

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

**Avaliação da construção e aplicação de modelos florestais de efeitos fixos e
efeitos mistos sob o ponto de vista preditivo**

Edgar de Souza Vismara

Tese apresentada para obtenção do título de Doutor em
Ciências, Programa: Recursos Florestais. Opção em:
Silvicultura e Manejo Florestal

**Piracicaba
2013**

Edgar de Souza Vismara
Engenheiro Florestal

**Avaliação da construção e aplicação de modelos florestais de efeitos fixos e
efeitos mistos sob o ponto de vista preditivo**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **JOÃO LUIS FERREIRA BATISTA**

Tese apresentada para obtenção do título de Doutor em
Ciências, Programa: Recursos Florestais. Opção em:
Silvicultura e Manejo Florestal

**Piracicaba
2013**

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

Vismara, Edgar de Souza

Avaliação da construção e aplicação de modelos florestais de efeitos fixos e efeitos mistos sob o ponto de vista preditivo / Edgar de Souza Vismara.- - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2013.
106 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2013.

1. Biomassa 2. Eucalipto 3. Máxima verossimilhança 4. Modelos matemáticos
5. Predição 6. Regressão linear I. Título

CDD 634.9734
V832a

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

AGRADECIMENTOS

Em primeiro lugar agradeço ao meu orientador professor doutor João Luis Ferreira Batista, pelo apoio, ensinamentos e confiança. Agradeço a todos os outros professores do Departamento de Ciências Florestais que auxiliaram nesse meu aprendizado, assim como os funcionários do departamento.

A minha esposa pelo carinho, compreensão e paciência.

A todos os meus colegas de pós-graduação pela convivência. Aos meus amigos de graduação pelo apoio, carinho e amizade. E a todos os outros meus amigos.

A minha mãe, pela força, carinho, amizade e amor. A toda minha família, pelo apoio durante minha vida. A minha querida irmã pelo companheirismo.

Aos colegas e amigos de laboratório, pelas risadas, cafés e amizade.

A todos aqueles que conheci durante esses dois anos de mestrado que por mais rápido que tenha sido o contato também contribuiu de certa forma com minha formação.

Não citarei nomes para ter certeza par não cometer nenhuma injustiça.

“Se a base da população for bastante grande, os acontecimentos mais extraordinários e mais acidentais se tornam objeto de cálculo fixo.

Não se poderia aventurar a predizer quando poderão nascer em Boston um capitão Bonaparte, uma cantora como Jenny Lind ou um navegador como Bowditch; mas com uma população de vinte ou de duzentos milhões de almas, pode-se chegar a predições aproximadas. Cada fenômeno relativo à espécie humana, considerada como um todo, pertence à ordem dos fatos físicos. Quanto maior for o número de indivíduos, mais desaparece o indivíduo, deixando a predominância a uma série de fatos gerais que dependem das causas pelas quais a sociedade existe e se mantém.”

Ralph Waldo Emerson

SUMÁRIO

RESUMO.....	10
ABSTRACT.....	12
1 INTRODUÇÃO	14
Referências.....	18
2 CONSTRUÇÃO E AVALIAÇÃO DE MODELOS EMPÍRICOS DE BIOMASSA LENHOSA, A PARTIR DA INCLUSÃO DE INFORMAÇÕES DENSIDADE BÁSICA....	22
.....	
Resumo.....	21
Abstract.....	21
2.1 Introdução.....	22
2.2 Material e Métodos	25
2.2.1 Área de estudo.....	25
2.2.2 Escolha das espécies	25
2.2.3 Determinação da biomassa lenhosa	26
2.2.4 Determinação das variáveis preditoras.	28
2.2.5 Construção dos modelos de predição de biomassa	28
2.2.6 Avaliação dos modelos testados	30
2.3 Resultados.....	33
2.3.1 Avaliação dos modelos testados quanto ao ajuste	33
2.3.2 Avaliação do desempenho preditivo dos modelos testados.....	35
2.4 Discussão.....	50
2.5 Conclusões.....	52
Referências.....	53
3 APLICAÇÃO DE MODELOS LINEARES DE EFEITO MISTO E CALIBRAÇÃO NA PREDIÇÃO DE VOLUME EM PLANTIOS REGIONAIS DE EUCALYPTUS GRANDIS.	56
.....	
Resumo.....	55
Abstract.....	55
3.1 Introdução.....	56
3.2 Material.....	58
3.3 Métodos.....	61

