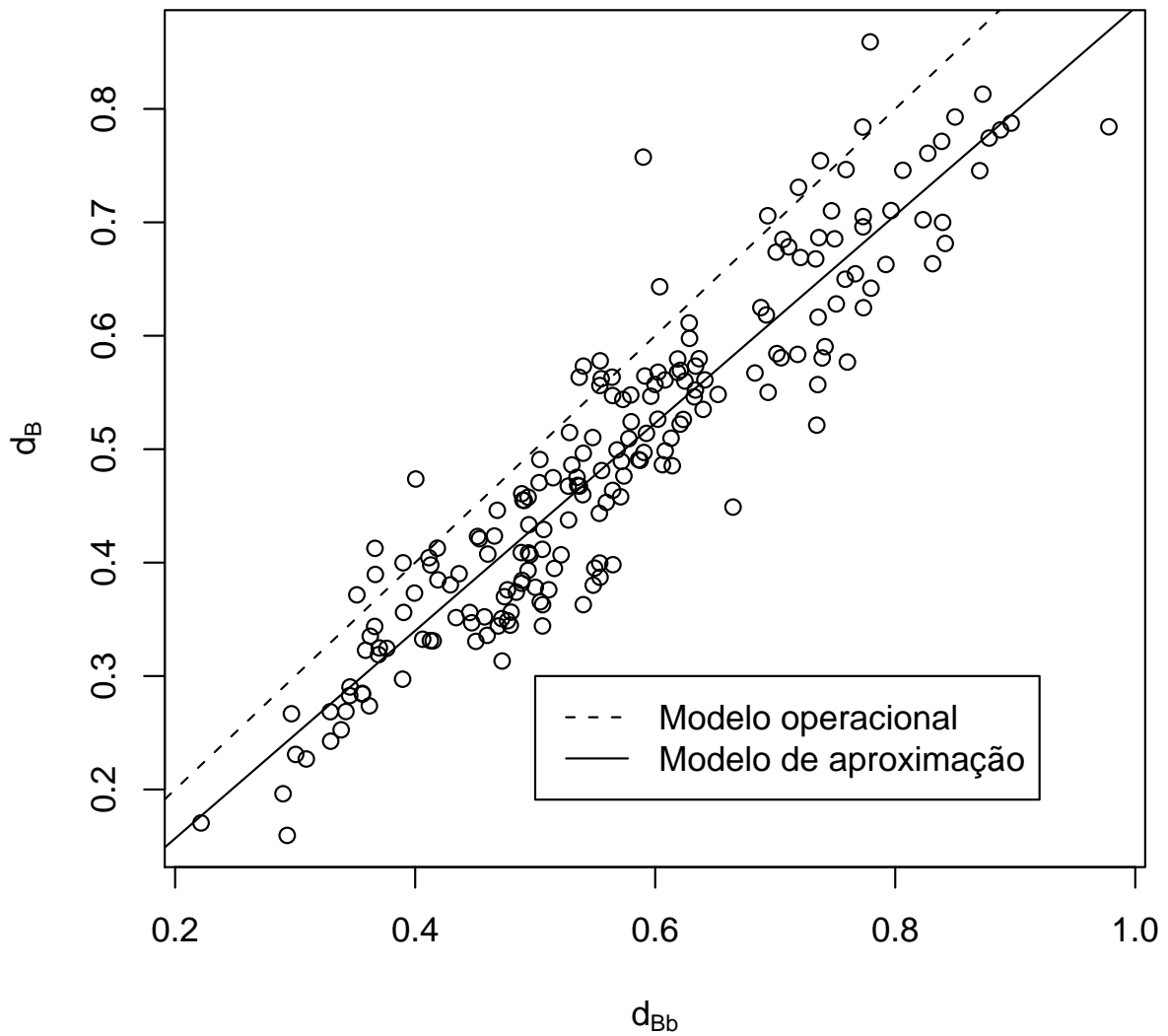


Figura 3.6 – Gráfico de dispersão da densidade da madeira do tronco e a densidade obtida pelo disco da base (d_{Bb}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

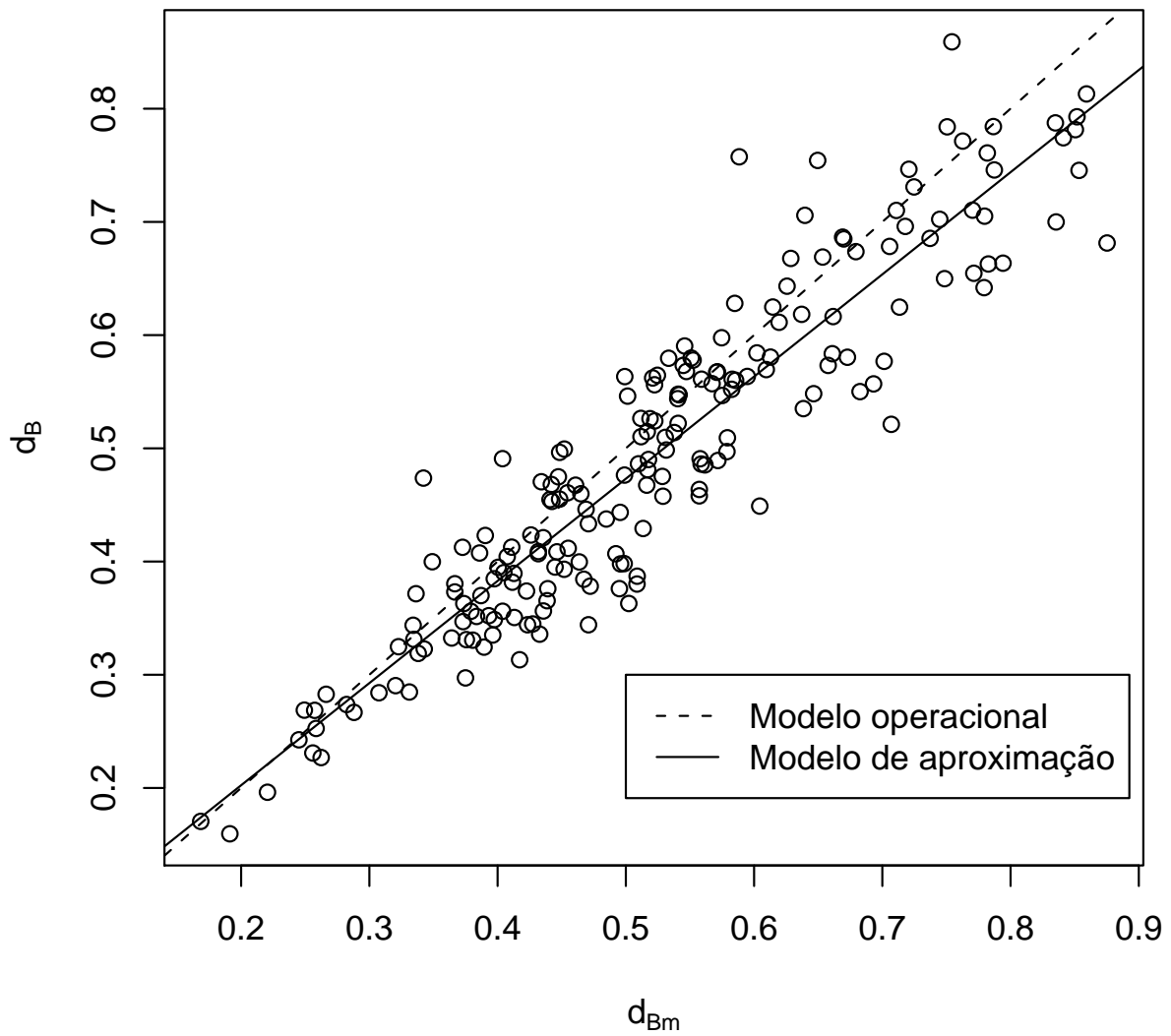


Figura 3.7 – Gráfico de dispersão da densidade da madeira do tronco e a densidade obtida pelo disco do meio (d_{Bm}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

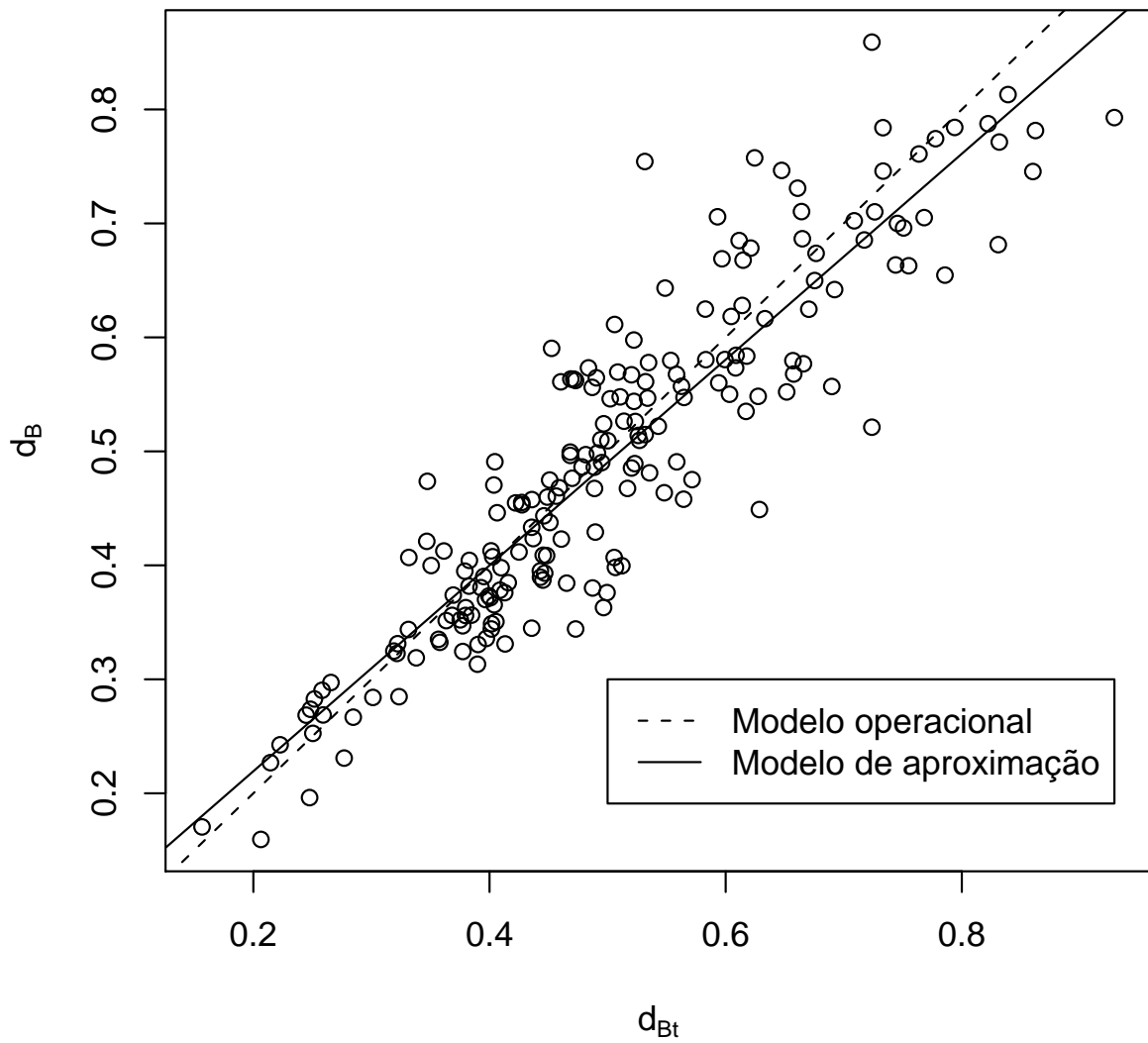


Figura 3.8 – Gráfico de dispersão da densidade da madeira do tronco e a densidade obtida pelo disco do topo (d_{Bt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

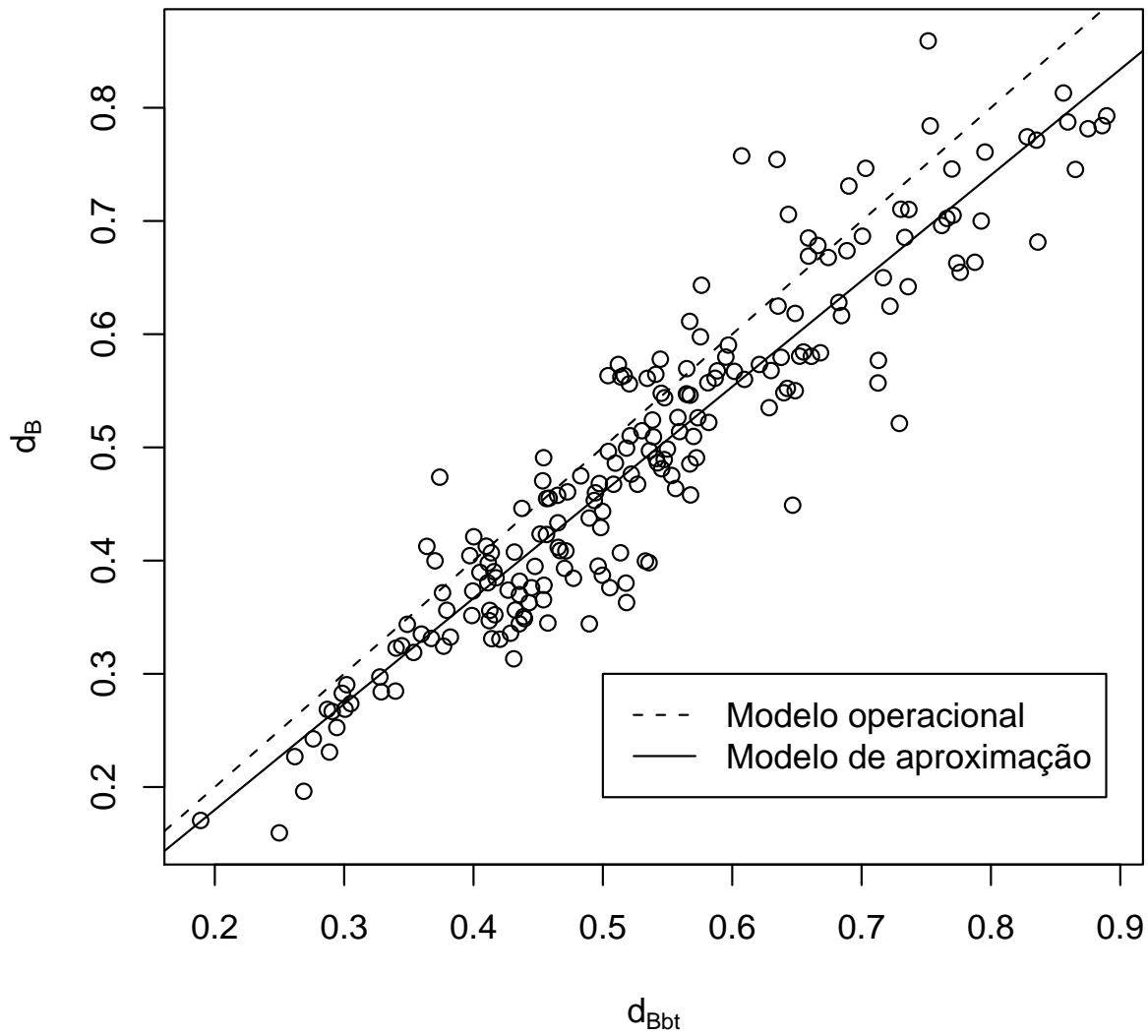


Figura 3.9 – Gráfico de dispersão da densidade da madeira do tronco e a média aritmética das densidade dos discos da base e do topo (d_{Bbt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

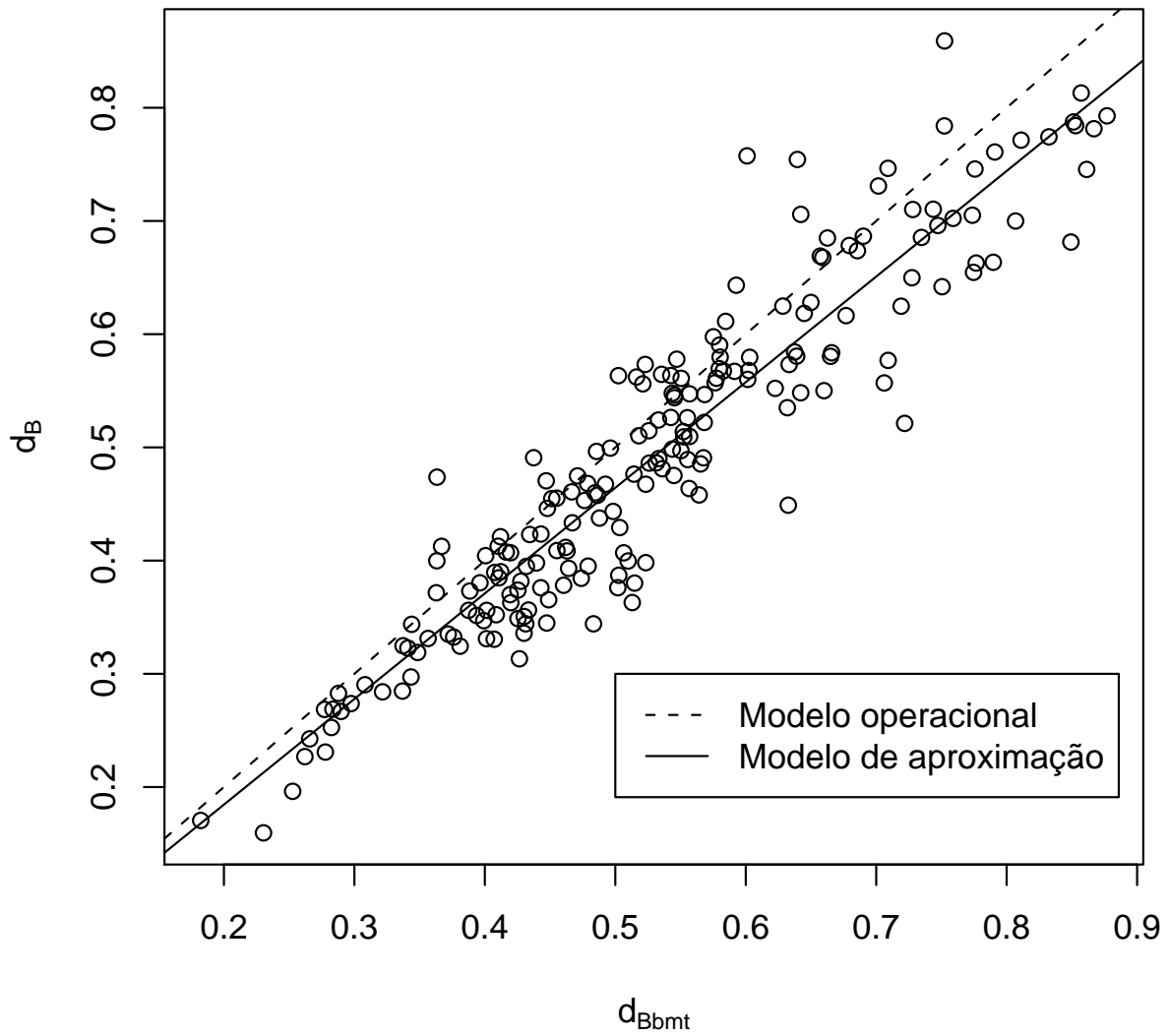


Figura 3.10 – Gráfico de dispersão da densidade da madeira do tronco e a média aritmética das densidade dos discos da base, do meio e do topo (d_{Bbmt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

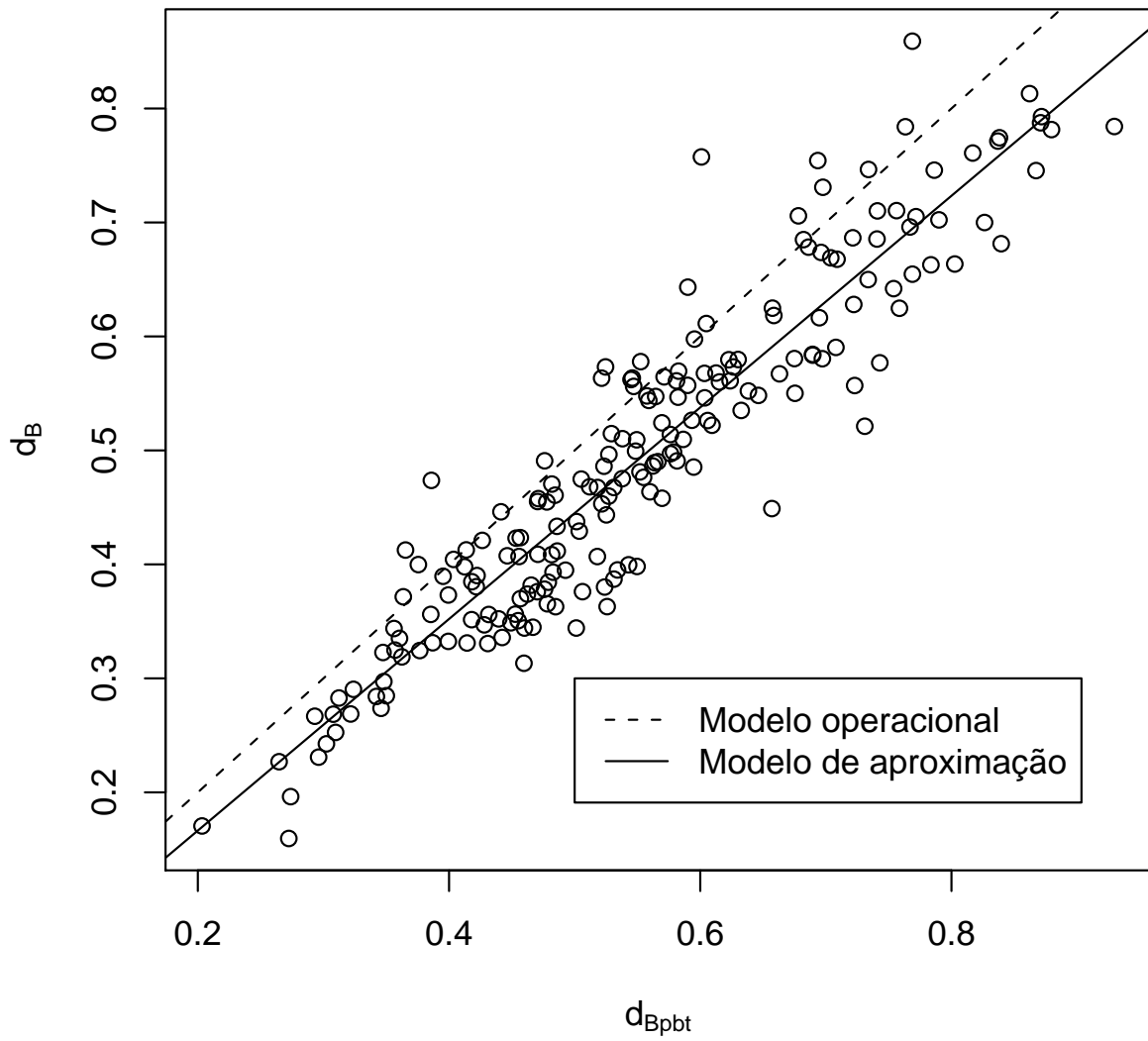


Figura 3.11 – Gráfico de dispersão da densidade da madeira do tronco e a média ponderada das densidade dos discos da base e do topo (d_{Bpbt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