3.3.1 Os modelos de volume	61
3.3.2 O modelo multivariado.....	62
3.3.3 Calibração dos modelos de efeitos mistos para as predições condicionais	64
3.4 Avaliação com dados empíricos.....	66
3.4.1 Estimação.....	66
3.4.2 Calibração.....	67
3.4.3 Predição.....	67
3.4.4 Implementação das abordagens de calibração.....	67
3.4.5 Critérios de avaliação	68
3.5 Resultados.....	68
3.6 Discussão e Conclusões.....	76
Referências	78
4 APLICAÇÃO DE MODELOS LINEARES DE EFEITO MISTO E CALIBRAÇÃO NA PREDIÇÃO DA BIOMASSA EM FLORESTAS TROPICAIS MULTIESPECÍFICAS	82
Resumo.....	81
Abstract.....	81
4.1 Introdução.....	82
4.2 Material.....	85
4.2.1 Área de estudo e coleta de dados.....	85
4.2.2 Medição das variáveis dendrométricas.....	87
4.2.3 Medição da biomassa.....	87
4.3 Métodos.....	90
4.3.1 O modelo de biomassa.....	90
4.3.2 Calibração com dados empíricos.....	92
4.3.3 Aplicação a dados empíricos	94
4.4 Resultados.....	96
4.5 Discussão e Conclusões.....	99
Referências	100
ANEXO	104

RESUMO

Avaliação da construção e aplicação de modelos florestais de efeitos fixos e efeitos mistos sob o ponto de vista preditivo

Neste trabalho procurou-se avaliar o processo de construção e aplicação de modelos preditivos no meio florestal. Para tanto, no primeiro artigo parte-se de uma amostra destrutiva de 200 indivíduos de dez espécies arbóreas distintas, originárias do bioma Atlântico, testando-se três modelos teóricos comumente usados na predição de volume e biomassa, sendo a esses adicionados preditores informativos da densidade básica da árvore. Para a avaliação os modelos ajustados foram simuladas três situações preditivas distintas. Os resultados demonstraram que aplicar o modelo em situações distintas a da amostra de ajuste gera viés nas predições que, no entanto, é reduzido com a entrada dos referidos preditores. O segundo artigo apresenta aplicações da calibração do modelo linear de efeito misto na predição do volume em plantios de *Eucalyptus grandis* em primeira e segunda rotação. Para tanto, partiu-se do modelo de Schumacher e Hall, em sua forma linearizada, para o desenvolvimento modelo de efeitos mistos, que considerou alguns de seus parâmetros como sendo aleatórios ao longo das diferentes fazendas. A calibração foi realizada em nível de fazenda partindo-se de um pequeno número de árvores-amostra. A abordagem foi desenvolvida para modelos univariados de primeira rotação, além de modelos bivariados de duas rotações. Os resultados mostraram que o procedimento de calibração fornece predições mais confiáveis que a dos modelos tradicionais de efeitos fixos em ambas as rotações. O terceiro artigo apresenta aplicações da calibração do modelo linear de efeito misto na predição da biomassa de árvores de espécies nativas numa floresta Ombrófila densa. Partiu-se do modelo de potência, em sua forma linearizada, para o desenvolvimento modelo de efeitos mistos e dois níveis: parcela e espécie. O ajuste do modelo foi feito considerando esses dois níveis, mas a calibração foi realizada em cada nível ignorando o efeito do outro, nível. Os resultados mostraram que o procedimento de calibração fornece predições mais confiáveis em nível de espécie que os modelos tradicionais. Em nível de parcela, a calibração não foi efetiva.

Palavras-chave: Modelos de predição; Estimador de máxima verossimilhança; Modelos de efeitos mistos; BLUP.