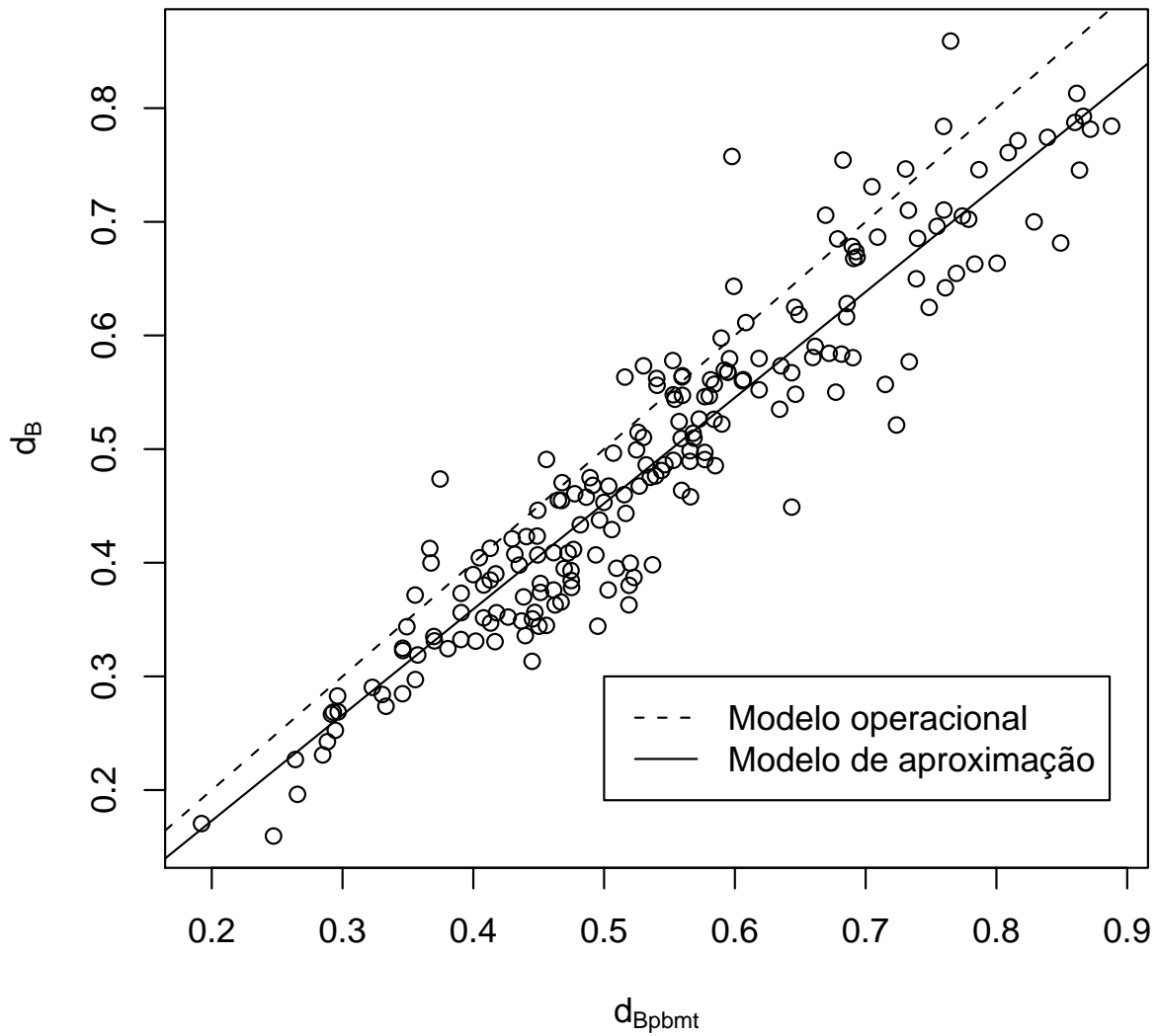


Figura 3.12 – Gráfico de dispersão da densidade da madeira do tronco e a média ponderada das densidade dos discos da base, do meio e do topo (d_{Bpbmt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

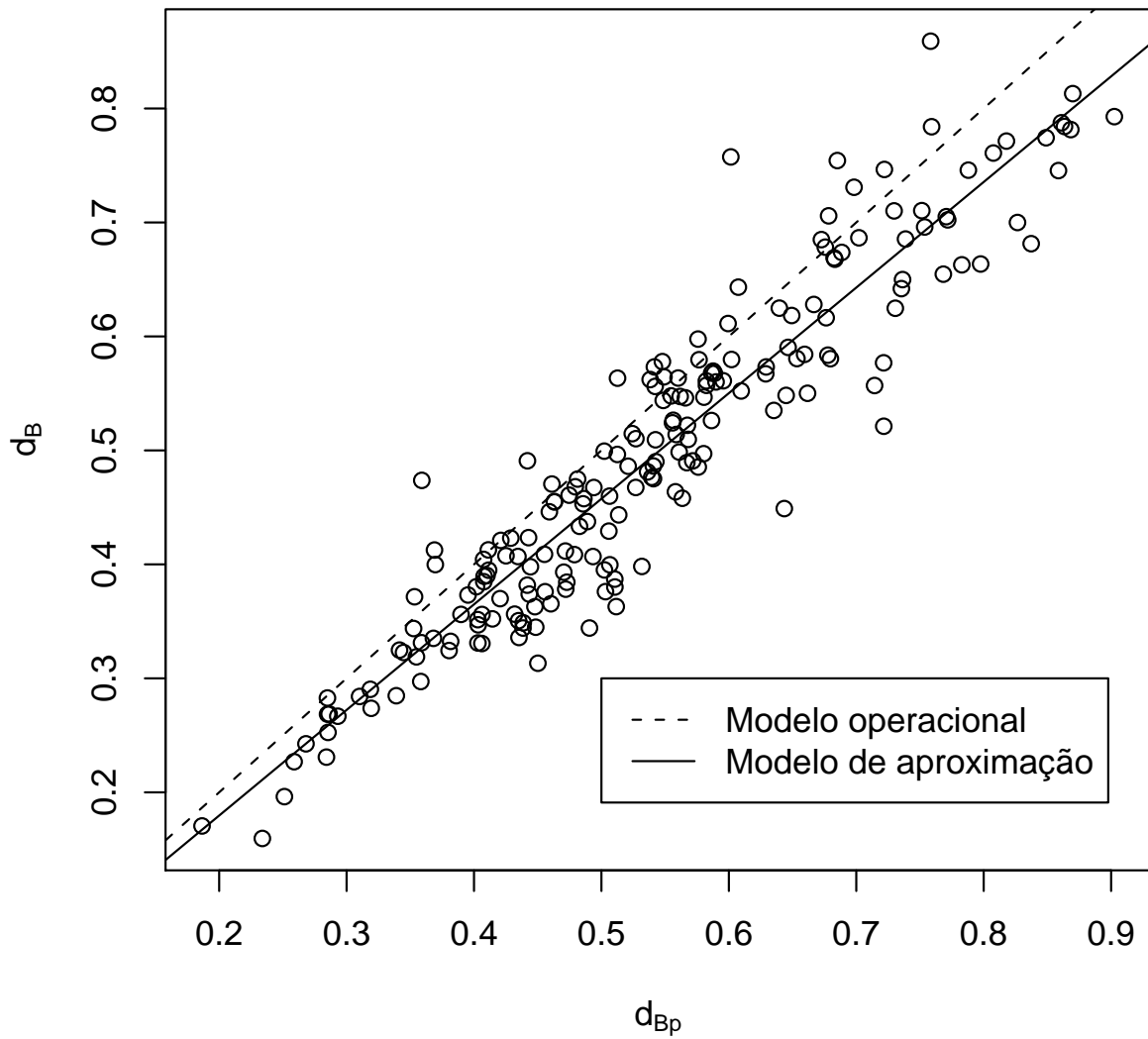


Figura 3.13 – Gráfico de dispersão da densidade da madeira do tronco e a média ponderada das densidade de todos os cinco discos amostrados (d_{Bp5}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

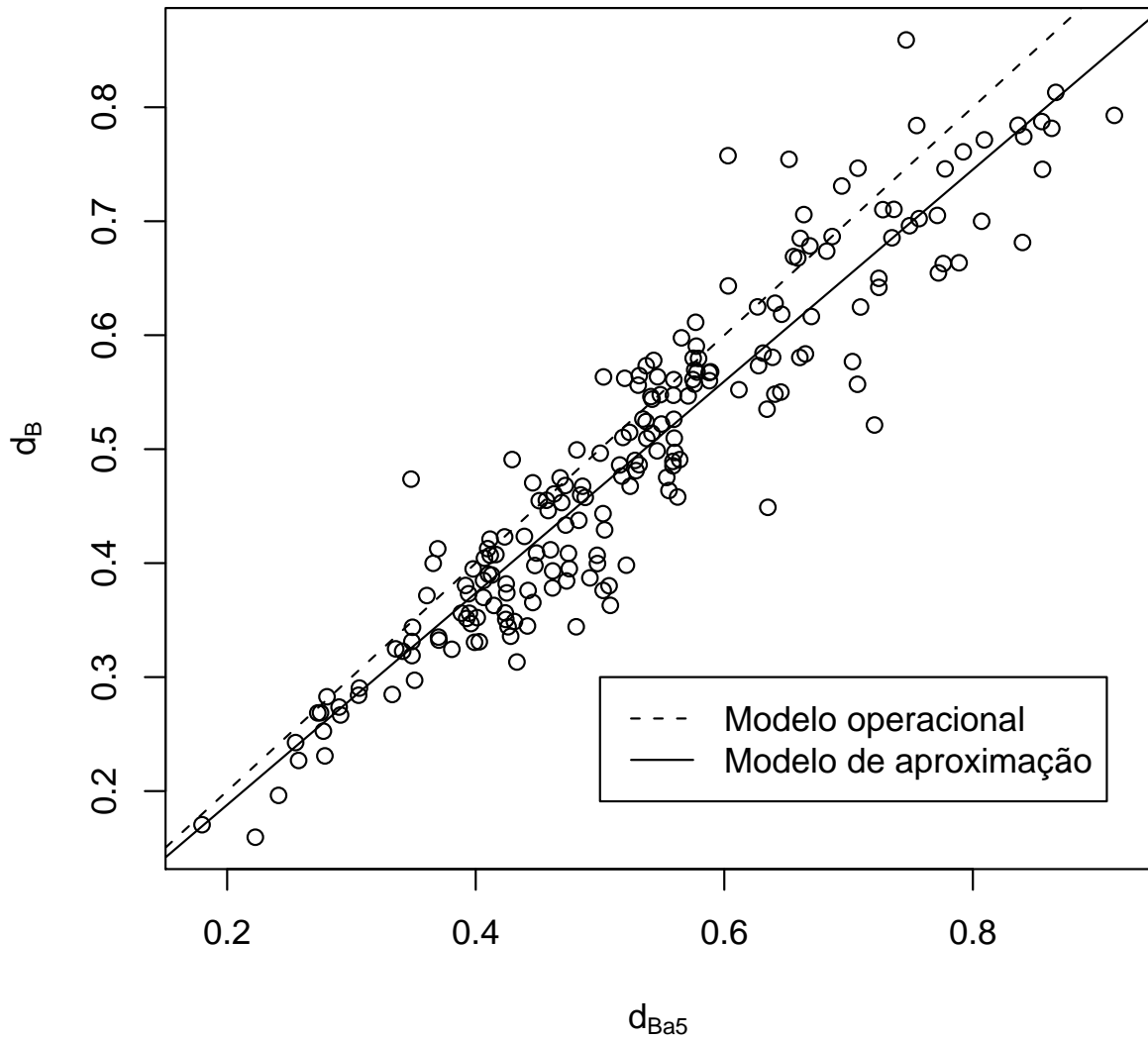


Figura 3.14 – Gráfico de dispersão da densidade da madeira do tronco e a média aritmética das densidade de todos os cinco discos amostrados (d_{Ba5}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

nados para determinação da densidade do tronco. Além disso, para um mesmo número de discos, o uso da média ponderada reduziu sensivelmente a discrepância de aproximação se comparada a média aritmética. Tanto que observando os resultados apresentados na tabela 3.4 nota-se que o melhor tratamento, no que se refere à aproximação, é justamente o que usa a média ponderada dos cinco discos amostrais.

Em contrapartida, o uso da média ponderada nos tratamentos provocou uma inflação muito grande na discrepância de estimação, tornando-os piores neste aspecto que os tratamentos que usaram a média aritmética. No balanço geral, pode-se concluir que o uso da média aritmética se mostrou superior ao uso da média ponderada, já que os valores da discrepância geral foram menores para os tratamentos que utilizaram média aritmética.

Quanto ao erro padrão da estimativa este classificou os tratamentos de acordo com o que foi observado na discrepância geral. Desta maneira, tratamentos com menor discrepância geral apresentaram menores valores de erro padrão.

Ao contrário do que foi observado na determinação do teor de umidade, a tendência da combinação de um maior número de discos amostrais resultar em menor discrepância geral não foi tão acentuada na determinação da densidade do tronco. Isso corrobora com a hipótese de que além da incerteza envolvida no método de obtenção da densidade do tronco a partir das densidades básicas dos discos, o processo analítico de determinação da densidade básica de cada disco é mais complexo e, portanto, mais sujeito a erros, quando comparado ao processo de determinação do teor de umidade.

3.3.3 Determinação da Biomassa do Tronco

Ao se utilizar a técnica volumétrica de determinação da biomassa do tronco, acrescenta-se aos erros de determinação da densidade do tronco os erros de determinação do volume do tronco, uma vez que a biomassa é obtida pelo produto dessas duas grandezas. É necessário, portanto, analisar como as técnicas de combinação da densidade básica dos discos para determinação da densidade do tronco influenciam a determinação da biomassa do tronco.

Na análise gráfica, a maioria das técnicas de determinação da densidade do tronco apresentaram desvio sistemático do modelo operacional em relação ao modelo de aproximação. Esse desvio diferiu em grau nos diferentes técnicas, apontando para a sub-estimação da biomassa do tronco em todas as técnicas (Figuras 3.15, 3.16, 3.17, 3.18, 3.19, 3.20, 3.21, 3.22 e 3.23), exceto para a técnica que utiliza a densidade média da espécie, que apresentou uma tendência de super-estimação da biomassa (Figura 3.24).

No caso da técnica que usa o disco do topo e a que usa a média aritmética dos discos da base e do topo, as linhas do modelo de aproximação e do modelo operacional são praticamente coincidentes, não sendo possível afirmar nada a respeito de desvio sistemático do modelo operacional em relação

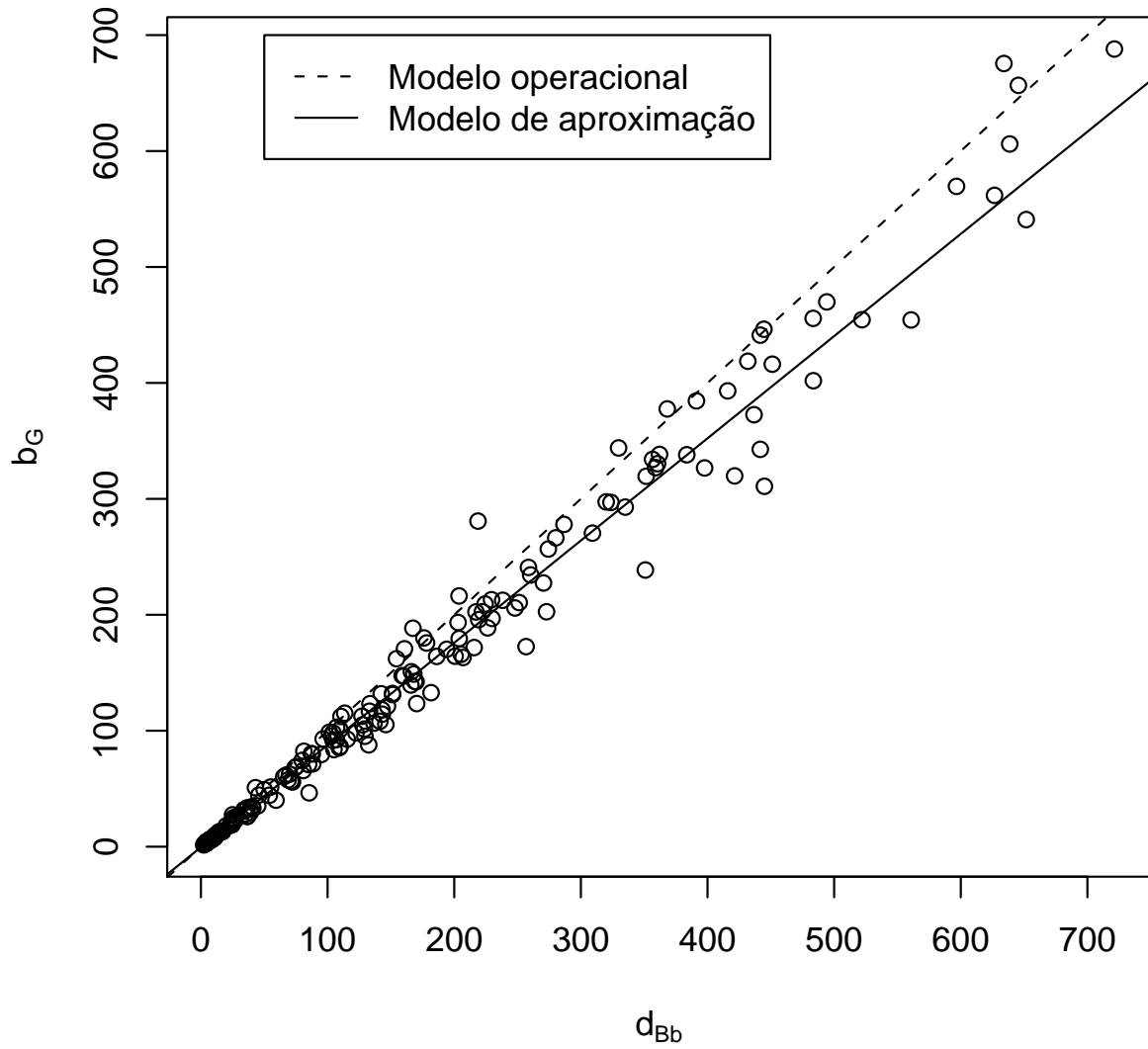


Figura 3.15 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pelo disco da base (d_{Bb}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

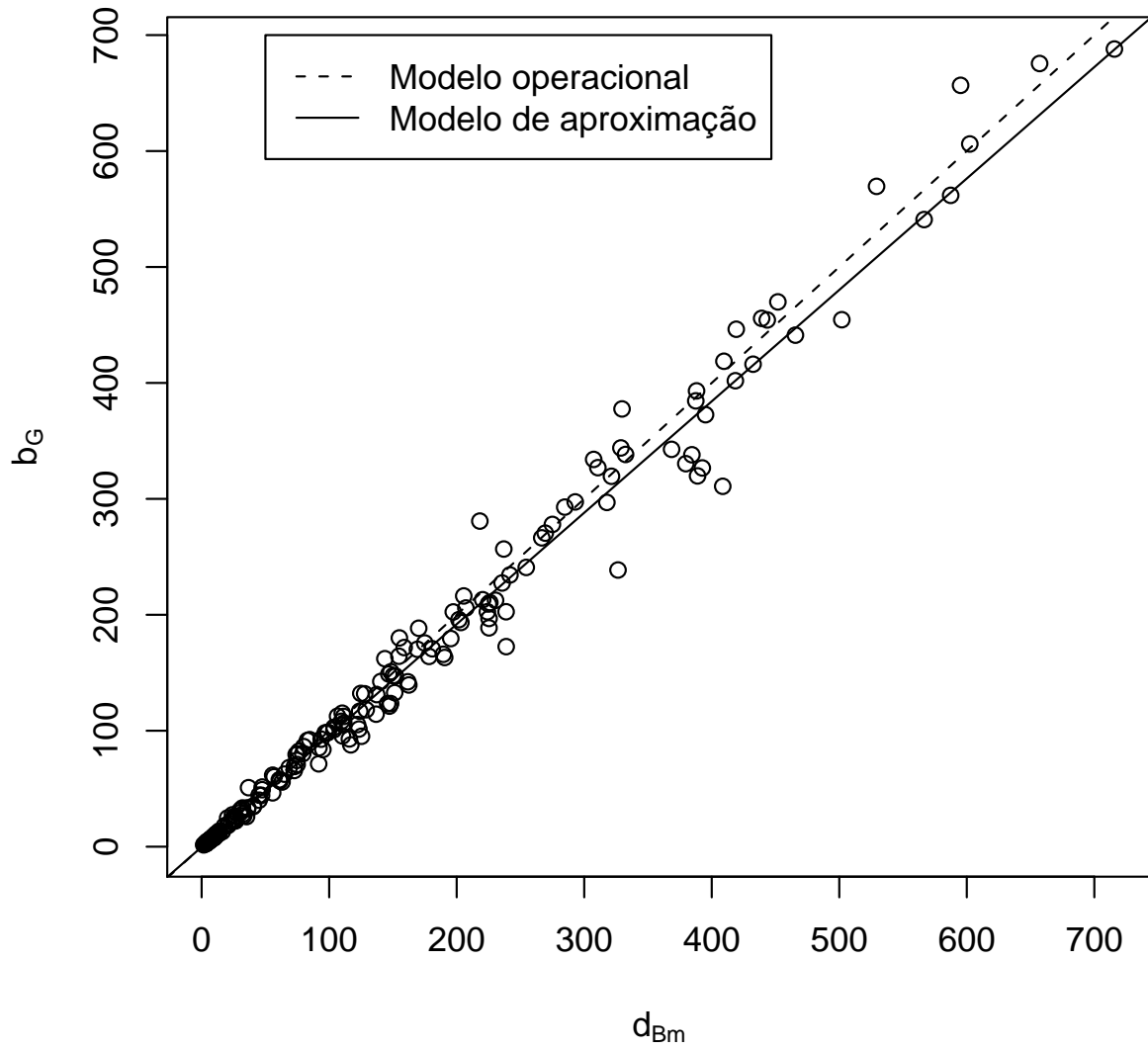


Figura 3.16 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pelo disco do meio (d_{Bm}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