ABSTRACT

Evaluation of goodness of fit and application of fixed and mixed effects models in forestry from the predictive point of view

In this study we tried to evaluate the process of construction and application of predictive models in forestry. Therefore, in the first paper we started from a destructive sample of 200 individuals from ten different tree species, originating from the Atlantic biome. We tested three theoretical models commonly used to predict volume and biomass, which was added predictors related to tree basic density. To evaluate the models were simulated three different predictive situations. The results showed that applying the model in different situations from the sample generates bias on predictions; however, it is reduced by adding the referred predictors. The second article presents applications of linear mixed-effects models and calibration to predict the volume in *Eucalyptus grandis* plantations in first and second rotation. Therefore, we started with the model of Schumacher and Hall, in their linearized form to develop the mixed-effects model, which considered some of its parameters as random throughout the different farms. The calibration was performed at the farm level and starting from a small number of sample trees. The approach was developed to first rotation univariate models, and a bivariate model of both rotations. The results showed that the calibration procedure provides more reliable predictions than the traditional fixed effects models in both rotations. The third article presents applications of linear mixed-effects model and calibration to predict the biomass in a rain forest. We started from the power model, in its linearized form, for developing the mixed-effects model considering two levels of grouping: plot and species, Model fitting was made considering these two levels, but the calibration was performed on each level ignoring the other level effect. The results showed that the calibration procedure provides more reliable predictions at species level than traditional models. On the plot level, the calibration was not effective.

Keywords: Predictive models; Maximum likelihood estimator; Mixed effects models; BLUP

1 INTRODUÇÃO

Estimativas de biomassa ou volume acima do solo em processos de inventário florestal são obtidas, geralmente, através da aplicação de modelos de predição construídos a partir de dados empíricos. A construção desses modelos envolve a medição destrutiva de uma série de indivíduos arbóreos representativos de uma determinada população, ajustados a um modelo ou relação funcional, em que a variável resposta (biomassa ou volume) medida é relacionada às variáveis preditoras, tais como, o diâmetro a 1,30 m do solo (DAP) e a altura total (h).

Estas relações funcionais, na sua maioria não lineares e com erro aditivo, sofrem, em geral, um processo de linearização a partir de transformação logarítmica. Essa transformação visa contornar a heterocedasticidade natural da variância e permitir que o ajuste seja feito através de procedimentos de mínimos quadrados ordinários do modelo linear clássico. Essa perspectiva é preferida pela maioria dos autores, pois, segundo Cunia (1996), garante segurança acerca do uso de critérios clássicos de seleção de modelos, tais como, os baseados nos testes de hipóteses, no coeficiente de determinação e no exame gráfico dos resíduos. No entanto, essa abordagem apesar de trazer certa segurança em termos do uso de critérios de seleção quanto o ajuste, limita o pesquisador a uma classe muito específica de modelos.

Um modelo de predição deve ser avaliado, em termos da sua função de distribuição, bem como, em termos das relações funcionais que descrevem o comportamento dos parâmetros da função de distribuição. O modelo deve ser avaliado não somente sob o ponto de vista da qualidade do seu ajuste, mas também, sob o ponto de vista do seu desempenho de aplicação (BATAGLIA; SANDS, 1998).

A abordagem clássica de seleção de modelos aplicada geralmente nos processos de construção de modelos preditivos, foca principalmente nos aspectos relacionados à qualidade do ajuste da função da média. Essa qualidade de ajuste é verificada através da amplitude da dispersão dos resíduos gerados da aplicação do modelo nos próprios dados de ajuste. Estes critérios, porém, não permitem considerar diferentes funções de distribuição, nem tampouco, comparar funções para os parâmetros que não sejam aninhadas ou hierárquicas quanto aos preditores. Além disso, aspectos a respeito da qualidade do desempenho dos modelos em realizar predições são poucas vezes considerados. Isso ocorre pela escassez de amostras destrutivas em grande quantidade para realização de procedimentos de validação e pela

natureza, intrinsecamente associada à visão de qualidade de ajuste, do arcabouço de seleção de modelos que compõe a análise do modelo linear clássico.

A abordagem de modelagem baseada na verossimilhança (HILBORN; MANGEL, 1997; ROYALL, 2000; BURNHAM; ANDERSON, 2002), por sua vez, pode melhorar o processo de construção dos modelos preditivos, garantindo ao modelador uma maior flexibilidade no processo. Isso porque, essa abordagem permite o ajuste e comparação de modelos Gaussianos ou Log-Normais com quaisquer relações funcionais, lineares e não lineares, para modelar a média e o desvio padrão em função de quaisquer preditores. Além disso, através do uso do critério de informação de Akaike (AIC) pode-se incluir uma perspectiva preditiva no processo de seleção de modelos, já que, Stone (1974 e 1977) e Akaike (1974), através de uma abordagem teórica e Davies et al. (1995), através de uma abordagem empírica, demonstraram a equivalência assintótica do AIC e da validação cruzada no processo de seleção modelos.

Quanto ao desempenho de aplicação, Vanclay (1994), afirma que as equações usadas para prever biomassa e volume são empíricas e sítio-específicas, por fornecer uma descrição do comportamento da biomassa sem tentar identificar as causas ou explicar os fenômenos. Para que essas equações forneçam estimativas precisas, devem ser formuladas de forma a se comportarem de modo biologicamente realista dentro de uma amplitude de condições limitadas pelo processo de obtenção dos dados e construção do modelo.

Apesar de diversos autores (BROWN et al., 1989; HIGUCHI; CARVALHO JR., 1994; KOEHLER et al., 2002; NOGUEIRA et al., 2007; VIEIRA et al., 2008) discutirem as limitações do emprego de modelos sítios- específicos em situações distintas àquelas em que estes foram criados, o problema de se considerar apenas a qualidade do ajuste dos modelos aos dados é que este procedimento pode gerar uma falsa impressão de generalidade de aplicação do modelo quando este é obtido a partir de uma amostra composta de indivíduos provenientes de uma vasta gama de espécies, tamanhos e procedências geográficas.

Os chamados modelos de predição de biomassa “pan-tropicais”, como o de Chave (2005), por exemplo, partem da premissa de que se amostra contiver um grande número de espécies provenientes de uma série de sítios localizados nos trópicos, este possa ser aplicado em diferentes regiões tropicais com maior segurança e menor viés que modelos construídos para uma região específica e aplicados em outra. O problema, no entanto, pode se mostrar muito mais complicado, já que, a variabilidade de situações de estoque de biomassa presentes no ambiente tropical é muito mais complexa que uma amostra de ajuste, por maior e mais diversificada que seja, possa incluir.

Vismara et al. (2010), demonstrou a natureza empírica e sítio-específica destes modelos de predição de biomassa publicados para a floresta Tropical. Estes utilizaram modelos obtidos para a região Amazônica, para o bioma atlântico, o modelo “pan-tropical” de Chave (2005), além de um modelo próprio para realizar predições de biomassa a partir do diâmetro em dados de biomassa obtidos numa área de reflorestamento localizada em região de predominância da floresta Atlântica localizada no estado do Espírito Santo. Estes constataram uma grande disparidade entre os coeficientes ditos "alométricos" da função obtida para amostra e os modelos da literatura ajustados para região Amazônica e o modelo de Chave (2005). O modelo de pior desempenho foi o modelo “pan-tropical” de Chave (2005), em contrapartida, o modelo obtido para o bioma Atlântico (BURGUER et al., 2005) apresentou um desempenho mais adequado (Figura 1.1).