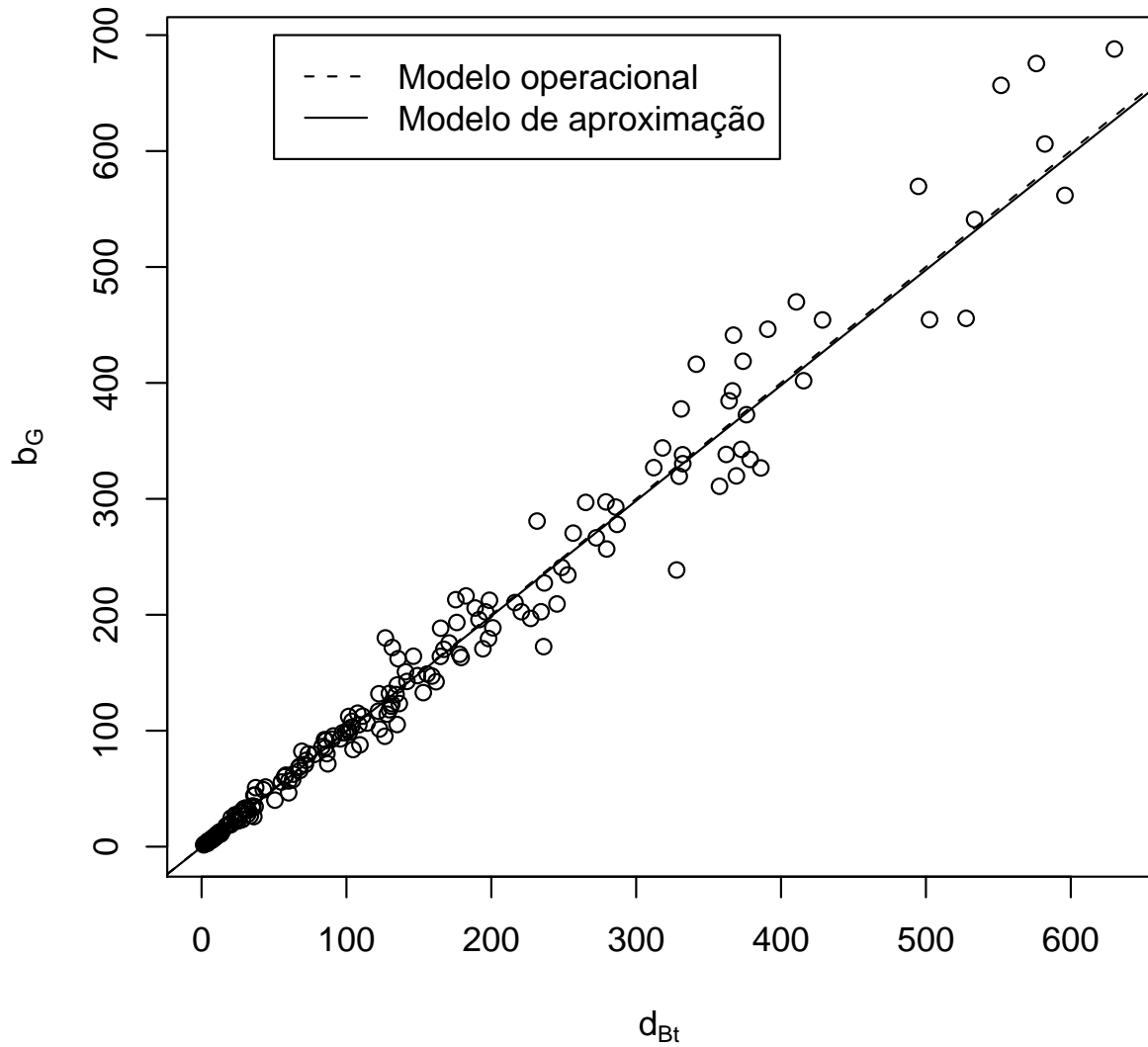


Figura 3.17 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pelo disco do topo (d_{Bt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

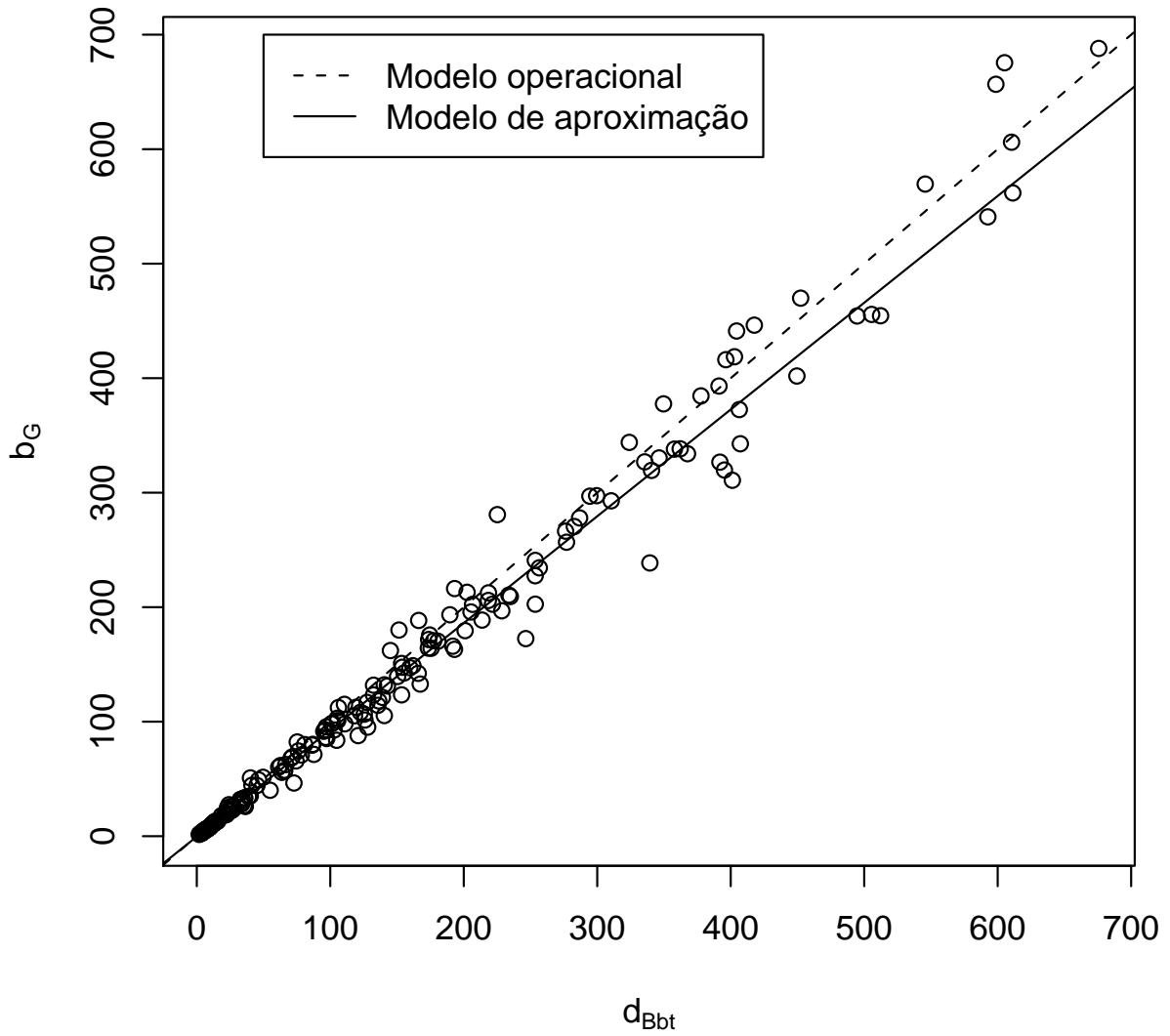


Figura 3.18 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pelos discos da base e do topo (d_{Bbt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

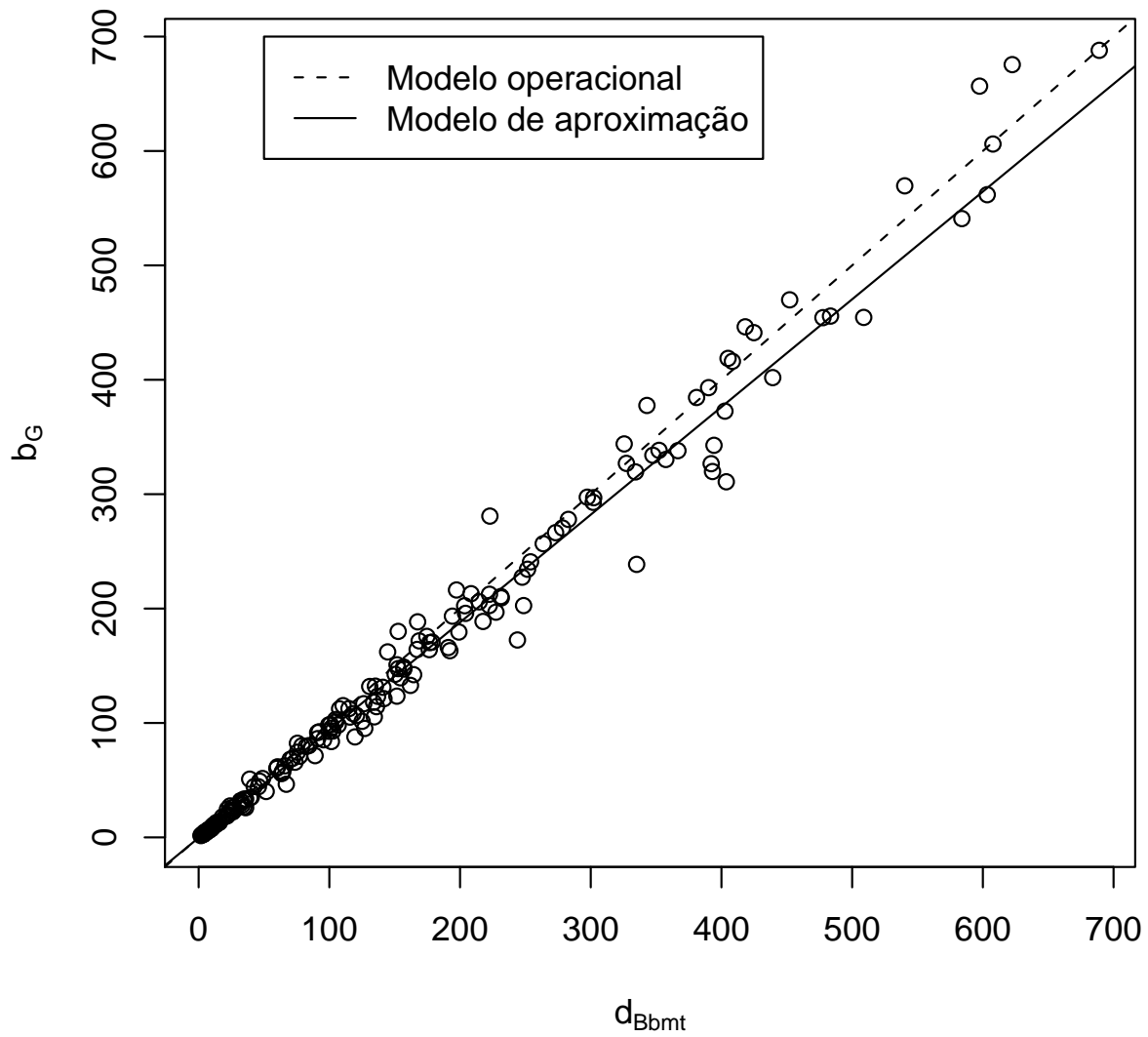


Figura 3.19 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pelos discos da base, do meio e do topo (d_{Bbmt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

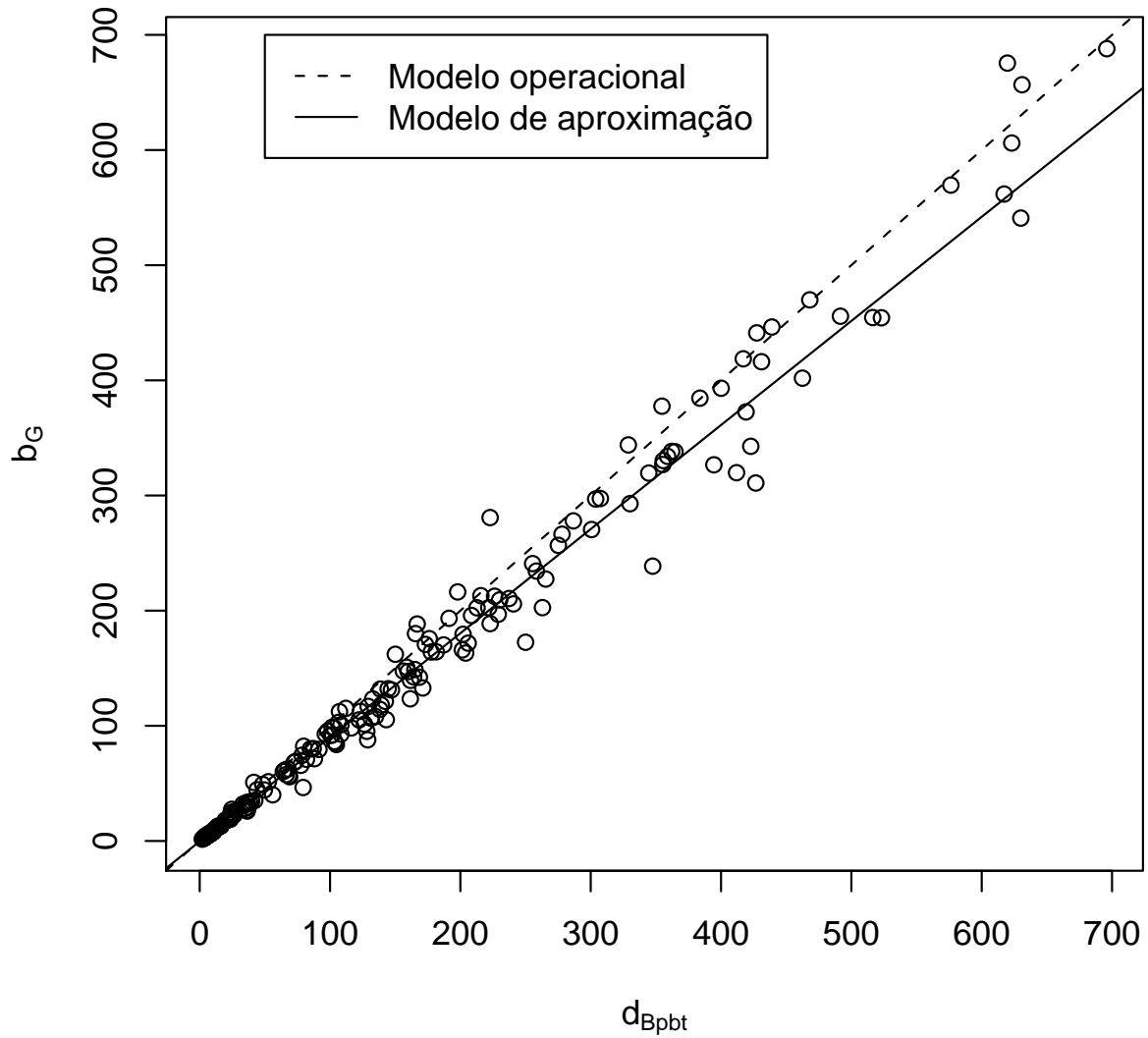


Figura 3.20 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pela média ponderada dos discos da base e do topo (d_{Bpbt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

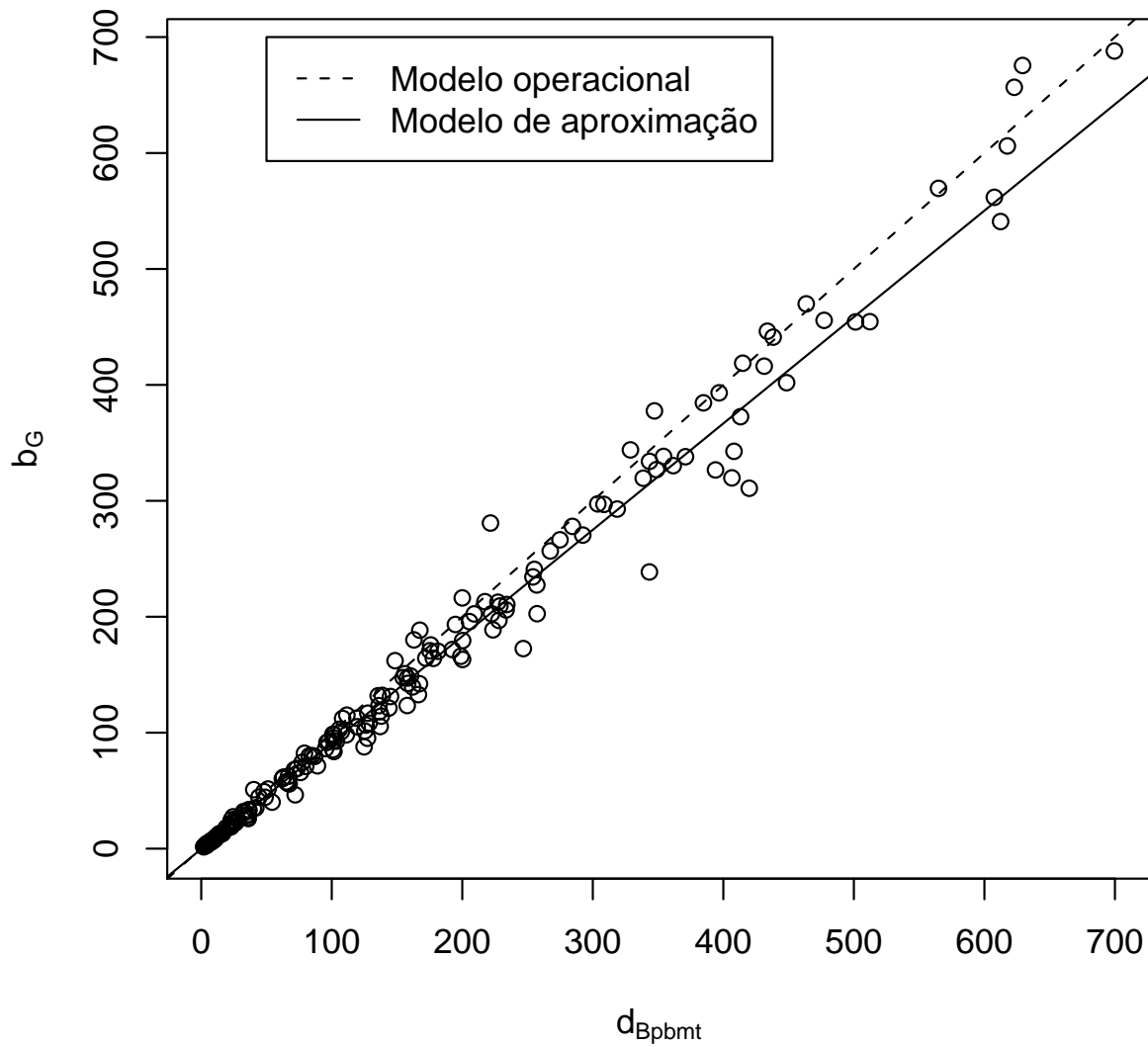


Figura 3.21 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pela média ponderada dos discos da base, do meio e do topo (d_{Bpbmt}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

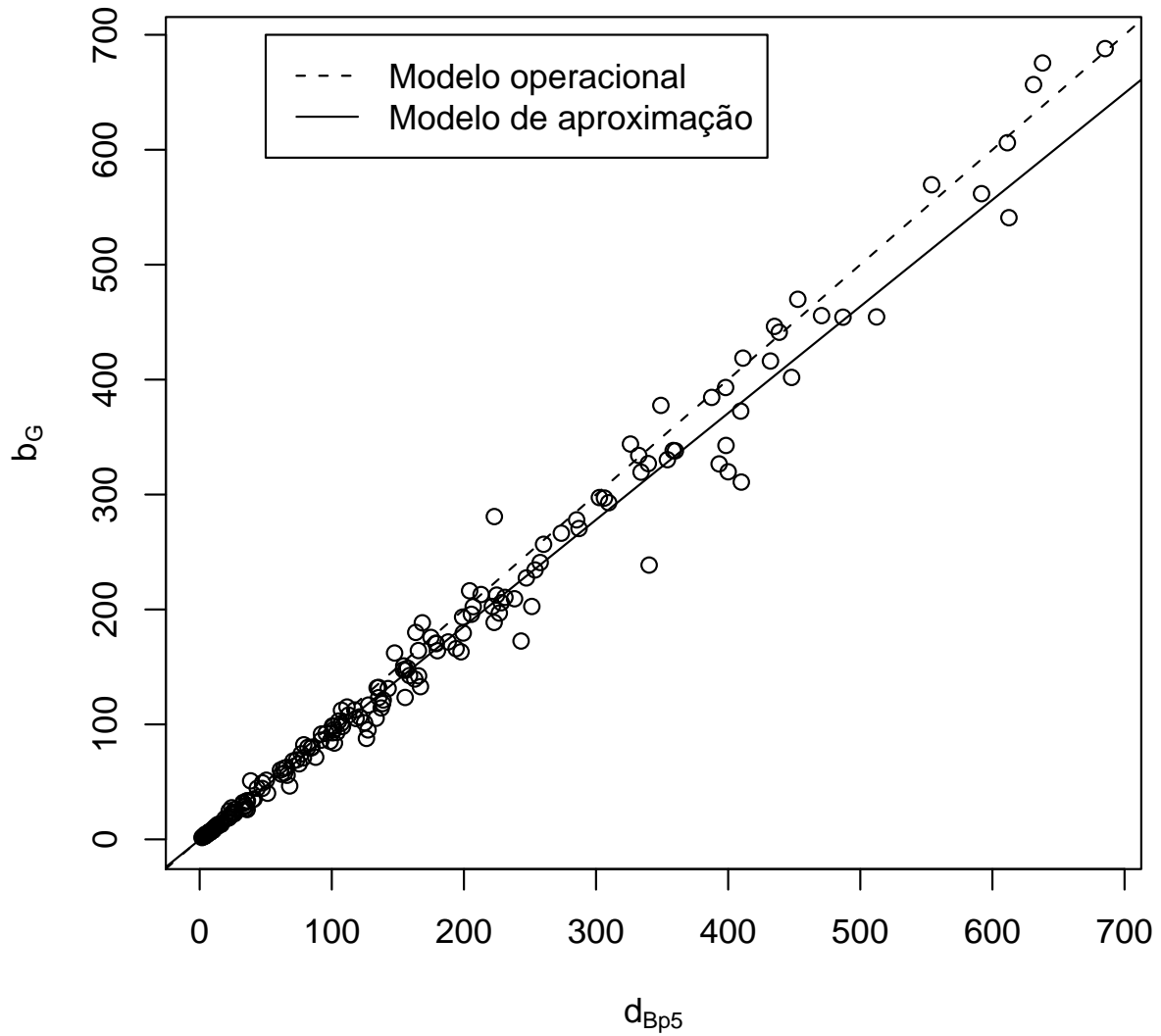


Figura 3.22 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pela média ponderada dos cinco discos (d_{Bp5}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

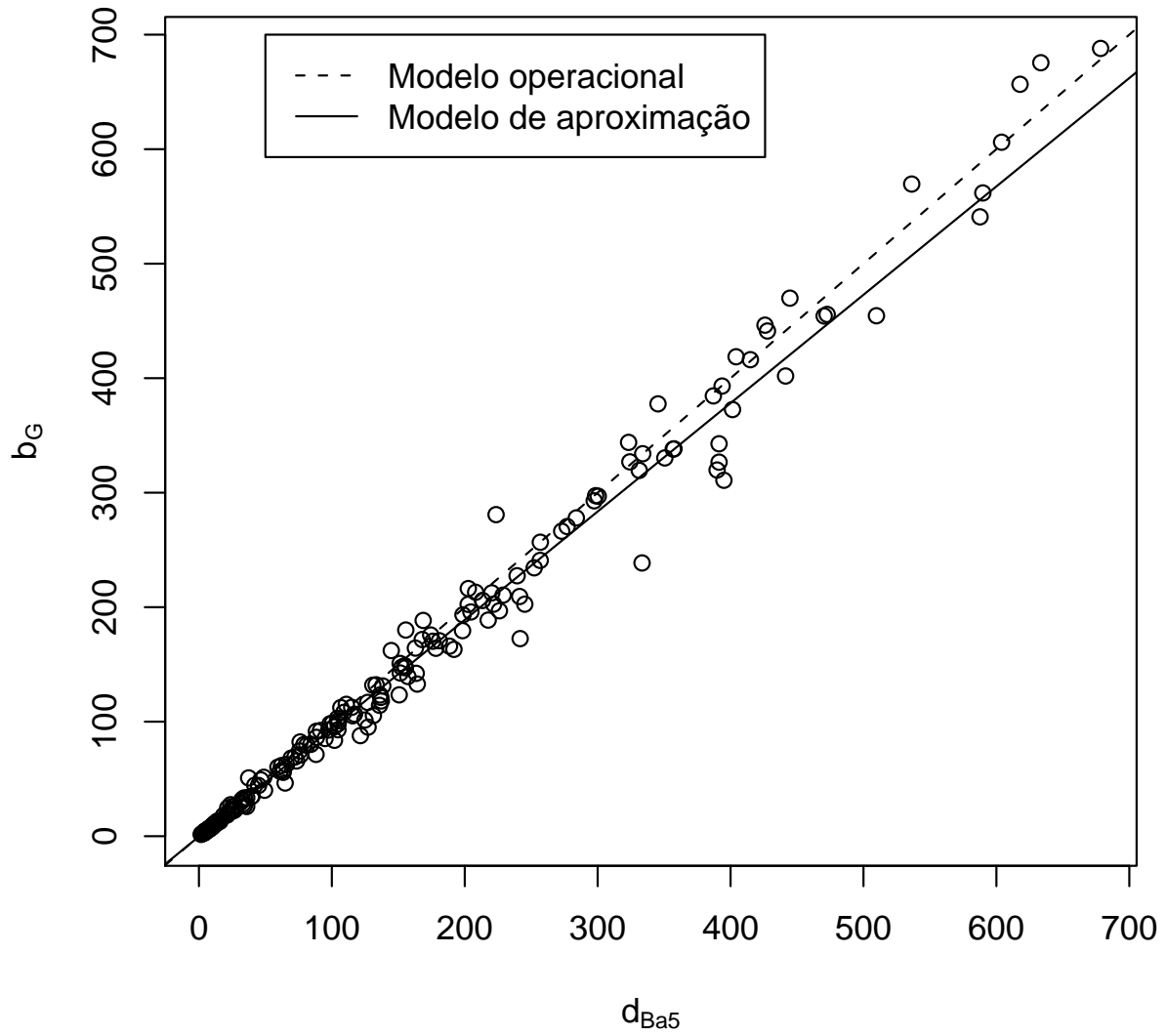


Figura 3.23 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade pela média aritmética dos cinco discos (d_{Ba5}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

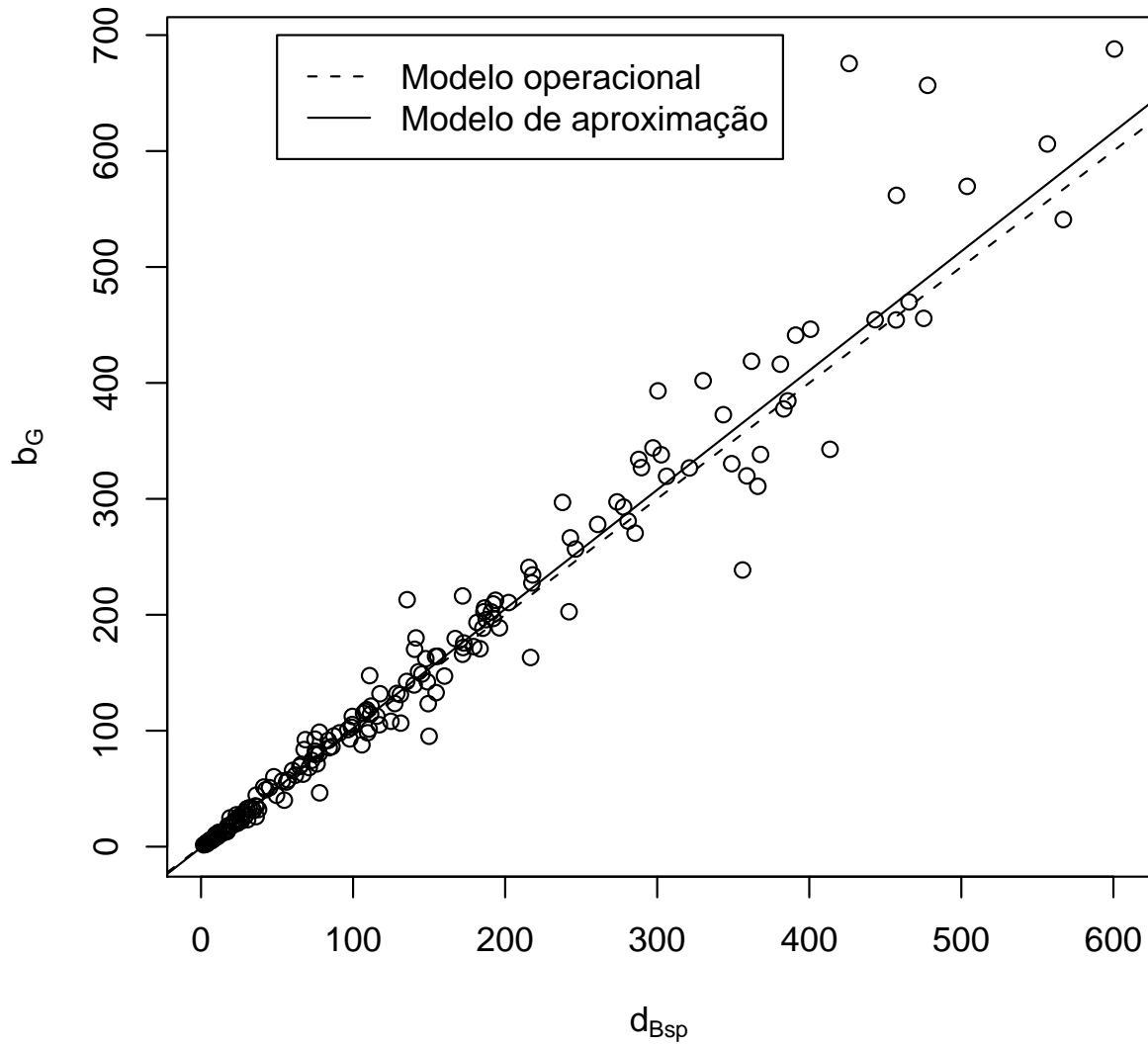


Figura 3.24 – Gráfico de dispersão da biomassa do tronco e a biomassa pelo método volumétrico, utilizando a determinação da densidade do tronco a densidade básica da espécie (d_{Bsp}), mostrando o modelo de aproximação (linha contínua) e o modelo operacional (linha tracejada)

ao modelo de aproximação. É interessante notar, no entanto, que apesar dessas duas técnicas não apresentarem, a partir da observação das Figura 3.17 e 3.18, desvios sistemáticos no modelo operacional, essas possuem grandes diferenças quanto a discrepância de estimação, sendo 2,976 para a técnica que usa apenas o disco do topo contra 108,127 para a técnica que usa da média aritmética dos disco da base e do topo. É provável que isto tenha ocorrido devido a desvios na linearidade da relação usada para o ajuste do modelo de aproximação da técnica que usa o disco da base e do topo e média aritmética. Qualquer desvio na linearidade é captado pelo AIC como um elemento de discrepância, embora incapaz de ser detectado por uma simples análise gráfica, como nesse caso.

Tabela 3.5 – Valores das discrepâncias e do erro padrão da estimativa dos modelos de determinação da biomassa do tronco, em função das técnicas de combinação das densidades básicas dos discos amostrados para determinar a densidade do tronco

Técnica	Discos	Média	Discrepâncias			Erro Padrão da Estimativa
			Geral	de Aproximação	de Estimação	
d_{Ba5}	todos	aritmética	1392,860	1313,017	79,843	13,82501
d_{Bb}	base	—	1567,799	1356,377	211,422	16,00633
d_{Bm}	meio	—	1383,842	1339,726	44,116	14,61496
d_{Bt}	topo	—	1398,453	1395,477	2,976	17,07760
d_{Bbt}	base e topo	aritmética	1436,477	1328,350	108,127	14,92169
d_{Bbmt}	base, meio e topo	aritmética	1406,933	1314,546	92,387	14,12631
d_{Bpbt}	base e topo	ponderada	1460,585	1315,318	145,267	14,30049
d_{Bpbmt}	base, meio e topo	ponderada	1502,690	1332,928	169,762	15,17482
d_{Bp5}	todos	aritmética	1431,135	1310,438	120,697	13,84542
d_{Bsp}	—	da espécie	1479,266	1440,401	38,866	20,00831

No que diz respeito à discrepância de estimação, a técnica que utiliza o disco do topo é o de menor discrepância, sendo seguida pela técnica com cinco discos e média ponderada, pela técnica que usa o disco do meio e pela técnica que utiliza os cinco discos e a média aritmética (Tabela 3.5). Com relação discrepância na aproximação, observa-se uma tendência de redução dessa discrepância à medida que o número de discos utilizados na determinação da densidade do tronco aumenta. O uso da média ponderada ou média aritmética resultou em pequena diferença, independente da quantidade de discos, com uma certa vantagem para a média ponderada, principalmente quando todos os cinco discos são utilizados, já que foi a técnica que resultou na menor discrepância na aproximação dentre todas as técnicas estudadas.

Ao contrario do que foi observado a análise da densidade, a discrepância geral nos modelos de biomassa foi menor para as técnicas que utilizam apenas um disco do que as técnicas que se utilizam de vários discos, com exceção da técnica que usa o disco da base. Dentre todas as técnicas estudadas, o uso do disco do meio do tronco gerou a menor discrepância geral (1383,8424), sendo claramente superior que a segunda menor discrepância, gerada pela técnica que se utiliza da média

aritmética de todos os cinco discos (1392, 860). Portanto, a determinação da densidade do tronco com base na densidade básica apenas do disco do meio é indicada para realizar a determinação de biomassa através da técnica volumétrica.

A determinação da densidade do tronco com base na densidade básica da espécie, apresentou a maior discrepância de aproximação e também o maior erro padrão da estimativa. Entretanto, o uso da densidade básica da espécie apresentou uma discrepância de estimação relativamente pequena, quando comparada às outras técnicas. Também é importante ressaltar que a biomassa é estimada na ordem de grandeza de centenas de quilogramas e, portanto, diferenças nos erros padrões na ordem de unidade de quilograma são pouco relevantes. De qualquer forma, a discrepância de aproximação mostra que a utilização de uma estimativa de densidade média por espécie para predição de biomassa é uma abordagem inferior se comparada a utilização de qualquer técnica de obtenção da densidade do tronco das árvores individualmente.

A inclusão dos dois fatores de erro nos modelos representantes da técnica volumétrica de determinação da biomassa do tronco surtiram efeitos muito distintos dos que se observa quando apenas as técnicas de determinação da densidade é considerada. Ao isolar o fator representado pelo método de determinação da densidade básica média, as técnicas que utilizaram um maior número de discos se apresentaram, de maneira geral, menos discrepantes e mais adequados para aplicação no modelo operacional. Neste caso, a utilização de discos individuais para determinação da densidade do tronco não se mostrou adequada. Por outro lado, ao incluir volume como fator de erro na determinação da biomassa do tronco, as técnicas que utilizaram um menor número de discos se mostraram mais adequadas. A inclusão do volume, aliás, além de inverter a ordem classificatória das técnicas, trouxe para os modelos de predição analisados o problema da heterocedasticidade da variância. Ou seja, a inclusão do volume trouxe um acréscimo na incerteza das predições de biomassa do modelo operacional para as árvores de maior tamanho.

3.4 Conclusões

Devido aos problemas apresentados pela técnica volumétrica na determinação da biomassa e, conseqüentemente, na construção de modelos de predição da biomassa do fuste em árvores individuais, conclui-se ser mais adequada a utilização da técnica gravimétrica em trabalhos de determinação da biomassa visando a construção de modelos empíricos para predição da biomassa de árvores em pé. Corroborar com essa conclusão o fato da técnica gravimétrica, apesar de exigir maior esforço de campo, simplificar os trabalhos de laboratório e obter resultados mais confiáveis.

No processo de construção do modelo referente à técnica gravimétrica, os resultados sugerem que para determinação do teor de umidade do tronco poderiam ter sido usados apenas três discos na determinação do teor de umidade do fuste, o que reduziria ainda mais os esforços de campo e laboratório.

Os resultados do presente trabalho mostram que a recomendação metodológica de usar a média ponderada da densidade básica dos cinco discos como estimativa da densidade tronco de uma árvore (Sanquetta, 2002) tem importância secundária na técnica volumétrica de determinação da biomassa do tronco, pois esta estimativa de densidade será multiplicada pela medida do volume nos modelos de predição da biomassa.

Para os casos onde a técnica gravimétrica é inviável de ser aplicada é preferível, de acordo com os resultados, determinar a densidade do fuste em nível de árvore e não utilizar estimativas de densidades médias da espécie, obtidas de forma empírica ou através de revisão de literatura.

Referências

ARAUJO, T.M.; HIGUCHI, N.; CARVALHO JR., J.A. Comparação de métodos para determinar biomassa na região amazônica, **Anais da Academia Brasileira de Ciência**, Rio de Janeiro, v. 68, n. 1, p. 35-41, ago. 1996.

BARBOSA, R.I.; FEARNSIDE, P. Wood density of trees in open savannas of the Brazilian Amazon. **Forest Ecology and Management**, Wisconsin, v. 199, n. 1, p. 115-123, May. 2004.

BROWN, S; GILLESPIE, A.J.R.; LUGO, A.E. Biomass estimation methods for tropical forest with applications to inventory data, **Forest Science**, Maryland, v. 35, n. 4, p. 881-902, Dec. 1989.

BROWN, S. Estimating biomass and biomass change of tropical forests: A primer. Food and Agriculture Organization of the United Nations (FAO), Rome, 1997. 55 p. (Forestry Paper 134).

BURHAM, K.P.; ANDERSON, D.R. Model selection and multimodel inference: practical information theoretic approach. 2nd.ed. New York: Springer, 2002. 488 p.

DRAPER, N.R.; SMITH, H. Applied regression analysis. 3rd.ed. New York: Wiley, 1998. 706 p.

FEARNSIDE, P.M. Wood density for estimating forest biomass in Brazilian Amazonia. **Forest Ecology and Management**, Wisconsin, v. 90, n. 2, p. 59-87, Oct. 1997.

HIGUCHI, N.; CARVALHO Jr., J. A. Fitomassa e conteúdo de carbono de espécies arbóreas da Amazônia. In: SEMINÁRIO DE EMISSÃO x SEQUESTRO DE CO_2 : UMA NOVA OPORTUNIDADE DE NEGÓCIOS PARA O BRASIL. 1994. Rio de Janeiro. **Anais...** Rio de Janeiro: Companhia Vale do Rio Doce, 1994. p. 125-153.

MOOD, A.M.; GRAYBILL, F.A.; BOES, D.C. Introduction to the theory of statistics. 3rd.ed. Auckland: MacGraw-Hill, 1974. 564p.

NOGUEIRA, E.M.; NELSON, B.W.; FEARNSTIDE, P.M. Wood density in dense forest in central Amazonia, **Forest Ecology and Management**, Wisconsin, n. 208, p. 261-286, May. 2005.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. Akaike information criterion statistics. Tokio: KTK Scientific Publisher/D. Riedel, 1986. 290 p.

SANQUETTA, C.R. Métodos de determinação de biomassa florestal. In: SANQUETTA, C.R.; WATZLAWICK, L.F.; BALBINOT, R.; ZILLIOTO, M.A.B.; GOMES, F.S. (Ed.). **As florestas e o carbono**: [s.n.], 2002, cap. 6, p. 119-140.

SANQUETTA, C.R., BALBINOT, R. Metodologia para determinação de biomassa florestal. In: SIMPÓSIO LATINO AMERICANO SOBRE FIXAÇÃO DE CARBONO. 2004. **Anais...** Curitiba, 2004. p. 77-94.

TEIXEIRA, L. M. **Influência da intensidade de exploração seletiva de madeira no crescimento e respiração do tecido Lenhoso das árvores em uma floresta tropical de terra-firme na região de Manaus**. 2003. 61p. Dissertação de Mestrado em Biologia Tropical e Recursos Naturais-Convênio entre Instituto Nacional de Pesquisas na Amazônia e Universidade Federal do Amazonas, Manaus, 2003.

WENDLING, J.L.G. **Modelos matemáticos de crescimento e de produção em biomassa para árvores de *Euterepe edulis Mart.* plantadas a céu aberto**. 1998. 171 p. Dissertação de Mestrado em Engenharia Florestal-Universidade Federal do Paraná, Curitiba, 1998.

VIEIRA, S.A.; ALVES, L.F.; AIDAR, M.; ARAUJO, L.S.; BAKER, T.; BATISTA, J.L.F.; CAMPOS, M.C.; CAMARGO, P.B.; CHAVE, J.; DELLITI, W.B.C.; HIGUSHI, N.; HONORIO, E.; JOLY, C.A.; KELLER, M.; MARTINELLI, L.A.; MATTOS, E.A.; METZKER, T.; PHILLIPS, O.; SANTOS, F.A.M.; SHIMABUKURU, M.T.; SILVEIRA, M.; TRUMBORE, S.E. Estimation of biomass and carbon stocks: The case of Atlantic Forest, **Biota Neotropica**, São Paulo, v. 8, n. 2, p. 21-29, jun. 2008.

VITAL, B.R. Métodos de determinação da densidade da madeira. **Boletim Técnico-SIF**, Viçosa, n. 1, p. 1-21, 1984.

4 AJUSTE E SELEÇÃO DE MODELOS PREDITIVOS DA BIOMASSA AÉREA

Resumo

O objetivo deste capítulo é analisar vários modelos de predição de biomassa aérea de árvores originárias da Floresta Tropical Atlântica, considerando seus diferentes compartimentos. Por meio de método destrutivo, foram coletadas 200 árvores de dez espécies nativas provenientes de reflorestamento, para determinação da sua biomassa através da técnica gravimétrica. As variáveis resposta utilizadas foram a biomassa total (b), do tronco (b_t), de galhos (b_g), de ramos (b_r) e de folhas (b_f). As variáveis preditoras foram o dap , as alturas total e comercial e uma medida de penetração no lenho. Foram ajustados sete modelos, a partir das funções de Spurr e Schumacher-Hall, com e sem a inclusão da altura como variável preditora. Destes modelos, quatro eram gaussianos e três eram lognormais. Estes mesmos sete modelos foram ajustados incluindo a medida de penetração como variável preditora, totalizando quatorze modelos testados. A seleção dos melhores modelos de predição da biomassa total e da biomassa dos diferentes compartimentos se deu por meio do critério de informação de Akaike (AIC). O modelo de Schumacher-Hall se mostrou, de maneira geral, superior ao modelo de Spurr. A altura só se mostrou efetiva na explicação da biomassa das árvores quando em conjunto com a medida de penetração. Os modelos selecionados foram do grupo que incluíram a medida de penetração no lenho como variável preditora. Todos os modelos selecionados, exceto o de predição da biomassa de folhas, se mostraram adequados para aplicação na predição da biomassa aérea em áreas de reflorestamento.

Palavras-chave: Biomassa aérea; Modelos de predição; Floresta tropical atlântica

Abstract

The purpose of this chapter is to examine various tree biomass prediction models from the Atlantic Rain Forest species, considering its different compartments. Through destructive method, 200 trees were collected to determine their biomass by gravimetric technique. The response variables used were total biomass (b), stem biomass (b_t), branches biomass (b_g), small branches biomass (b_r) and leaves biomass (b_f). The predictor variables were the dbh , the total and stem height and a wood penetration measure. Seven models, as variations of Spurr and Schumacher-Hall volume equation models, were tested for the different tree components: wood (stem and large branches), little branches, leaves and total biomass. In general, Schumacher-Hall models were better than Spurr based models, and models that included only diameter (DBH) information performed better than models with diameter and height measurements. When a measure of penetration in the wood, as a surrogate of wood density, was added to the models, the models with the three variables: diameter, height and penetration, became the best models.

Keywords: Aboveground biomass; Prediction models; Atlantic rain forest

4.1 Introdução

As florestas tropicais vem despertando, nos últimos anos, grande interesse da sociedade e da comunidade científica, no que diz respeito ao seu valor ambiental. Este bioma tem sofrido pressões que alteram a sua estrutura e que até mesmo levam, em alguns casos, a sua completa destruição (SANTOS et al., 2001).

Esse crescente interesse vêm, principalmente, do fato da influência que a floresta tropical exerce no balanço dos gases de efeito estufa e suas conseqüências para as mudanças climáticas do planeta. Segundo (HIGUSHI; CARVALHO JR., 1994) este balanço é representado pelas emissões quando a floresta é queimada ou pelo potencial seqüestro quando do crescimento de um povoamento florestal.

Nestes dois contextos, para que se possa quantificar esse balanço se faz necessário, obter estimativas fidedignas da biomassa florestal, e por conseqüência do carbono, presente neste bioma (SANQUETTA, 2002).

Os inventários florestais, tem sido usados como ponto de partida para estimar a biomassa aérea nas florestas tropicais no Brasil (BROWN; GILLESPIE; LUGO, 1989). Os dados dendrométricos desses inventários, associados às estimativas de densidade básica média das espécies, são usados então para construir equações alométricas de predição (SEGURA, 2005). O uso dessa metodologia, apesar da simplicidade, pode, segundo Fearnside (1997) e Nogueira et al. (2007), levar a estimativas muito discrepantes. Essa discrepância se deve aos erros associados à obtenção das variáveis dendrométricas e ao erro associado à aplicação da densidade básica em nível de espécie. Esse último erro é mais importante já que, segundo Nogueira, et al. (2007), envolve a extrapolação das densidades básicas por espécie para diferentes áreas sem levar em consideração a variabilidade intraespecífica dessas densidades. Em alguns casos, no entanto, as equações de predição da biomassa são obtidas a partir da mensuração de árvores individuais, onde as variáveis dendrométricas, como diâmetro a altura do peito (dap) e altura total ou comercial, são medidas e associadas à biomassa aérea, determinada por métodos diretos (gravimetria) e indiretos (volumetria), através de técnicas de regressão (ARAUJO; HIGUSHI; CARVALHO JR., 1996; HIGUSHI, et al., 1998; NELSON et al., 1999 e SANTOS et al., 2001).

Essas equações, obtidas através da mensuração de uma amostra de indivíduos representativos da espécie ou de um grupo de espécies tem sido desenvolvidas no Brasil, principalmente, para região Amazônica. Para floresta Atlântica, porém, segundo (VIEIRA et al., 2008), o número de trabalhos que utilizam amostras destrutivas de árvores individuais para construção de equações de predição é muito reduzido. Vieira et al. (2008) demonstraram que as relações alométricas das árvores da floresta atlântica são muito distintas das da região amazônica. Desta forma a aplicação dessas equações para Floresta Atlântica é muito complicada, podendo levar a estimativas de biomassa muito errôneas. Isso reforça a necessidade de se estabelecer equações de predição de biomassa

sítio-específicas para floresta Atlântica, visando sua conservação e futuros projetos de seqüestro de carbono.