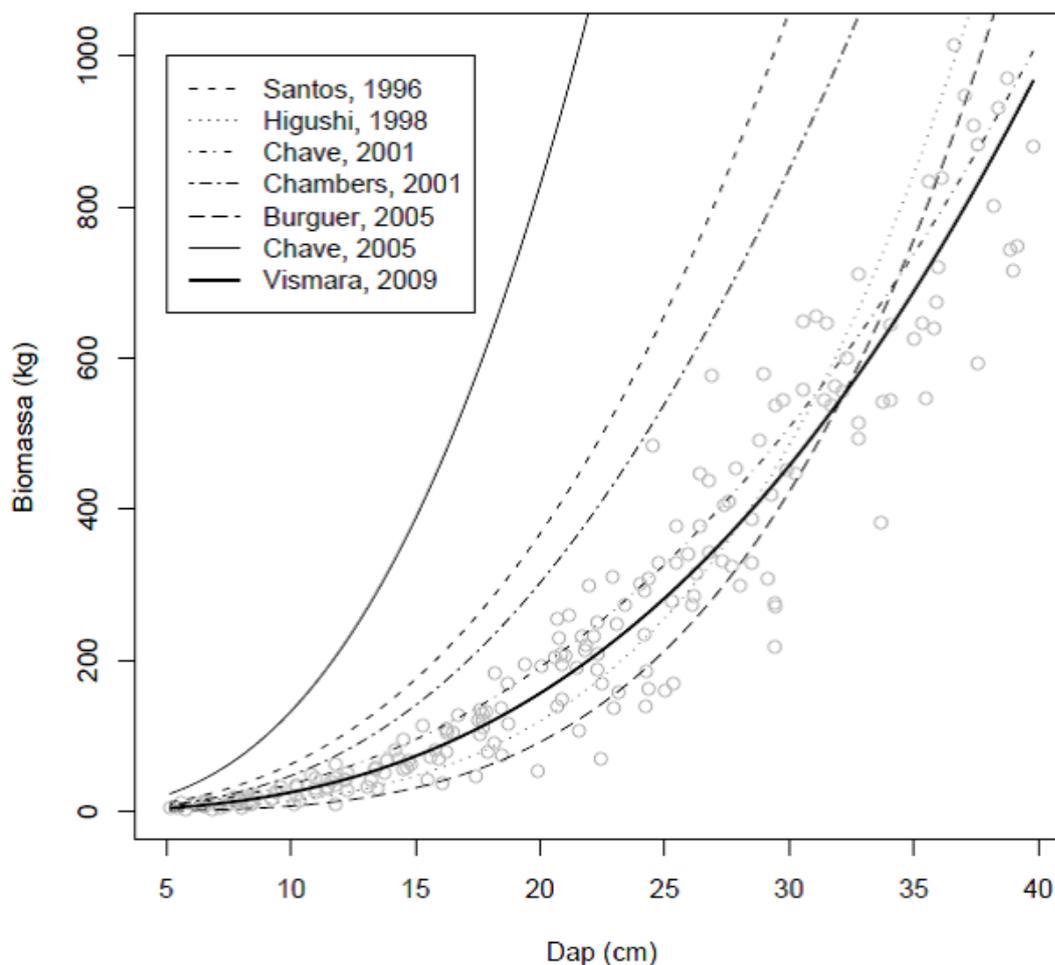


Figura 1.1 - Gráfico das curvas dos diferentes modelos de literatura e do modelo de Vismara (2009), aplicados aos dados de Vismara (2009).

Essa disparidade entre equações obtidas na floresta Amazônica e floresta Atlântica já havia sido observada por Vieira et al. (2008) e Sotomayor et al. (2010), comprovando a natureza eminentemente empírica destes modelos. Um aspecto particularmente importante nos erros de predição gerados pelos modelos empíricos é que estes são sistemáticos e não aleatórios, ou seja, o impacto destes erros observados em nível de árvore tende a ser mais importantes quando aplicados em nível de povoamento.

Outros resultados que demonstram a natureza empírica e a dificuldade de se estabelecer uma relação "alométrica" em que os parâmetros demonstrem um comportamento estável para diferentes espécies em diferentes situações pode ser visto em Zianis e Radoglou (2006). Esse trabalho comparou modelos obtidos de dados empíricos com os coeficientes alométricos de literatura e observaram uma discrepância importante entre os coeficientes alométricos e os coeficientes obtidos após ajuste aos dados, demonstrando a natureza eminentemente empírica relacionada às equações de predição de biomassa.