No entanto, na floresta Atlântica, a obtenção de amostras destrutivas, em áreas naturais é um processo complicado no que se refere à legislação, pois se trata de um bioma protegido e, atualmente de extensão muito reduzida. Para se estabelecer equações de predição de biomassa para a floresta atlântica se faz necessário tomar amostras de áreas de reflorestamento onde espécies representativas desse bioma foram inseridas. É claro, que a aplicação de equações, obtidas a partir de amostras de reflorestamento, em áreas naturais apresenta limitações semelhantes as da aplicação de equações desenvolvidas na Amazônia na floresta Atlântica. Isso ocorre pois, em áreas de reflorestamento os indivíduos arbóreos das espécies tropicais apresentam arquitetura diferenciada quando comparados aos indivíduos arbóreos encontrados na floresta primária. Porém, projetos de reflorestamento que visam à conservação do bioma Atlântico ou obtenção de créditos de carbono vem ganhando importância nos últimos anos e se faz necessário estabelecer equações de predição de biomassa adequadas a essa situação.

Sendo assim, o presente trabalho tem como objetivo estabelecer equações de predição da biomassa total e dos diferentes compartimentos da árvore (tronco, galhos, ramos e folhas), a partir de uma amostra destrutiva de 200 árvores de dez espécies distintas, originárias do bioma Atlântico. Para isso, foram testados diversos modelos comumente usados na literatura para predição de volume e biomassa, incluindo uma variável preditora referente à medida de penetração no lenho, já que, em tese, essa possui uma boa relação com a densidade das árvores.

4.2 Material e Métodos

4.2.1 Área de estudo

O estudo foi realizado em área pertencente à Companhia Vale do Rio Doce, recuperada com essências nativas e localizada ao norte do Estado do Espírito Santo, predominantemente, no município de Linhares com uma pequena abrangência nos de Sooretama e de Jaguaré. Geograficamente situa-se entre os paralelos 19° 06'-19° 18' de latitude sul e entre os meridianos 39° 45'-40° 19' de longitude oeste, contando com uma área de 22.000 ha.

4.2.2 Escolha das espécies

A escolha das espécies para realização deste estudo seguiu dois critérios que juntos buscavam uma certa generalidade quanto a aplicação deste em áreas distintas àquela onde o trabalho foi realizado. Primeiramente buscou-se espécies de ampla ocorrência geográfica e entre essas foram escolhidas espécies que possuíam, segundo a literatura, densidades básicas bem distintas. Desta forma foram escolhidas dez espécies a saber: *Caesalpinia ferrea*, *Pteropcarpus violaceus*, *Zeyhera*

tuberculosa, *Schizolobium parayba*, *Pterygota brasiliensis*, *Cordia trichotoma*, *Astronium graveolens*, *Joannesia princeps*, *Cariniana legalis* e *Tabebuia serritifolia*.

Os indivíduos arbóreos usados no trabalho foram coletados, segundo um processo de amostragem estratificada, sendo cada classe de diâmetro definida com um estrato. Foram definidas 10 classes de diâmetro por espécie, segundo a distribuição diamétrica, da espécie, observada no povoamento. Em cada classe de diâmetro da espécie foram amostrados dois indivíduos.

Como foram selecionadas 10 espécies com 10 classes de diâmetro em cada espécie e dois indivíduos por classe de diâmetro, a amostragem totalizou 200 indivíduos, cuja frequência por classe diamétrica, considerando toda a amostra, é apresentada na Tabela 4.1.

Tabela 4.1 – Distribuição diamétrica da amostra, com a frequência absoluta e a frequência relativa observada

Classe de Diâmetro	Frequências		
	Absoluta	Relativa (%)	Acumulada (%)
(5 ; 8,5]	26	13,0	13,0
(8,5 ; 12]	21	10,5	23,5
(12 ; 15,5]	20	10,0	33,5
(15,5 ; 19]	25	12,5	46,0
(19 ; 22,5]	24	12,0	58,0
(22,5 ; 26]	20	10,0	68,0
(26 ; 29,5]	24	12,0	80,0
(29,5 ; 33]	15	7,5	87,5
(33 ; 36,5]	13	6,5	94,0
(36,5 ; 40]	12	6,0	100,0
Total	200	100,0	-

A distribuição das árvores por classe de diâmetro não ficou uniforme na amostra, já que cada espécie possuía uma estrutura diamétrica distinta e a amostragem seguiu essa estrutura. Desta maneira, a amostragem conta com mais indivíduos de pequeno porte e menos indivíduos de médio e grande porte.

4.2.3 Determinação da biomassa

A biomassa foi determinada para os diferentes compartimentos das árvores e esses compartimentos foram definidos de acordo com a literatura (Higushi et al. (1998); Sanquetta (2002); Soares et al. (2002) e Segura (2005)).

Copa

Cada árvore foi desfolhada, sendo suas folhas pesadas no campo com célula de carga com capacidade máxima de 200 kg e precisão na casa de centenas de gramas. Do conjunto de folhas destacadas foi retirada uma amostra de aproximadamente 500g, para determinação da massa seca, após secagem em estufa de circulação forçada de ar a 75°C.

Em seguida, cada árvore foi desgalhada separando-se os galhos em dois compartimentos. O compartimento com diâmetros entre 10 cm e 5 cm foi chamado de galho e o compartimento com diâmetro inferior a 5 cm foi chamado de ramo.

Os dois compartimentos de galhos foram separadamente pesados no campo com dinamômetro, inclusive a ponta da árvore com diâmetro menor que 5 cm. No conjunto de ramos e galhos foram retirados fragmentos de galhos localizados em diferentes posições da copa da árvore, os quais foram misturados, formando-se uma amostra composta.

Desta amostra, foi retirada uma subamostra de aproximadamente 500 g de ramos e 500 g de galhos, para determinação da sua massa seca, após secagem em estufa de circulação forçada de ar a 103°C. A biomassa de folhas, ramos galhos foram obtidas de acordo com a seguinte expressão:

$$b_{ci} = m_i (1 - u_i/100) \quad (4.1)$$

onde: i é o índice que se refere às árvores; b_{ci} é a biomassa do biomassa de folhas ou ramos ou galhos (kg); m_i é a massa verde biomassa de folhas ou ramos ou galhos (kg); e u_i é o teor de umidade de folhas ou ramos ou galhos (%).

Tronco

O tronco da árvore foi seccionado e pesado com auxílio de uma célula de carga com capacidade máxima de 200 kg e precisão na casa de centenas de gramas. Discos de madeira, com casca, de aproximadamente 5 cm de espessura foram extraídos nas posições relativas 0, 25, 50, 75 e 100 % da altura comercial. A altura comercial foi definida como altura até a primeira inserção de galhos ou até um diâmetro mínimo de 10 cm. Os discos extraídos foram pesados para determinação da sua massa seca, após secagem em estufa de circulação forçada de ar a 103°C. A biomassa do tronco das árvores, com casca foi obtida de acordo com a seguinte expressão:

$$b_{ti} = m_i (1 - u_i/100) \quad (4.2)$$

onde: i é o índice que se refere às árvores; b_{ti} é a biomassa do tronco (kg); m_i é a massa verde do tronco (kg); e u_i é o teor de umidade do tronco (%).

Devido aos critérios usados na compartimentação, as árvores com dap menores que 10 cm foram consideradas desprovidas de galhos, tendo apenas tronco, ramos e folhas. Desta forma, a

variável resposta biomassa de galhos foi obtida num menor número de árvores (165). Além disso, os trabalhos de campo ocorreram durante um período de queda das folhas sendo que algumas árvores apresentavam pouca ou nenhuma folha. Da mesma forma que os galhos, então, essa variável foi coletada num número menor de árvores (176).

4.2.4 Determinação das variáveis preditoras

Com as árvores ainda em pé, foram medidos o diâmetro a altura do peito *DAP* com suta (cm), altura total (m) e altura comercial (m) com clinômetro eletrônico vertex III. Considerou-se como tronco, a medida da base até a primeira inserção significativa de galhos. Além disso, foi usado um penetrômetro para medir a grau de penetração (mm) no lenho na altura do dap, como variável preditora discriminatória da densidade da madeira.

4.2.5 Ajuste e Seleção de Modelos

Segundo Parresol (1999), os modelos de predição de biomassa apresentam, geralmente, erros heteroscedásticos, ou seja, a variância não é constante. Para contornar esse problema, a abordagem convencional consiste em transformar a variável resposta, nesse caso a biomassa, ou utilizar a regressão ponderada que, em termos do modelo ajustado, também pode ser considerada uma transformação da variável resposta.

O ajuste de um modelo com a variável resposta transformada implica que a modelagem será realizada numa escala distinta da escala em que a variável foi medida e em que as predições do modelo serão realizadas na prática. Embora esse problema da heteroscedasticidade seja corrigido, a transformação resulta em duas sérias limitações. Primeiramente, certas propriedades estatísticas das estimativas geradas na escala transformada não se aplicam na escala original, como ausência de viés e variância constante. Em segundo lugar, vários dos critérios de comparação dos modelos não são comparáveis quando os modelos estão em escalas distintas, como por exemplo: erro padrão da estimativa, coeficiente de determinação e AIC.

A abordagem de modelagem nesse trabalho consistiu em adaptar os modelos ao problema da heteroscedasticidade. Na primeira classe de modelos, designada de *modelos Gaussianos heteroscedásticos*, a biomassa é modelada como uma distribuição Gaussiana com a seguinte função de densidade probabilística

$$f(y) = \frac{1}{\sigma(x_1, \dots, x_k)\sqrt{2\pi}} \exp \left\{ -\frac{[y - \mu(x_1, \dots, x_k)]^2}{2(\sigma(x_1, \dots, x_k))^2} \right\}, \quad -\infty < y < \infty$$

onde

y é a biomassa;

$\mu(x_1, \dots, x_k)$ é uma expressão que descreve o comportamento da média em função de k variáveis preditoras; e

$\sigma(x_1, \dots, x_k)$ é uma expressão que descreve o comportamento do desvio padrão em função de k variáveis preditoras.

Note que a heteroscedasticidade foi incorporada ao modelo e que os diferentes modelos dessa classe serão definidos conforme as expressões da média ($\mu(\cdot)$) e do desvio padrão ($\sigma(\cdot)$).

A segunda classe de modelos, designada de *modelos Log-Normais*, consiste em utilizar a distribuição Log-normal para modelar a biomassa. A transformação logarítmica é uma transformação comumente utilizada na abordagem convencional, para modelar a biomassa ou o volume das árvores. Nessa classe de modelos, a variável resposta passa a ter distribuição de densidade probabilística

$$f(y) = \frac{1}{y \sigma \sqrt{2\pi}} \exp \left\{ -\frac{[\ln(y) - \mu(x_1, \dots, x_k)]^2}{2\sigma^2} \right\}, \quad 0 \leq y < \infty$$

onde

y é a biomassa;

$\mu(x_1, \dots, x_k)$ é uma expressão que descreve o comportamento da média em função de k variáveis preditoras, que nesse caso está na escala logarítmica;

σ é o desvio padrão constante, uma vez que o modelo Log-normal é assumido como homoscedástico.

Note que nesse modelo, a modelagem da média ocorre na escala logarítmica sem a transformação da variável resposta, pois essa é modelada como Log-normal. Os diferentes modelos dessa classe se distinguem pela expressão que modela a média e a tabela 4.2 apresenta as expressões estudadas.

Esses modelos foram utilizados para modelar tanto a biomassa total da árvore, quanto a biomassa dos componentes da árvore (tronco, galhos, ramos e folhas). No caso da biomassa do tronco, no lugar da altura total utilizou-se a altura comercial.

Nos modelos Gaussianos heteroscedásticos, o desvio padrão foi sempre modelado como uma função de potência do diâmetro da árvore:

$$\sigma(d) = \alpha_0 d^{\alpha_1}.$$

As diferenças de capacidade preditiva entre os modelos dessa classe refletem, portanto, apenas a adequação da expressão que modela a média.

Todos os modelos foram ajustados pelo Método da Máxima Verossimilhança (MOOD; GRAYBILL; BOES; 1974) utilizando um algoritmo de otimização numérica em todos os casos. Nos casos

Tabela 4.2 – Expressões para modelagem da média nos modelos Gaussianos heteroscedásticos e nos modelos Log-normais. d — DAP; h — altura total ou altura comercial; p — medida de penetração no lenho

MODELO	FORMA DA EXPRESSÃO DA MÉDIA	
	TRADICIONAL (a)	C/ MEDIDA DE PENETRAÇÃO (b)
<i>Modelos Gaussianos Heteroscedásticos</i>		
1 - Diamétrico	$\beta_0 d^{\beta_1}$	$\beta_0 d^{\beta_1} p^{\beta_2}$
2 - Spurr	$\beta_0 + \beta_1 (d^2 h)$	$\beta_0 + \beta_1 (d^2 h) + \beta_2 p$
3 - Spurr-potência	$\beta_0 (d^2 h)^{\beta_1}$	$\beta_0 (d^2 h)^{\beta_1} p^{\beta_2}$
4 - Schumacher-Hall	$\beta_0 d^{\beta_1} h^{\beta_2}$	$\beta_0 d^{\beta_1} h^{\beta_2} p^{\beta_3}$
<i>Modelos Log-Normais</i>		
5 - Diamétrico	$\beta_0 + \beta_1 \ln(d)$	$\beta_0 + \beta_1 \ln(d) + \beta_2 \ln(p)$
6 - Spurr	$\beta_0 + \beta_1 \ln(d^2 h)$	$\beta_0 + \beta_1 \ln(d^2 h) + \beta_2 \ln(p)$
7 - Schumacher-Hall	$\beta_0 + \beta_1 \ln(d) + \beta_2 \ln(h)$	$\beta_0 + \beta_1 \ln(d) + \beta_2 \ln(h) + \beta_3 \ln(p)$

em as estimativas de máxima verossimilhança são coincidentes com as estimativas de quadrados mínimos, ou seja, nos modelos com erro Gaussiano homoscedástico, as estimativas obtidas pelo algoritmo de otimização numérica foram idênticas às estimativas de quadrados mínimos. O ajuste foi implementado no pacote “stats4” do software estatístico “R”, utilizando a função de ajuste “mle” (*maximum likelihood estimation*), que utiliza por default a função “optim” para otimização numérica.

4.2.6 Análise dos modelos selecionados

Os parâmetros das funções da média ($\mu(\cdot)$) e desvio padrão ($\sigma(\cdot)$) foram ajustados pelo método da máxima verossimilhança. Por se tratar de um algoritmo iterativo de otimização, valores iniciais dos parâmetros foram obtidos pelo método dos quadrados mínimos. Isso foi realizado para garantir a convergência do processo iterativo para um mínimo global da função de verossimilhança.

A seleção dos modelos foi feita através do Critério de Informação de Akaike (AIC)(AKAIKE, 1974, 1985), sendo que quanto menor o valor de AIC obtido pelo modelo, melhor sua qualidade para realizar previsões de biomassa, considerando os dados coletados. Segundo Burnham e Anderson (2002), o critério de informação de Akaike classifica, quanto ao seu ajuste, vários modelos concorrentes *a priori* selecionados, independente de sua função de distribuição e da relação funcional considerada para descrever a média. No entanto, se todos os modelos testados forem pobres no que se refere ao ajuste, o *AIC*, ainda assim, indicará o melhor entre eles.

Para verificar a qualidade do ajuste dos modelos selecionados por este critério, foi realizada uma análise gráfica dos resíduos percentuais e da normalidade dos resíduos padronizados. O exame gráfico dos resíduos visou analisar o desempenho do modelo na predição do próprio conjunto de dados a fim de avaliar a magnitude do erro na predição da biomassa de árvores de diferentes tamanhos.

Outra medida que foi usada para avaliar o ajuste do modelo selecionado é o coeficiente de determinação empírico:

$$R_e^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}, \quad (4.3)$$

onde y_i é a biomassa observada, \hat{y}_i é biomassa predita pelo modelo e \bar{y} é a média da biomassa observada.

Este fornece uma medida de superioridade do modelo selecionado em relação ao modelo mais simples que é a média amostral, sendo um indicativo da qualidade do ajuste do mesmo.

4.3 Resultados e Discussão

4.3.1 Análise descritiva das variáveis usadas no ajuste dos modelos

As estatísticas descritivas dos dados utilizados como variáveis resposta para construção de modelos de predição podem ser observados na Tabela 4.3.

Tabela 4.3 – Estatísticas descritivas das variáveis resposta utilizadas na construção dos diferentes modelos de predição, onde Min. é o valor mínimo, Max. é valor máximo, s é o desvio padrão, CV é o coeficiente de variação e n é o tamanho da amostra

	Min.	Max.	Mediana	Média	s	CV (%)	n
Biomassa total	2,54	1014,81	162,29	250,11	253,40	101,32	200
Biomassa do tronco	1,58	688,02	95,53	142,66	151,70	106,34	200
Biomassa de ramos	0,22	172,08	23,26	36,75	38,12	103,73	200
Biomassa de galhos	0,59	410,48	31,81	75,95	89,76	118,17	165
Biomassa de folhas	0,02	52,69	5,52	8,73	9,43	108,02	176

Foi notável a contribuição da biomassa do fuste e da biomassa de galhos em relação à biomassa total. A biomassa dos ramos tiveram contribuição menor na biomassa total e as folhas uma contribuição muito pequena, haja visto que a coleta transcorreu num período de queda de folhas das árvores coletadas.

A distribuição das frequências dos dados de diâmetro em função das frequências dos dados de altura são apresentados na Tabela 4.4.

Tabela 4.4 – Distribuição das freqüência do dados de diâmetro em função dos dados de altura total

Classe de diâmetro	Classe de altura						
	(3 ; 6]	(6 ; 9]	(9 ; 12]	(12 ; 15]	(15 ; 18]	(18 ; 21]	(21 ; 24]
(5 ; 8,5]	13	12	1	0	0	0	0
(8,5 ; 12]	1	9	10	1	0	0	0
(12 ; 15,5]	0	3	8	8	1	0	0
(15,5 ; 19]	0	0	4	13	6	2	0
(19 ; 22,5]	0	0	0	6	11	7	0
(22,5 ; 26]	0	0	0	4	7	9	0
(26 ; 29,5]	0	0	0	2	11	8	3
(29,5 ; 33]	0	0	0	1	3	8	3
(33 ; 36,5]	0	0	0	1	0	6	4
(36,5 ; 40]	0	0	0	0	0	4	8

A distribuição dos diâmetros em relação as alturas totais na amostra coletada demonstra a estrutura alométrica característica de povoamentos multiespecíficos, onde, a partir das classes interme-

diárias de altura, indivíduos de classes de diâmetros muito distintas, ocorrem numa mesma classe de altura (Tabela 4.4).

Esta observação é reforçada a partir do coeficiente de variação apresentado pela variável altura total, menor que o coeficiente de variação do diâmetro das mesmas (Tabela 4.5). Além disso o coeficiente de variação da altura do fuste demonstra a variação das arquiteturas da árvores amostradas, também típica de amostras compostas de indivíduos de diversas espécies e com diferentes arquitetura de copa.

Tabela 4.5 – Estatísticas descritivas das variáveis preditoras utilizadas na construção dos modelos de predição da biomassa total, biomassa do tronco, biomassa de galhos, biomassa de ramos e biomassa de folhas. Onde Min. é o valor mínimo, Max. é valor máximo, s é o desvio padrão, CV é o coeficiente de variação e n é o tamanho da amostra

	Min.	Max.	Mediana	Média	s	CV (%)
DAP	5,16	39,79	20,85	20,76	9,47	45,61
Altura total	3,30	24,40	15,10	14,59	5,15	35,31
Altura do fuste	0,79	16,93	7,50	7,56	4,04	53,37
Penetração	3,00	40,00	10,00	11,11	6,99	62,93

4.3.2 Seleção dos modelos

A Tabela 4.6 apresenta os valores de AIC para os diversos modelos ajustados para predição da biomassa total, biomassa do tronco, biomassa de galhos, biomassa de ramos e biomassa de folhas.

Com exceção da biomassa das folhas, os modelos diamétricos sem a informação de penetração no lenho se mostraram superiores aos modelos que utilizaram o diâmetro e a altura como variáveis preditoras. Esses resultados mostram que, para as árvores estudadas, a inclusão da informação altura não implica necessariamente em melhores predições de biomassa.

Isso já era esperado, já que, a análise descritiva das variáveis preditoras indicaram uma variação alométrica muito acentuada entre os indivíduos componentes da amostra. Além disso, trata-se de uma amostra composta por diferentes espécies de densidade básica muito distintas, onde a relação entre os atributos de tamanho das árvores tem relação muito variável com a biomassa das mesmas.

A inclusão da variável resistência a penetração foi um fator importante, no que se refere, à melhoria na qualidade do ajuste, dos modelos de predição de todas as variáveis resposta consideradas, inclusive para a predição da biomassa de folhas. Além disso, os modelos com funções cujas variáveis preditoras não eram combinadas (Schumacher-Hall e Diamétricos) também apresentaram melhor desempenho geral. Este resultado é equivalente ao obtido por Santos et al. (2001) trabalhando com modelos de predição de biomassa total na Amazônia. Estes autores ajustaram modelos

baseados na função de Shumacher-Hall e na função de Spurr com a variável combinada, concluindo que o modelo da variável combinada de Spurr apresentava pior ajuste que o de Schumacher-Hall.

Tabela 4.6 – Valores de *AIC* para os modelos testados de predição da biomassa total, biomassa do tronco, biomassa de galhos, biomassa de ramos e biomassa de folhas

CLASSE	MODELO	BIOMASSA				
		TOTAL	TRONCO	GALHOS	RAMOS	FOLHAS
Gaussiano	1.a - Diamétrico	1879,163	1838,663	1642,649	1585,237	1252,669
	2.a - Spurr	2027,124	1826,610	1652,628	1622,236	1096,001
	3.a - Spurr-potência	2031,657	1818,630	1650,789	1616,841	1093,535
	4.a - Schumacher-Hall	2004,986	1748,758	1647,992	1594,528	1074,615
Log Normal	5.a - Diamétrico	2054,754	1850,776	1560,237	1620,329	2661,127
	6.a - Spurr	2070,334	1813,040	1566,634	1635,604	1076,870
	7.a - Schumacher-Hall	2051,280	1763,350	1562,136	1619,343	3294,920
Gaussiano	1.b - Diamétrico	2072,081	2286,295	1622,580	1531,550	1086,367
	2.b - Spurr	2007,262	1806,930	1653,100	1608,678	1088,311
	3.b - Spurr-potência	1890,945	1669,170	1817,703	1561,481	1078,833
	4.b - Schumacher-Hall	1855,158	1618,51	1624,040	1540,900	1060,792
Log Normal	5.b - Diamétrico	1904,650	1801,746	1543,323	3987,563	3239,023
	6.b - Spurr	1895,654	1659,030	1546,783	1546,768	1065,720
	7.b - Schumacher-Hall	1873,350	1597,260	1546,338	1534,232	3124,281

Considerando somente os modelos de predição da biomassa total, estes acompanharam a tendência geral, com valores de *AIC* menores para as funções onde a variável não era combinada e que utilizaram a medida de penetração como variável preditora. Valores de *AIC* mais elevados para os modelos tradicionais (grupo a) que incluíram a altura total, se devem ao fato de que conforme observado no processo de coleta dos dados a relação entre diâmetro, altura e biomassa é muito complexa. No processo de coleta dos dados foram selecionadas espécies que possuíam grande variabilidade quanto à arquitetura e quanto à densidade básica. Desta forma, dentro da amostra, árvores com alturas muito semelhantes apresentavam diâmetros e densidades muito diferentes e a inclusão da informação da altura não se mostrou efetiva em agregar poder explicativo nos modelos tradicionais (grupo a).

Por outro lado, considerando os modelos onde a variável preditora resistência à penetração foi inserida, a inclusão da altura passa a ter significativo papel na explicação da variável biomassa total. Essa variável de acordo com os resultados parece capaz de classificar as espécies quanto sua densidade básica, tornando a altura total relevante na predição da biomassa total. Desta maneira, o melhor modelo para predição de biomassa total é, segundo os valores de *AIC* da Tabela 4.6, o modelo de Schumacher com inclusão da medida de penetração (4.b). O fato do modelo selecionado possuir erros gaussianos demonstra que a função considerada para descrever o comportamento de

σ foi efetiva em contornar o problema da heterocedasticidade da variância.

Considerando agora somente os modelos de predição da biomassa do tronco, pode-se notar que os modelos Log-normais obtiveram, de maneira geral, valores de AIC menores que os modelos gaussianos. Da mesma forma que para os modelos de predição da biomassa total, os modelos que incluíram a medida de penetração obtiveram melhor desempenho geral na predição da biomassa do tronco, principalmente nos modelos da classe Log normal. A inclusão da variável altura comercial para os modelos de predição da biomassa do tronco, assim como a inclusão da altura total nos modelos de predição da biomassa total, também só foi efetiva quando em conjunto com a variável resistência penetração. Desta forma, o modelo mais adequado dentre os testados é o modelo de Schumacher-Hall da classe Log-normal e com a inclusão da medida de penetração (7.b).

No que se refere, aos modelos de predição de galhos, os modelos que incluíram a medida de penetração (grupo b) obtiveram também melhor ajuste. No entanto, o que chama mais atenção, no ajuste desses modelos aos dados de biomassa de galhos, é o fato dos modelos da classe Log-normal terem se mostrado muito superiores que os gaussianos, já que, independente da função considerada para a média, estes tiveram valores de AIC menores que dos modelos gaussianos. Ao se observar os valores de AIC na Tabela 4.6, nota-se que três modelos da classe Log-normal, com diferentes funções de média e que incluíram a medida de penetração, obtiveram valores de AIC muito semelhantes. Estes foram os modelos diamétrico (AIC=1543, 323), Spurr (AIC=1546, 783) e Schumacher-Hall (AIC=1546, 338). Mas, como diferenças de AIC, entre modelos concorrentes, superiores a dois indicam que os modelos diferem quanto à qualidade do seu ajuste, o modelo diamétrico (5.b) é o mais adequado, entre todos os modelos testados.

Para os modelos de predição da biomassa de ramos, os modelos que incluíram a medida de penetração também obtiveram, na sua maioria, menores valores de AIC (Tabela 4.6). No entanto, o modelo diamétrico Log-normal que inclui a medida de penetração (5.b) se mostrou muito inferior em relação aos demais modelos considerados para predição desta variável. O melhor modelo para predição da biomassa de ramos foi o modelo diamétrico gaussiano (1.b). É importante notar que esses dois modelos só diferem quanto a classe, sendo que a função considerada para modelar a média é a mesma.

A altura total, mais uma vez, não foi eficaz em explicar as variações de biomassa de galhos e ramos nas diferentes observações, o que já era esperado, pois, a amostra composta de várias espécies, apresentava árvores de mesma altura, mas com tamanhos de copa muito distintos.

Por fim, considerando apenas os modelos de predição de biomassa das folhas, os piores modelos foram, segundo os valores da Tabela 4.6, os diamétricos que não incluem a altura total como variável preditora. O fato da inclusão da altura total melhorar a qualidade do ajuste destes modelos contradiz o observado em relação aos outros compartimentos componentes da copa da árvore (galhos e ramos). Essa contradição pode ser explicada pelo fato, já comentado, da coleta dos dados

ter se transcorrido por longo período de tempo e durante a época de queda de folhas de algumas espécies. Desta forma, uma boa relação esperada entre biomassa de galhos e biomassa das folhas não foi observada. É provável que isso tenha ocorrido devido ao fato de árvores de mesma espécie e com biomassa de galhos semelhantes, apresentarem diferentes valores de biomassa de folhas. A informação da medida de penetração, a exemplo de todos os compartimentos das árvores, se mostrou efetiva na melhoria do ajuste do modelo de predição de biomassa de folhas. Sendo assim, o melhor modelo para predição da biomassa de folhas foi o modelo de Schumacher-Hall com a inclusão da medida de penetração (4.b).

4.3.3 Análise dos modelos selecionados

A partir dos modelos selecionados estes foram avaliados no que se refere ao seu desempenho em relação a média e em relação ao comportamento gráfico dos seus resíduos. Os modelos selecionados, para predição da biomassa total e biomassa dos diferentes compartimentos, na sua forma de aplicação e com respectivo coeficiente de determinação empírico, são apresentados na Tabela 4.7.

Tabela 4.7 – Modelos selecionados, para predição da biomassa total e biomassa dos diferentes compartimentos, na sua forma de aplicação e com seu respectivo coeficiente de determinação empírico (R_e^2). Onde b é a biomassa total, b_t é a biomassa do tronco, b_g é a biomassa de galhos, b_r é a biomassa de ramos, b_f é a biomassa de folhas, d é o diâmetro a 1,30 m do solo, h é a altura total, hc é altura comercial e p é a medida de penetração no lenho

MODELO	FORMA DE APLICAÇÃO	R_e^2
Schumacher-Hall Gaussiano	$b = -2,40687d^{2,33325}h^{0,48620}p^{-0,37715}$	0,9479
Schumacher-Hall Log normal Diamétrico	$b_t = -2,49248 + 2,19684\ln(d) + 0,65486\ln(hc) - 0,41824\ln(p)$	0,9314
Log normal Diamétrico	$b_g = -4,23231 + 2,88043\ln(d) - 0,47757\ln(p)$	0,4962
Gaussiano	$b_r = -2,12251d^{2,16163}p^{-0,48472}$	0,6962
Schumacher Hall Gaussiano	$b_f = -1,80482d^{2,17953}h^{-0,83708}p^{-0,23314}$	0,4417

Todos os modelos selecionados foram do grupo que incluíram a medida de penetração, sendo que a aplicação destes, em qualquer circunstância, depende do uso de um penetrômetro. Isso torna a aplicabilidade dos resultados obtidos neste trabalho restrita à situações onde um equipamento

muito pouco usual esteja disponível. Desta forma, as estimativas dos coeficientes, com respectivo desvio padrão, de todos os modelos testados, bem como os valores de AIC desses modelos são apresentados nos anexos A, B, C, D e E.

Modelo de predição da biomassa total

Ao se observar o coeficiente de determinação empírico (R_e^2) do modelo de predição de biomassa total de 0,9479 (Tabela 4.7), pode-se afirmar que este modelo obteve um ótimo ajuste em relação à média amostral. A distribuição dos resíduos padronizados deste modelo pode ser considerada bem próxima da distribuição normal, já que observando a Figura 4.1, pode-se notar que os quantis amostrais estão bem próximos aos quantis da distribuição normal padronizada.

Uma análise do resíduo percentual do modelo selecionado (Figura 4.2), mostra que a amplitude do erro de predição para maioria das observações se encontra na faixa de $\pm 40\%$ em torno da média zero. Essa amplitude erro foi maior para as observações de menor biomassa e diminuiu para as árvores de maior biomassa, com uma tendência de superestimativa das observações de menor biomassa.

Este resultado está de acordo com o que foi observado por Santos et al. (2001) em estudo realizado na Amazônia ajustando, a mesma função para média só que na forma linearizada e sem a inclusão da variável preditora resistência à penetração. Outro resultado semelhante foi obtido por Reis et al. (1998) utilizando modelos de razão para predição da biomassa total de árvores individuais, também para região Amazônica. Os resultados mostram que o modelo selecionado é adequado e pode ser usado para realizar predições de biomassa total em áreas onde reflorestamentos com essas espécies tenham sido estabelecidos.

Modelo de predição da biomassa do tronco

Com relação à biomassa do tronco o modelo selecionado, obteve um valor de R_e^2 , um pouco inferior ao modelo de predição de biomassa total (0,9314), mas ainda sim, indicando um ótimo ajuste em relação à média amostral. (Tabela 4.7).

A distribuição dos resíduos padronizados do modelo apresentou também um bom comportamento, com os quantis amostrais bem próximos dos quantis teóricos da distribuição normal padronizada (Figura 4.3).

A análise gráfica do resíduo percentual (Figura 4.4) indica também um bom ajuste, com resultados semelhantes aos obtidos para o modelo de predição da biomassa total. Desta maneira, a amplitude do erro de predição ficou em torno de $\pm 40\%$ em torno da média zero, estando também de acordo com os resultados obtidos por Santos et al. (2001) e Reis et al. (1998) na predição deste compartimento para região amazônica. Desta forma, os resultados obtidos para o modelo selecionado indicam sua adequação para ser aplicado na predição da biomassa do tronco em áreas de

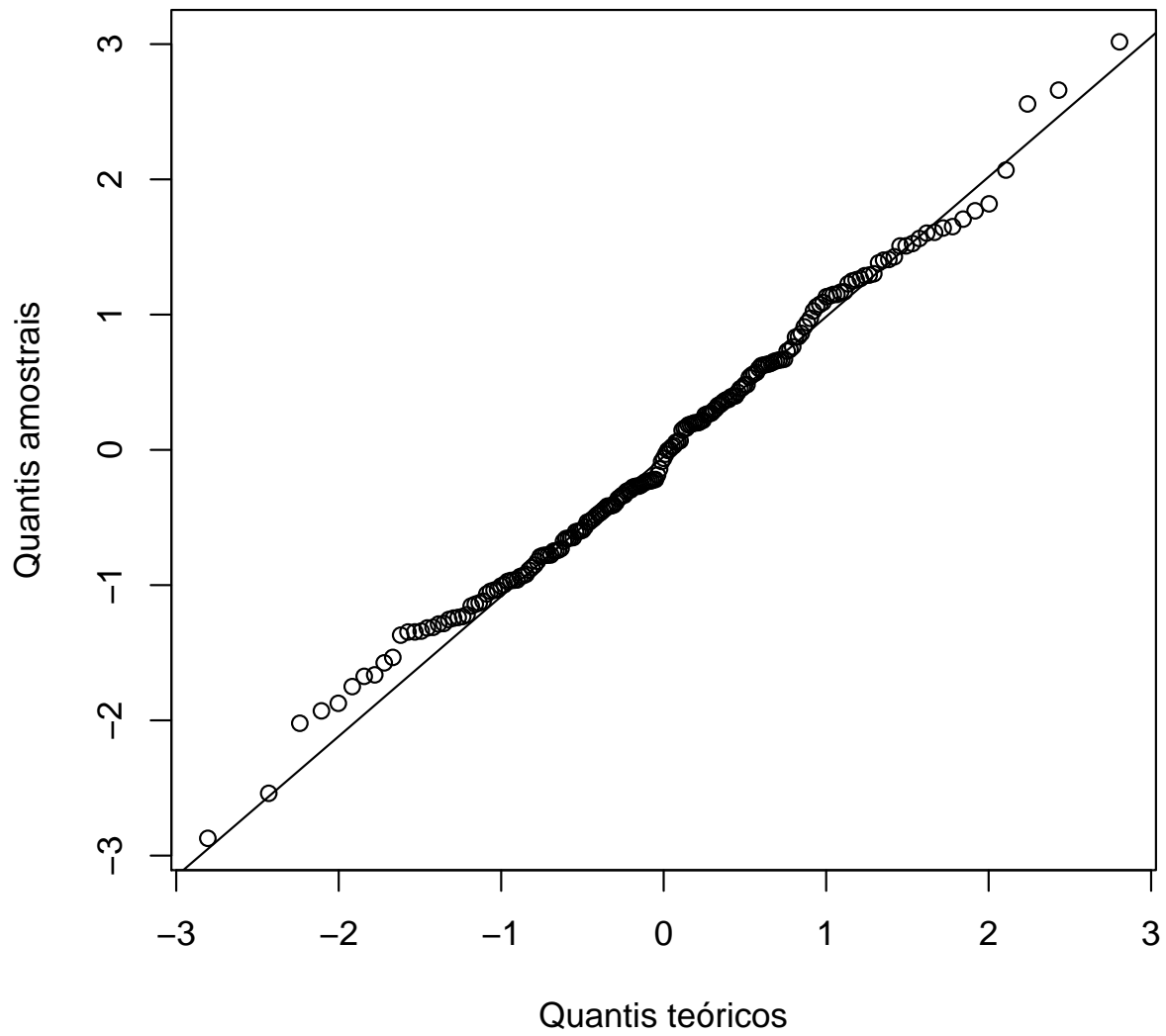


Figura 4.1 – Gráfico dos quantis dos resíduos padronizados do modelo de predição da biomassa total em relação aos quantis teóricos da distribuição normal padronizada

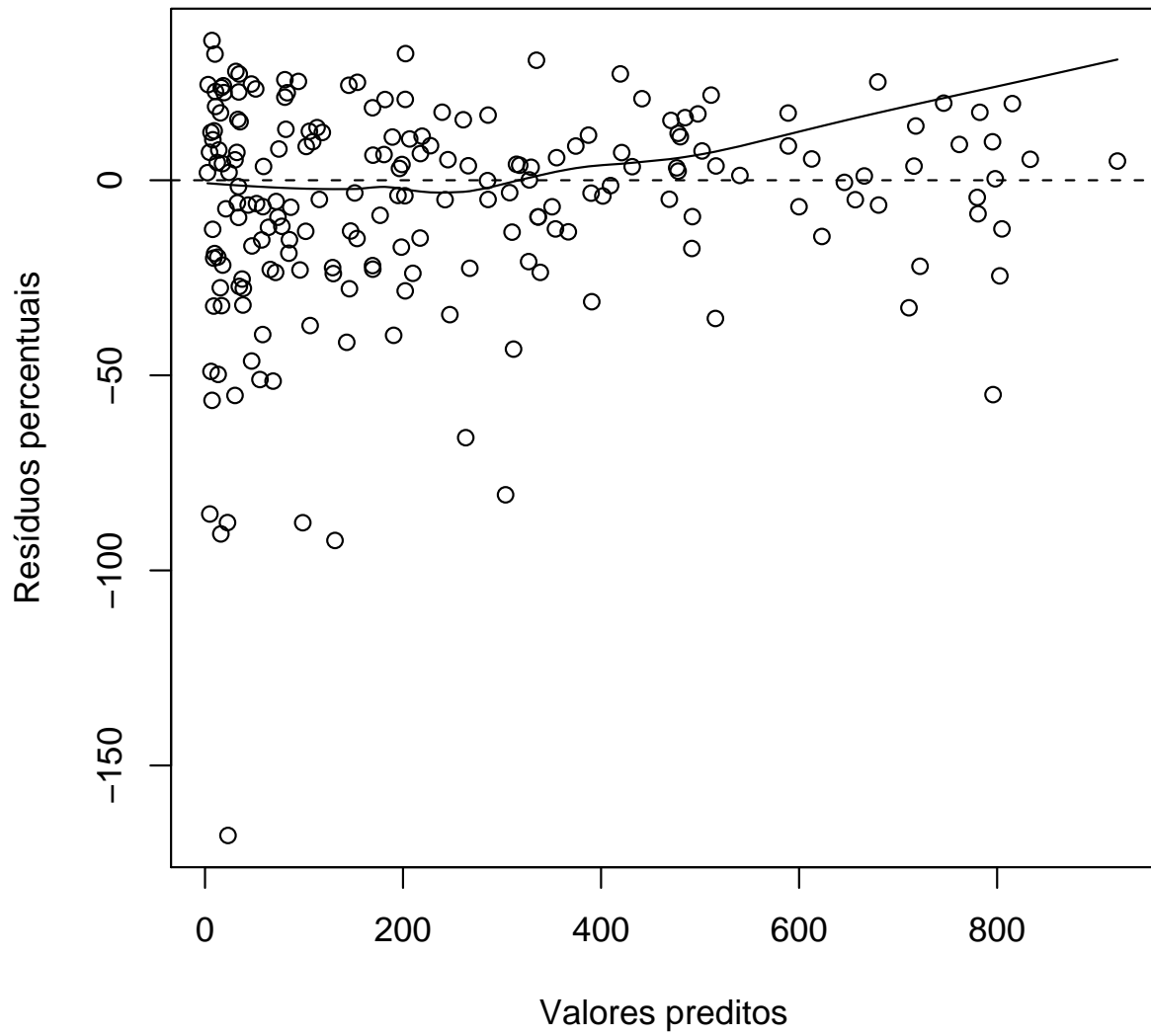


Figura 4.2 – Gráfico de dispersão dos resíduos percentuais em função dos valores ajustados do modelo de predição da biomassa total

reflorestamentos com espécies tropicais originárias da mata Atlântica.

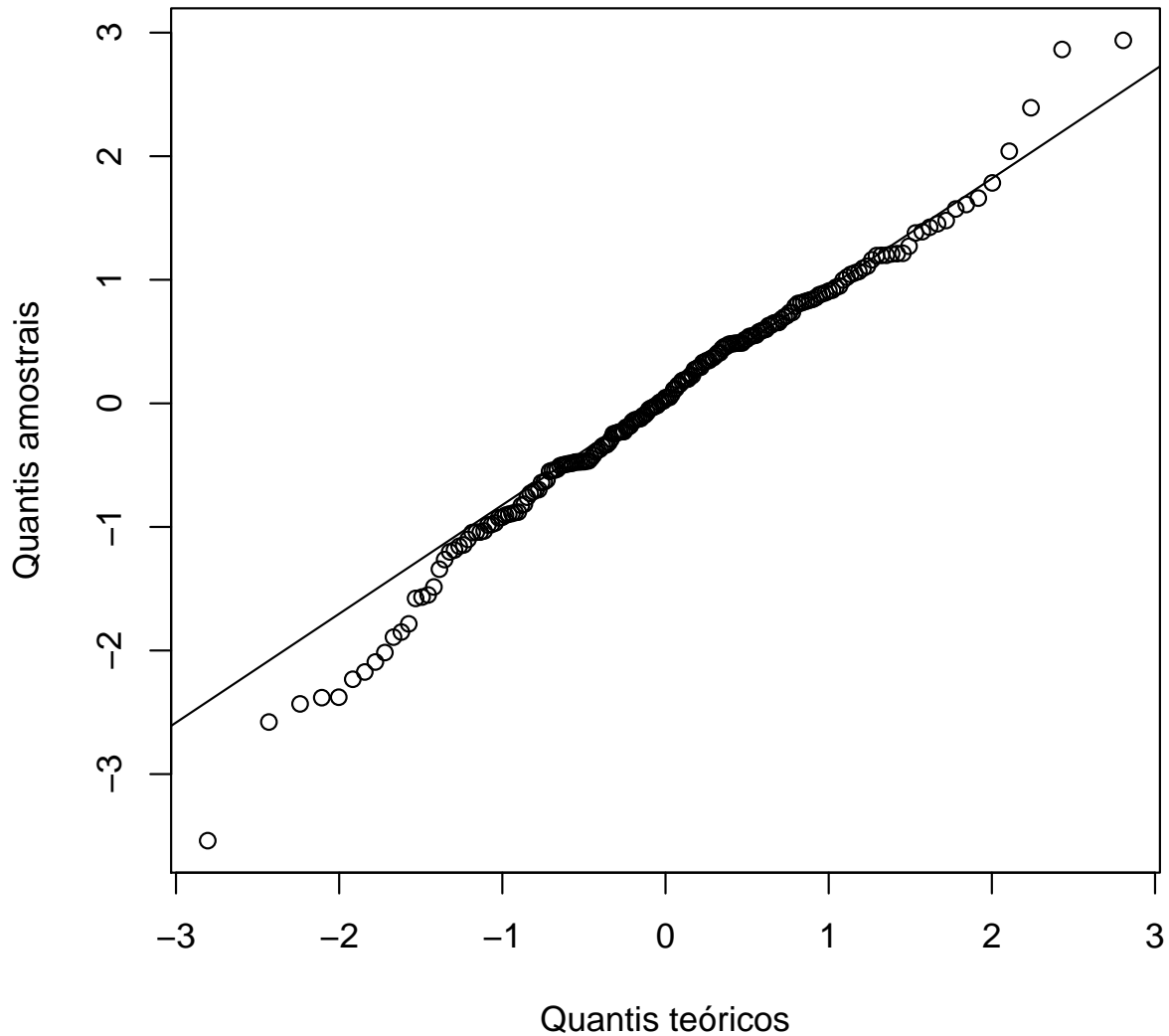


Figura 4.3 – Gráfico dos quantis dos resíduos padronizados do modelo de predição da biomassa do tronco em relação aos quantis teóricos da distribuição normal padronizada

Modelo de predição da biomassa de galhos

O desempenho do modelo selecionado para predição da biomassa de galhos no que se refere ao R_e^2 de 0,4962, foi bem inferior se comparado aos modelos de predição de biomassa total e biomassa

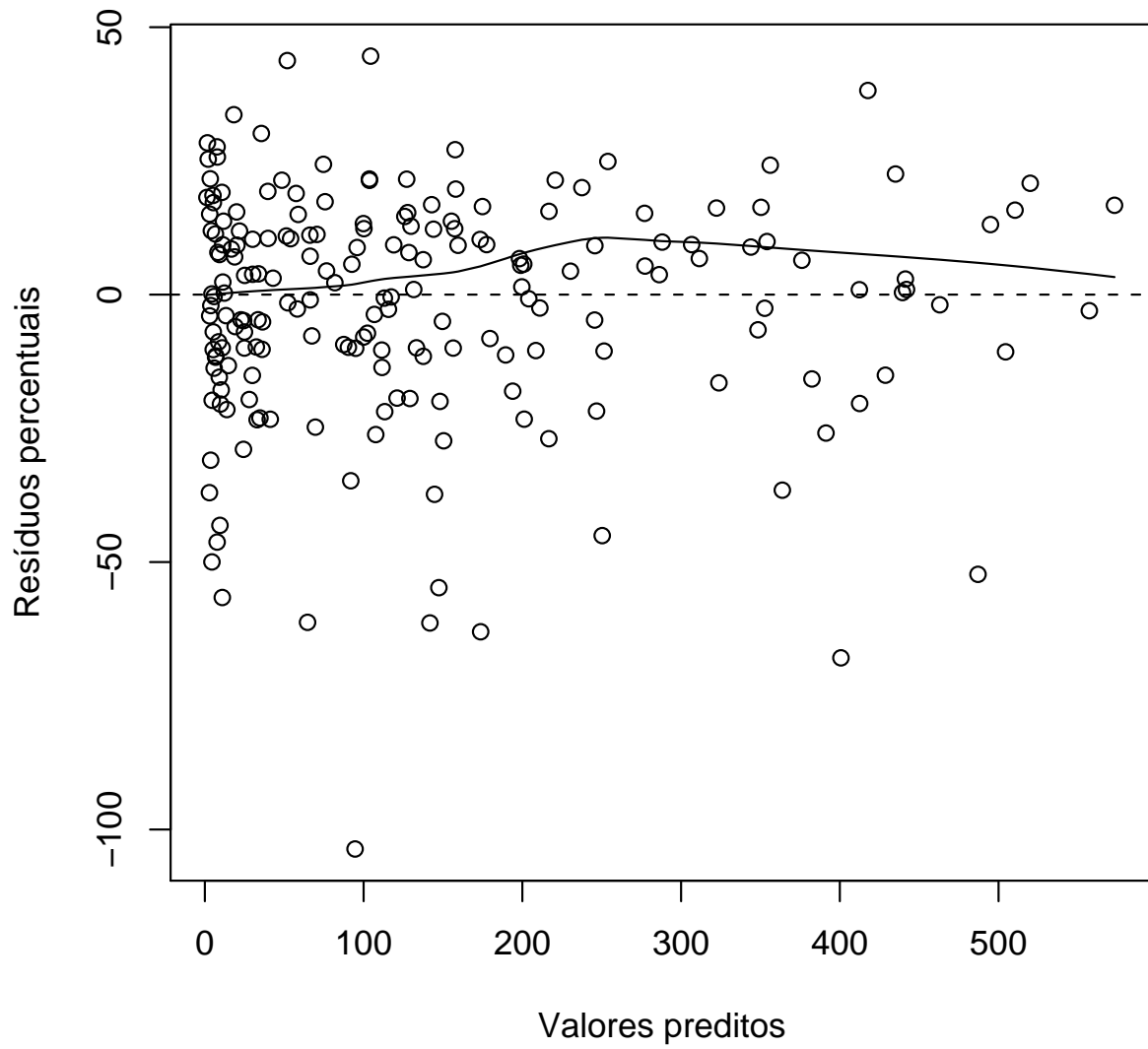


Figura 4.4 – Gráfico de dispersão dos resíduos percentuais em função dos valores ajustados do modelo de predição da biomassa do tronco

do tronco (Tabela 4.7). Este resultado indica que o uso das variáveis preditoras DAP e resistência à penetração foi pouco efetivo para explicar a variabilidade da biomassa de galhos.

O comportamento da distribuição dos resíduos padronizados deste modelo (Figura 4.5) indicam uma ligeira assimetria à esquerda na cauda da distribuição normal.

A análise gráfica do resíduo percentual (Figura 4.6) indica que a amplitude do erro de predição para a maioria das observações ficou em torno de $\pm 100\%$ em torno da média zero. Além disso, a análise gráfica aponta para uma tendência de superestimação principalmente das observações de menor valor de biomassa. Sendo assim, os resultados obtidos para o modelo selecionado indicam que aplicação do modelo selecionado é problemática, principalmente no que se refere às árvores de menor tamanho.

Modelo de predição da biomassa de ramos

Observando a Tabela 4.6 nota-se que o valor do R_e^2 , obtido para o modelo de predição da biomassa de ramos selecionado de 0,6962 é bastante razoável (Tabela 4.7). Este resultado se torna ainda mais relevante se for considerada a baixa relação esperada entre as variáveis preditoras utilizadas (dap e resistência à penetração) com a biomassa dos ramos, principalmente se for considerada a variação natural das arquiteturas de copa das espécies estudadas. No entanto, ao observar a distribuição do resíduo padronizado, verifica-se uma assimetria a direita importante (Figura 4.7), que aponta para um comportamento do resíduo do modelo que escapa à normalidade.

A distribuição dos resíduos percentuais mostra uma dispersão muito elevada em torno da média com valores de até 1500 % para algumas observações (Figura 4.8), além de tendência de superestimativa para árvores pequenas. Santos et al. (2001), estabelecendo equações de predição da biomassa da copa obteve modelos com alto valor de R^2 , mas com comportamento do resíduo semelhante aos obtidos neste trabalho na predição da biomassa dos ramos. Seus resultados apontam para erros de $\pm 200\%$ em modelos de predição da biomassa da copa a partir do diâmetro e altura total, com valor de R^2 de 0,8882, mostrando a dificuldade de se obter modelos precisos de predição desse tipo de variável a partir das variáveis dendrométricas tradicionais ou com inclusão de medida de penetração.

Modelo de predição da biomassa de folhas

Com relação ao modelo de predição de biomassa de folhas selecionado, seu ajuste é o pior em relação aos modelos de predição dos demais componentes da copa considerados neste trabalho. O valor do seu R_e^2 (Tabela 4.7) é bem próximo ao modelo de predição da biomassa de galhos, mas o comportamento do seu resíduo é muito inferior. A partir da análise do gráfico do resíduo padronizado (Figura 4.9), é possível observar uma acentuada assimetria à direita, indicando um importante desvio em relação à normalidade.

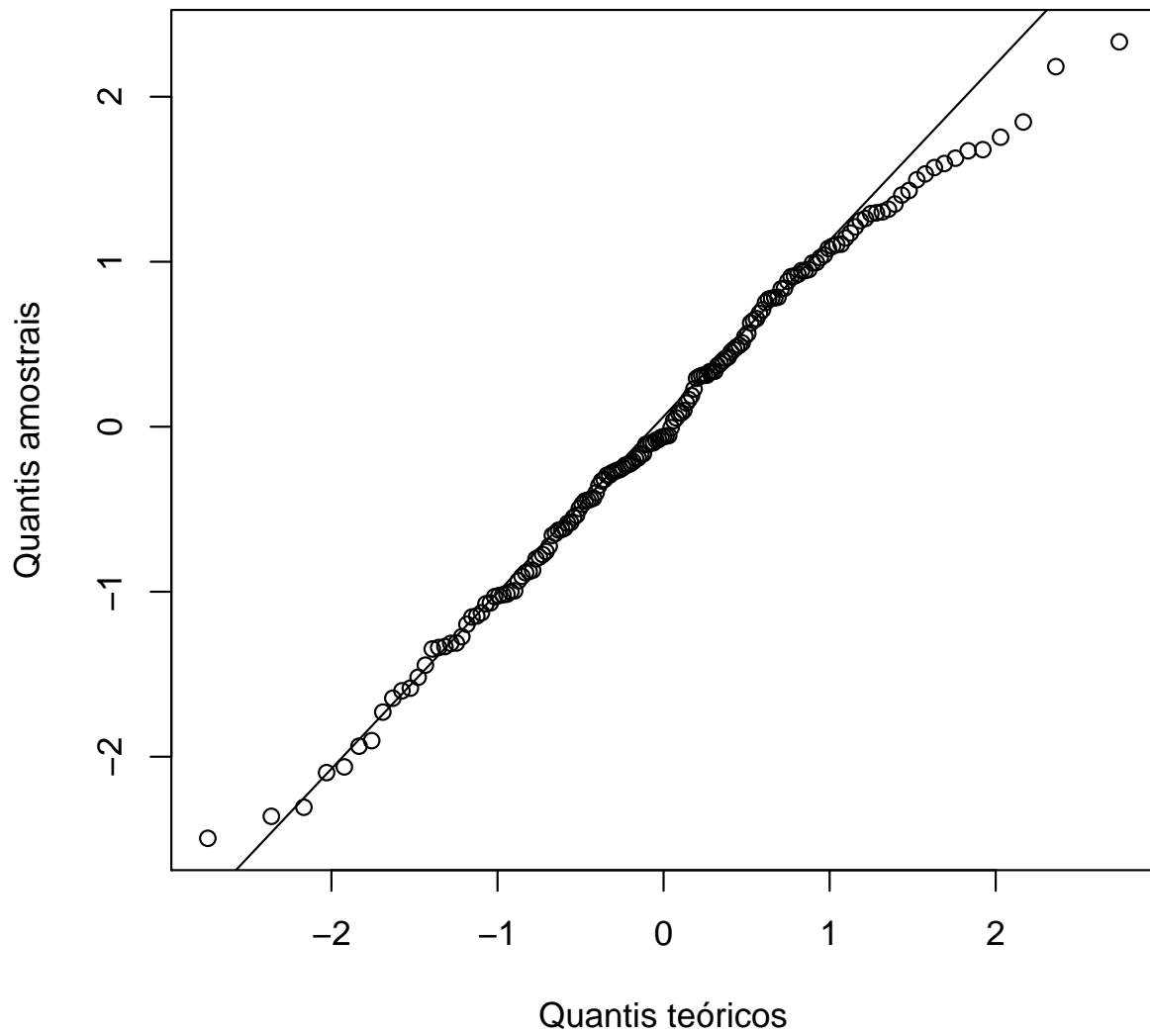


Figura 4.5 – Gráfico dos quantis dos resíduos padronizados do modelo de predição da biomassa de galhos em relação aos quantis teóricos da distribuição normal padronizada

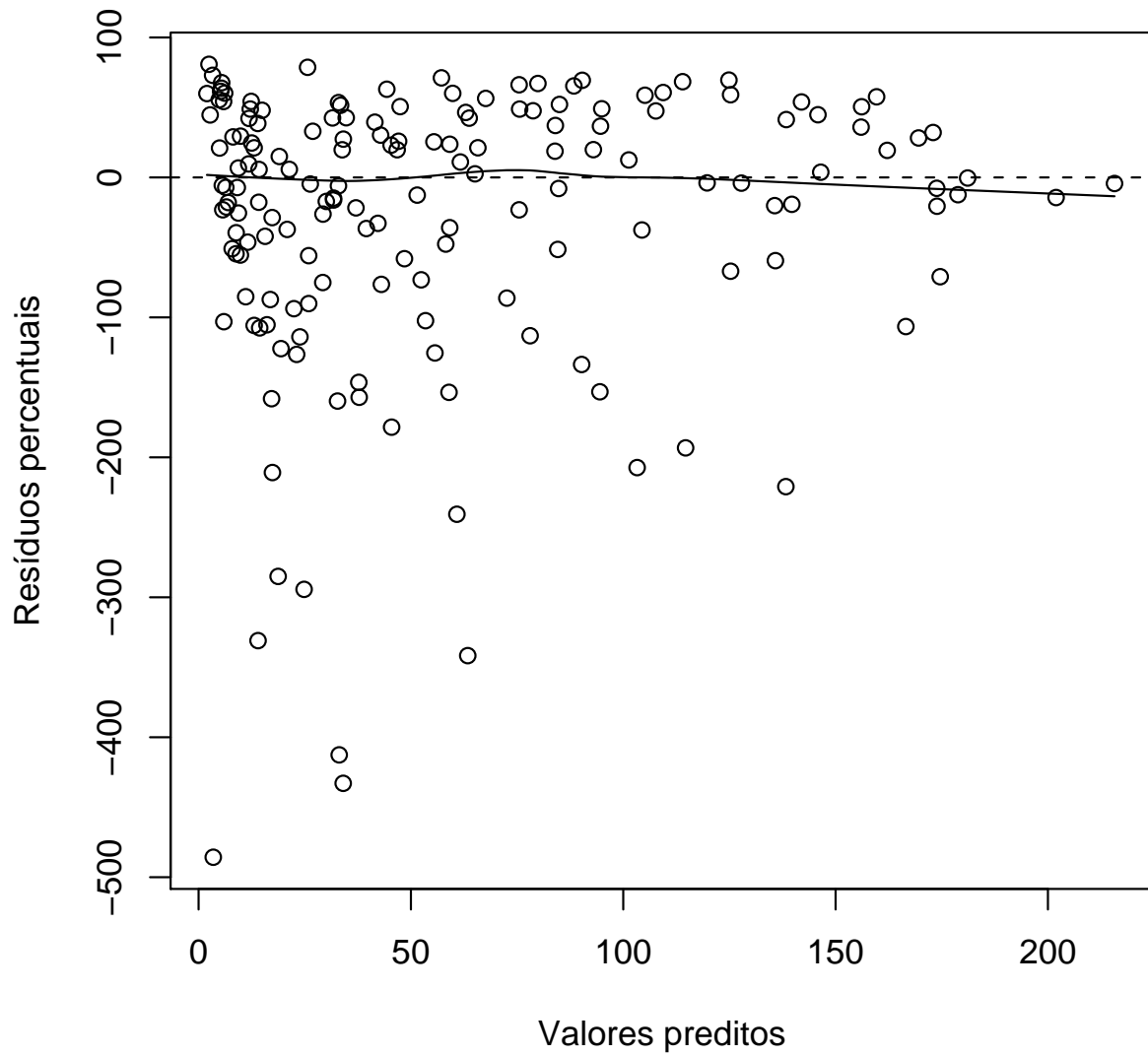


Figura 4.6 – Gráfico de dispersão dos resíduos percentuais em função dos valores ajustados do modelo de predição da biomassa de galhos

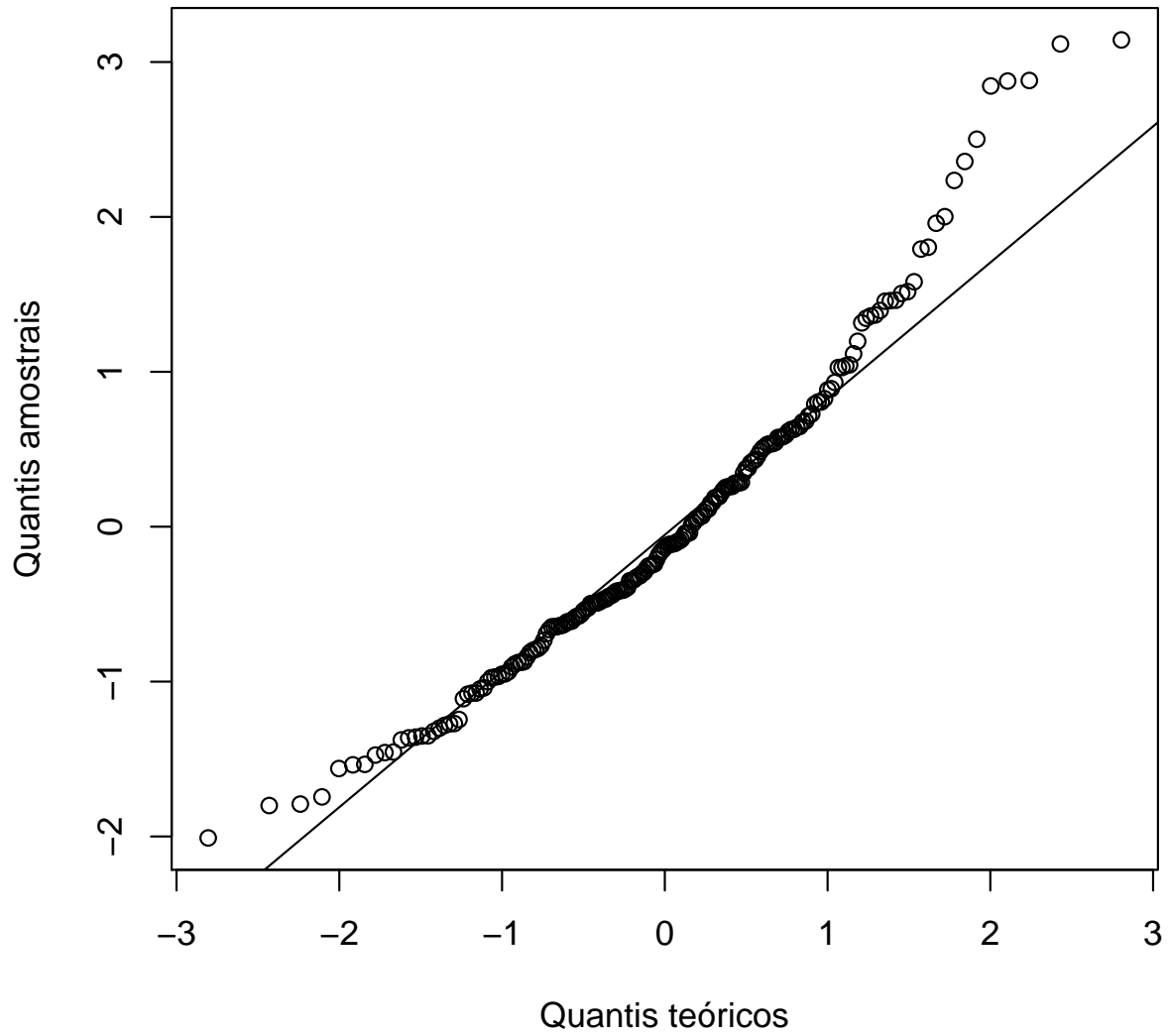


Figura 4.7 – Gráfico dos quantis dos resíduos padronizados do modelo de predição da biomassa de ramos em relação aos quantis teóricos da distribuição normal padronizada

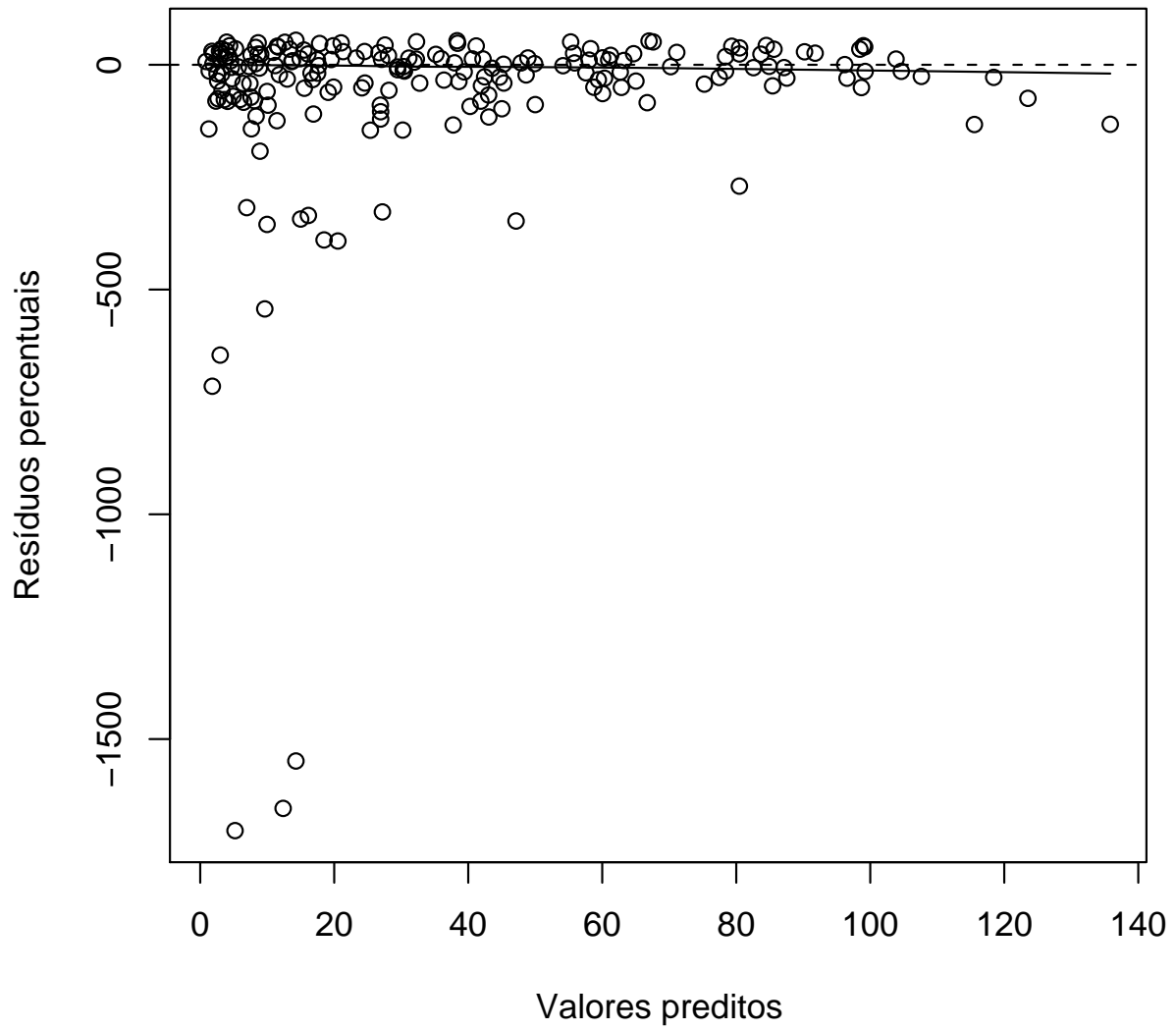


Figura 4.8 – Gráfico de dispersão dos resíduos percentuais em função dos valores ajustados do modelo de predição da biomassa de ramos

No entanto, o comportamento do resíduo percentual é ainda mais problemático. A Figura 4.10 mostra que os resíduos estão acentuadamente mais dispersos em torno da média que os resíduos dos modelos de predição da biomassa de galhos e ramos. Esse comportamento de dispersão dos resíduos invalida o modelo selecionado para aplicação em predições de biomassa em qualquer circunstância.

O fato do modelo de predição de biomassa das folhas obtido ter tido desempenho tão aquém em relação aos demais modelos, pode ter relação com o fato, já discutido, de que a coleta se transcorreu durante o período de queda de folhas de algumas espécies, não possibilitando o estabelecimento de relação alguma entre o tamanho das árvores e o conteúdo de folhas da mesma.

4.4 Conclusões

A variável preditora resistência à penetração no lenho se mostrou efetiva em incrementar o poder explicativo dos modelos independente das relações funcionais consideradas para descrever a média no presente trabalho. A relevância desta variável é reforçada pelo fato de todos os modelos selecionados, neste trabalho, para predição da biomassa nos diferentes compartimentos considerados, tiveram como variável preditora, entre outras a resistência à penetração no lenho.

De maneira geral, modelar o desvio padrão a partir de uma relação de potência com o DAP se mostrou mais eficiente em contornar o problema de heteroscedasticidade da variância, que considerar o modelo Log-normal para descrever o comportamento de variação dos resíduos. Isto pode ser afirmado, já que, na maioria dos casos, os modelos com distribuição gaussiana dos resíduos tiveram melhor comportamento que os modelos com distribuição Log-normal na análise feita pelo *AIC*. A única exceção foi o modelo selecionado para predição da biomassa do tronco, onde o modelo Log-normal foi superior ao modelo gaussiano no que se refere à qualidade do seu ajuste.

Os modelos de Schumacher tiveram desempenho superior aos modelos de variável combinada de Spurr. A variável altura, total ou do tronco, não foi efetiva em explicar a variabilidade da biomassa total ou da biomassa do tronco da árvore sem que no modelo estivesse presente a variável resistência à penetração. Desta forma, a variável altura, considerando diferentes espécies com diferentes relações alométricas, só teve relação com a biomassa da árvore a partir da introdução de uma variável relacionada à densidade básica do lenho.

Os modelos selecionados para biomassa total e biomassa do tronco apresentaram bom ajuste e podem ser utilizados como ferramenta de predição destes atributos em áreas de reflorestamento com espécies originárias da floresta Atlântica. Já quanto aos modelos de predição de biomassa de galhos e ramos, estes não obtiveram o ajuste adequado e sua utilização é mais restrita. Por fim, o modelo selecionado para predição de biomassa de folhas, devido ao problema relacionado à queda natural das folhas durante o processo de coleta dos dados, teve sua aplicação inviabilizada.

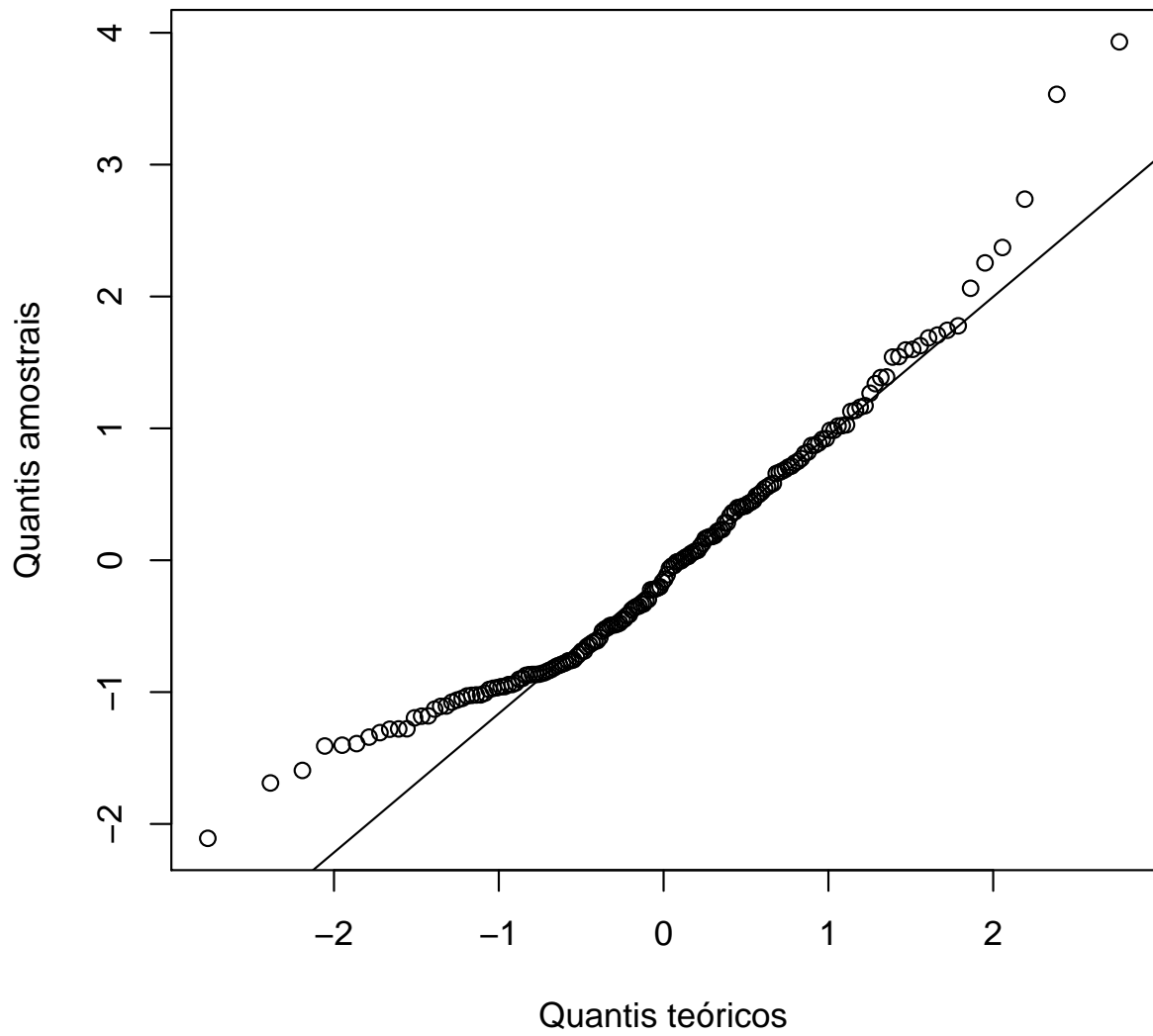


Figura 4.9 – Gráfico dos quantis dos resíduos padronizados do modelo de predição da biomassa de folhas em relação aos quantis teóricos da distribuição normal padronizada

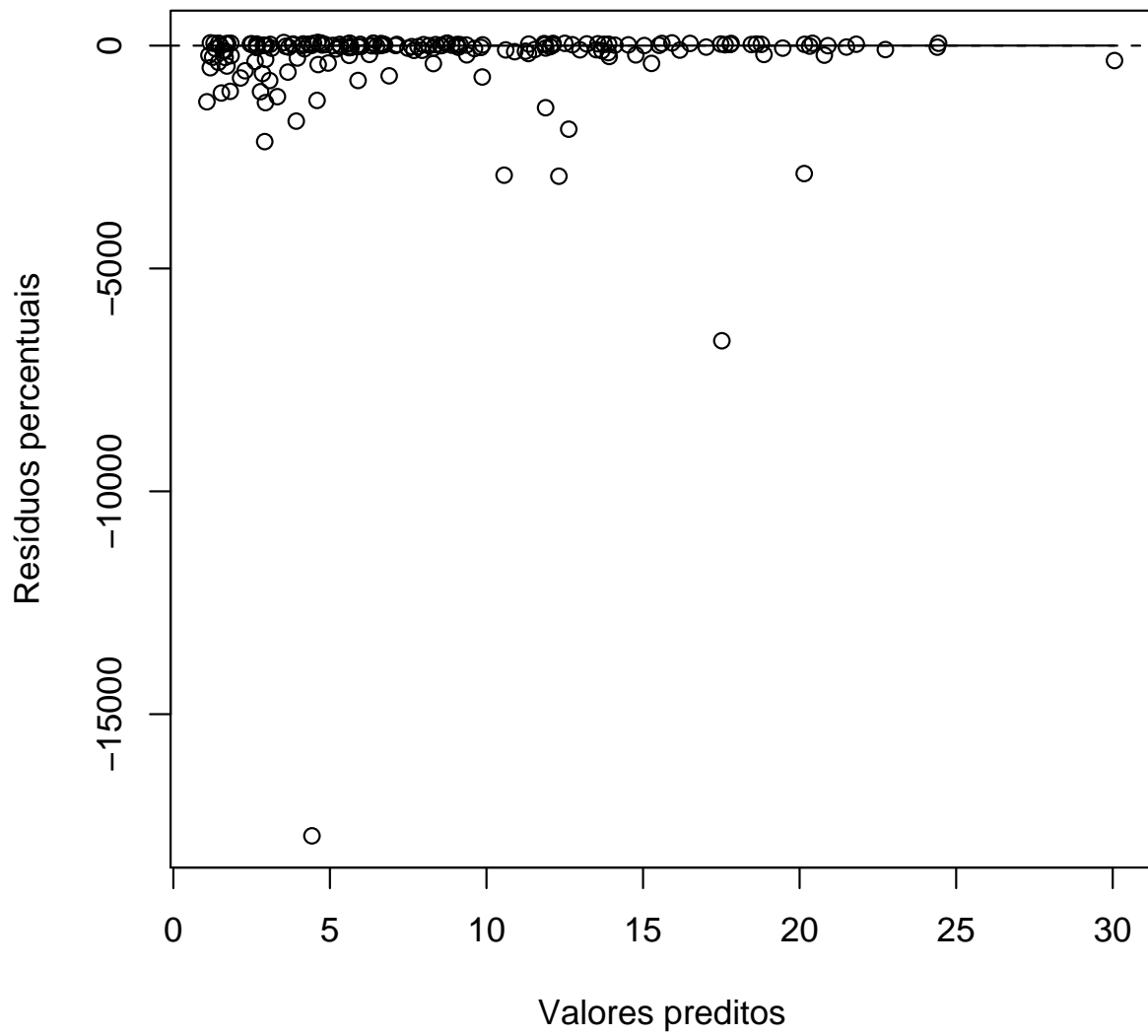


Figura 4.10 – Gráfico de dispersão dos resíduos percentuais em função dos valores ajustados do modelo de predição da biomassa de folhas

Referências

- AKAIKE, H. A new look at statistical model identification, **IEEE Transactions on Automatic Control**, Tokio, v. 19, n. 6, p. 717-723, Dec. 1974.
- AKAIKE, H. Likelihood of a model and information criteria, **Journal of Econometrics**, Amsterdam, v. 16, n.1, p. 3-14, May. 1981.
- ARAÚJO, T.M.; HIGUCHI, N.; CARVALHO JR., J.A. Comparação de métodos para determinar biomassa na região amazônica, **Anais da Academia Brasileira de Ciência**, Rio de Janeiro, v. 68, n. 1, p. 35-41, ago. 1996.
- BROWN, S.; GILLESPIE, A.J.R.; LUGO, A.E. Biomass estimation methods for tropical forest with applications to inventory data, **Forest Science**, Maryland, v. 35, n. 4, p. 881-902, Dec. 1989.
- BURNHAM, K.P.; ANDERSON, D.R. Model selection and multimodel inference: Practical information theoretic approach. 2nd.ed. New York: Springer, 2002. 488 p.
- FEARNSIDE, P.M. Wood density for estimating forest biomass in Brazilian Amazonia. **Forest Ecology and Management**, Winsconsin, n. 90, p. 59-87. Oct. 1997.
- HIGUCHI, N.; CARVALHO Jr., J. A. Fitomassa e conteúdo de carbono de espécies arbóreas da Amazônia. In: SEMINÁRIO DE EMISSÃO x SEQUESTRO DE CO_2 : UMA NOVA OPORTUNIDADE DE NEGÓCIOS PARA O BRASIL. 1994. Rio de Janeiro. **Anais...** Rio de Janeiro: Companhia Vale do Rio Doce, 1994. p. 125-153.
- HIGUCHI, N.; SANTOS, J.; RIBEIRO, R.J.; MINETTE, Y.B.M. Biomassa da parte aérea da vegetação da floresta tropical úmida de terra-firme da Amazônia brasileira, **Acta Amazônica**, Manaus, v. 28 n. 2, p. 153-166, mar. 1998.
- NELSON, B.W.; MESQUITA, R.; PEREIRA, J.L.G.; SOUZA, S.G.S.A.; BATISTA, G.T.; COUTO, L.B. Allometric regressions for improved estimate of secondary forest biomass in the central Amazon, **Forest Ecology and Management**, Winsconsin, n. 117, p. 149-167, Ago. 1998.
- NOGUEIRA, E.M.; FEARNESIDE, P.M.; NELSON, B. Wood density in forests of Brazil's "arc of deforestation": Implications for biomass and flux of carbon from land-use change in Amazonia.

Forest Ecology and Management, Wisconsin, n. 248, p. 119-135, Mar. 2007.

MOOD, A.M; GRAYBILL, F.A.; BOES, D.C. Introduction to the theory of statistics. 3rd.ed. Auckland: MacGraw-Hill, 1974. 564p.

PARRESOL, B.R. Assessing tree and stand biomass: A review with examples and critical comparisons. **Forest Science**, Lawrence, v. 45, n. 4, p. 573-593, Aug. 1999.

REIS, E.; PAULA NETO, F.; SOUZA, L. A.; LEITE, H.G. Modelos de razão para estimar a biomassa de árvores individuais da floresta tropical úmida de terra-firme da Amazônia brasileira, **Revista árvore**, Viçosa, v. 22, n. 2, p. 163-177, ago. 1998.

SANQUETTA, C.R. Métodos de determinação de biomassa florestal. In: SANQUETTA, C.R.; WATZLAWICK, L.F.; BALBINOT, R.; ZILLIOTO, M.A.B.; GOMES, F.S. (Ed.). **As florestas e o carbono**: [s.n.], 2002, cap. 6, p. 119-140.

SANTOS, J.; PAULA NETO, F.; HIGUSHI, N.; LEITE, H.G.; SOUZA, L.A.; VALE, B.A. Modelos estatísticos para estimar a fitomassa acima do nível do solo da floresta tropical úmida da Amazônia central, **Revista árvore**, Viçosa, v. 25, n. 4, p. 445-454, out. 2001.

SEGURA, M. Allometric models for tree volume and total aboveground biomass in a tropical humid forest in Costa Rica, **Biotropica**, Baton Rouge, v. 37, n. 1, p. 2-8, Oct. 2005.

SOARES, C.P.B.; OLIVEIRA, M.L.R. Equações para estimar o carbono na parte aérea de árvores de eucalipto em Viçosa, Minas Gerais, **Revista Árvore**, Viçosa, v. 26, n. 5, p. 533-539, set. 2002.

VIEIRA, S.A.; ALVES, L.F.; AIDAR, M.; ARAUJO, L.S.; BAKER, T.; BATISTA, J.L.F.; CAMPOS, M.C.; CAMARGO, P.B.; CHAVE, J.; DELLITI, W.B.C.; HIGUSHI, N.; HONORIO, E.; JOLY, C.A.; KELLER, M.; MARTINELLI, L.A.; MATTOS, E.A.; METZKER, T.; PHILLIPS, O.; SANTOS, F.A.M.; SHIMABUKURU, M.T.; SILVEIRA, M.; TRUMBORE, S.E. Estimation of biomass and carbon stocks: The case of Atlantic Forest, **Biota Neotropica**, São Paulo, v. 8, n. 2, p. 21-29, jun. 2008.

ANEXOS

ANEXO A - Estimativas dos coeficientes com respectivo desvio padrão (entre parênteses), dos diversos modelos testados para predição da biomassa total

Modelo	Classe	Coeficientes				AIC
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	
1.a - Diamétrico	Gaussiano	-3,8627 (0,1290)	2,8861 (0,0418)	-	-	1879,163
2.a - Spurr	Gaussiano	1,7992 (0,7064)	0,0264 (0,006)	-	-	2027,124
3.a - Spurr-potência	Gaussiano	-3,4047 (0,1742)	0,9769 (0,0191)	-	-	2031,657
4.a - Schumacher-Hall	Gaussiano	-3,0729 (0,1529)	2,4470 (0,0965)	0,2969 (0,1214)	-	2004,986
5.a - Diamétrico	Log Normal	-3,1100 (0,1413)	2,7061 (0,0478)	-	-	2054,754
6.a - Spurr	Log Normal	-3,4565 (0,1536)	0,9765 (0,01798)	-	-	2070,334
7.a - Schumacher-Hall	Log Normal	-3,2527 (0,1519)	2,4634 (0,11355)	0,3271 (0,1388)	-	2051,280
1.b - Diamétrico	Gaussiano	-9,2781 (0,7581)	3,4533 (0,4527)	8,5461 (0,6542)	-	1904,650
2.b - Spurr	Gaussiano	4,8130 (0,8382)	0,0262 (0,0005)	-0,2186 (0,0451)	-	2007,262
3.b - Spurr-potência	Gaussiano	-2,6000 (0,1321)	0,9864 (0,0137)	-0,4034 (0,0285)	-	1890,945
4.b - Schumacher-Hall	Gaussiano	-2,4068 (0,1173)	2,3332 (0,06535)	0,4862 (0,0880)	-0,3771 (0,0271)	1855,158
5.b - Diamétrico	Log Normal	-2,1337 (0,1160)	2,7256 (0,03264)	-0,4590 (0,0304)	-	2072,081
6.b - Spurr	Log Normal	-2,4203 (0,1161)	0,9871 (0,0115)	-0,5017 (0,0297)	-	1895,654
7.b - Schumacher-Hall	Log Normal	-2,3182 (0,1110)	2,3270 (0,0726)	0,5370 (0,0892)	-0,4828 (0,0282)	1873,350

ANEXO B - Estimativas dos coeficientes com respectivo desvio padrão (entre parênteses), dos diversos modelos testados para predição da biomassa do tronco

Modelo	Classe	Coeficientes				AIC
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	
1.a - Diamétrico	Gaussiano	-3,6121 (0,1550)	2,6999 (0,0501)	-	-	1838,663
2.a - Spurr	Gaussiano	1,2737 (0,3182)	0,0297 (0,0008)	-	-	1826,612
3.a - Spurr-potência	Gaussiano	-2,8131 (0,1280)	0,9184 (0,0158)	-	-	1818,634
4.a - Schumacher-Ha	Gaussiano	-3,4504 (0,1371)	2,3146 (0,0561)	0,5009 (0,0422)	-	1748,758
5.a - Diamétrico	Log Normal	-3,8347 (0,1529)	2,7525 (0,0517)	-	-	1850,776
6.a - Spurr	Log Normal	-2,9435 (0,1234)	0,9277 (0,0157)	-	-	1813,041
7.a - Schumacher-Hall	Log Normal	-3,4559 (0,1272)	2,2741 (0,0611)	0,5469 (0,0514)	-	1763,352
1.b - Diamétrico	Gaussiano	-3,8862 (0,3262)	0,9032 (0,0259)	-2,994 (0,0417)	-	2286,295
2.b - Spurr	Gaussiano	2,4328 (0,3550)	0,0301 (0,008)	-0,1039 (0,0206)	-	1806,931
3.b - Spurr-potência	Gaussiano	-2,1955 (0,0891)	0,9456 (0,0104)	-0,4150 (0,0301)	-	1669,172
4.b - Schumacher-Hall	Gaussiano	-2,5437 (0,0923)	2,1970 (0,0420)	0,6326 (0,0401)	-0,3746 (0,0267)	1618,510
5.b - Diamétrico	Log Normal	-3,1479 (0,1619)	2,7651 (0,0455)	-0,3235 (0,0424)	-	1801,746
6.b - Spurr	Log Normal	-2,0282 (0,1024)	0,9423 (0,0106)	-0,4576 (0,0297)	-	1659,032
7.b - Schumacher-Hall	Log Normal	-2,4923 (0,1023)	2,1968 (0,0403)	0,6548 (0,0343)	-0,4182 (0,0257)	1597,267

ANEXO C - Estimativas dos coeficientes com respectivo desvio padrão (entre parênteses), dos diversos modelos testados para predição da biomassa de galhos

Modelo	Classe	Coeficientes				AIC
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	
1.a - Diamétrico	Gaussiano	-5,0403 (0,5202)	2,8775 (0,1667)	-	-	1642,649
2.a - Spurr	Gaussiano	-0,6461 (1,0186)	0,006 (0,0004)	-	-	1652,628
3.a - Spurr-potência	Gaussiano	-8,3672 (0,4534)	1,356 (0,0499)	-	-	1650,789
4.a - Schumacher-Hall	Gaussiano	-5,6552 (0,5182)	2,838 (0,3230)	0,2679 (0,3835)	-	1647,992
5.a - Diamétrico	Log Normal	-5,3931 (0,4874)	2,905 (0,1567)	-	-	1560,237
6.a - Spurr	Log Normal	-6,1304 (0,5423)	1,087 (0,0604)	-	-	1566,634
7.a - Schumacher-Hall	Log Normal	-5,4966 (0,5858)	2,831 (0,2809)	0,1210 (0,3801)	-	1562,136
1.b - Diamétrico	Gaussiano	-4,1914 (0,5182)	2,976 (0,1605)	-0,5295 (0,1088)	-	1622,580
2.b - Spurr	Gaussiano	0,6040 (1,4468)	0,006 (0,0004)	-0,1278 (0,1055)	-	1653,100
3.b - Spurr-potência	Gaussiano	1,3789 (0,7243)	-1,4702 (0,6947)	-0,9796 (0,3844)	-	1817,703
4.b - Schumacher-Hall	Gaussiano	-4,3593 (0,5729)	2,8203 (0,2724)	0,2482 (0,3390)	-0,5429 (0,1116)	1624,040
5.b - Diamétrico	Log Normal	-4,2323 (0,5283)	2,8804 (0,1481)	-0,4775 (0,1066)	-	1543,323
6.b - Spurr	Log Normal	-4,8716 (0,5586)	1,0873 (0,0560)	-0,5528 (0,1067)	-	1546,783
7.b - Schumacher-Hall	Log Normal	-4,5896 (0,5828)	2,5606 (0,2699)	0,5183 (0,3668)	-0,5130 (0,1089)	1546,338

ANEXO D - Estimativas dos coeficientes com respectivo desvio padrão (entre parênteses), dos diversos modelos testados para predição da biomassa de ramos

Modelo	Classe	Coeficientes				AIC
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	
1.a - Diamétrico	Gaussiano	-3,1296 (0,2521)	2,1443 (0,0825)	-	-	1585,237
2.a - Spurr	Gaussiano	1,9597 (0,3753)	0,0038 (0,0001)	-	-	1622,236
3.a - Spurr-potência	Gaussiano	-3,3241 (0,2929)	0,7664 (0,0332)	-	-	1616,841
4.a - Schumacher-Hall	Gaussiano	-2,9753 (0,2401)	2,3915 (0,1827)	-0,3346 (0,2145)	-	1594,528
5.a - Diamétrico	Log Normal	-3,6142 (0,3085)	2,2366 (0,1043)	-	-	1620,329
6.a - Spurr	Log Normal	-3,7993 (0,3351)	0,7948 (0,0392)	-	-	1635,604
7.a - Schumacher-Hall	Log Normal	-3,3831 (0,3339)	2,6301 (0,2494)	-0,5290 (0,3050)	-	1619,343
1.b - Diamétrico	Gaussiano	-2,1225 (0,2133)	2,1616 (0,0651)	-0,4847 (0,0618)	-	1531,580
2.b - Spurr	Gaussiano	3,2898 (0,4995)	0,0038 (0,0001)	-0,1049 (0,0269)	-	1608,678
3.b - Spurr-potência	Gaussiano	-2,2698 (0,2382)	0,7764 (0,0259)	-0,5227 (0,0658)	-	1561,481
4.b - Schumacher-Hall	Gaussiano	-2,0213 (0,2145)	2,3745 (0,1515)	-0,2931 (0,1876)	-0,4664 (0,0617)	
5.b - Diamétrico	Log Normal	-1,3636 (33,332)	5,1866 (65,607)	-11,011 (61,074)	-	3987,563
6.b - Spurr	Log Normal	-2,0152 (0,3144)	0,8132 (0,0312)	-0,8639 (0,0805)	-	1546,768
7.b - Schumacher-Hall	Log Normal	-1,7922 (0,3073)	2,3979 (0,2011)	-0,1714 (0,2468)	-0,8224 (0,0781)	1534,232

ANEXO E - Estimativas dos coeficientes com respectivo desvio padrão (entre parênteses), dos diversos modelos testados para predição da biomassa de folhas

Modelo	Classe	Coeficientes				AIC
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	
1.a - Diamétrico	Gaussiano	-0,9840 (0,1465)	-5,2095 (0,1263)	-	-	1252,669
2.a - Spurr	Gaussiano	1,7676 (2,9322)	0,0007 (0,0001)	-	-	1096,001
3.a - Spurr-potência	Gaussiano	-2,7261 (0,4405)	0,5485 (0,0493)	-	-	1093,535
4.a - Schumacher-Hall	Gaussiano	-2,2802 (0,3371)	2,1672 (0,2572)	-0,8327 (0,2990)	-	1076,370
5.a - Diamétrico	Log Normal	-1,8101 (78,533)	0,8156 (26,482)	-	-	2661,127
6.a - Spurr	Log Normal	-3,5792 (0,5461)	0,5916 (0,0637)	-	-	1074,615
7.a - Schumacher-Hall	Log Normal	-4,4534 (503,58)	-7,4221 (394,54)	-8,1798 (473,67)	-	3294,920
1.b - Diamétrico	Gaussiano	-2,1508 (0,4083)	1,5883 (0,1217)	-0,2790 (0,0966)	-	1086,367
2.b - Spurr	Gaussiano	2,7655 (0,4362)	0,0007 (0,0006)	-0,1075 (0,0344)	-	1088,311
3.b - Spurr-potência	Gaussiano	-2,3242 (0,4821)	0,5742 (0,0519)	-0,3068 (0,1039)	-	1078,833
4.b - Schumacher-Hall	Gaussiano	-1,8048 (0,3672)	2,1795 (0,2418)	-0,8370 (0,2941)	-0,2331 (0,0917)	1060,792
5.b - Diamétrico	Log Normal	-3,4610 (465,01)	-3,4408 (137,40)	-11,592 (136,58)	-	3239,023
6.b - Spurr	Log Normal	-2,2611 (0,5858)	0,6397 (0,0609)	-0,7931 (0,1683)	-	1065,720
7.b - Schumacher-Hall	Log Normal	-1,5524 (349,03)	0,1028 (242,61)	-1,5215 (294,37)	-3,6284 (98,835)	3124,281