Dentro deste contexto de natureza empírica dos modelos de predição da biomassa, mesmo um critério preditivo e com propriedades de validação cruzada como o AIC esta sujeito à estrutura limitada pelas características particulares da amostra de ajuste. Sendo assim se faz necessário analisar os modelos obtidos por esse critério do ponto de vista do seu desempenho em situações em que a estrutura da floresta seja distinta à da amostra de ajuste e que permita, portanto, sua avaliação adequada.

Os procedimentos de seleção de modelos deveriam, sempre que possível, incluir a comparação das predições com dados empíricos de biomassa específicos da situação em estudos, ainda que para uma amostra pequena, uma vez que as informações obtidas na literatura são raramente suficientes para que se possa dizer de antemão qual dos modelos publicados resultará em predições adequadas ou inadequadas na situação de estudo.

Para isso, no entanto, seria necessário dispor de dados provenientes de amostras destrutivas em áreas onde os modelos de predição fossem aplicados. Diante da impossibilidade e incoerência desta abordagem e dentre as diversas possibilidades para contornar esse problema, o presente trabalho propõe duas abordagens. A primeira delas seria, através de procedimentos de reamostragem no próprio conjunto de dados de ajuste, simular cenários distintos para que uma avaliação mais abrangente e realista possa ser aplicada no processo de seleção destes modelos preditivos. A segunda saída proposta é a utilização de procedimentos de calibração da relação empírica regional à uma situação local específica, através do uso de modelos de efeitos mistos.

Desta forma, este trabalho tomando a abordagem da verossimilhança como base do processo de modelagem, apresenta no capítulo 2 um estudo sobre o processo de construção e avaliação de modelos de predição de biomassa a partir de dados empíricos sob a ótica do ajuste e do desempenho usando o artifício da simulação por "bootstrap". No capítulo 3, propõe-se o uso de procedimentos de ajuste de modelos de efeitos mistos e calibração aplicados a dados regionais de volume *Eucalyptus grandis*. O capítulo 4 traz novamente a abordagem de modelos de efeitos mistos e calibração mas aplicados à dados de floresta tropical multiespecífica. A abordagem de modelos de efeito misto, apesar de simplista do ponto de vista da linearidade, atenua as características sítio-específicas do modelo, pelo fato de tratar do comportamento aleatório de certos coeficientes associados aos preditores do modelo que podem variar e ser calibrados de acordo com a situação de aplicação.

Referências

- AKAIKE, H. A new look at statistical model identification. **IEEE Transactions on Automatic Control**, Tokyo, v. 19, n. 6, p. 717-723, Dec. 1974.
- BATTAGLIA, M.; SANDS, P.J. Process-based forest productivity models and their application in forest management. **Forest Ecology and Management**, Wisconsin, v. 102, n. 1, p. 13-32, Mar. 1998.
- BROWN, S.; GILLESPIE, A. J. R.; LUGO, A. E. Biomass estimation methods for tropical forest with applications to inventory data. **Forest Science**, Bethesda, v. 35, n. 4, p. 881-902, Dec. 1989.
- BURGER, D.M. **Modelos alométricos para a estimativa da fitomassa de Mata Atlântica na Serra do Mar, SP**. 2005. 112 p. Tese (Doutorado em Ecologia) – Instituto de Biociências, Universidade de São Paulo, São Paulo, 2005.
- BURGER, D.M.; DELITTI, W.B.C. Allometric models for estimating the phytomass of a secondary Atlantic Forest area of southeastern Brazil. **Biota Neotropica**, Campinas, v. 8, n. 4, p. 131-136, Oct./Dec. 2008.
- BURNHAM, K.P.; ANDERSON, D.R. **Model selection and multimodel inference: a practical-theoretic approach**. 2nd ed. New York: Springer-Verlag, 2002. 488 p.
- CHAMBERS, J.Q.; SANTOS, J.; RIBEIRO, R.J.; HIGUCHI, N. Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. **Forest Ecology and Management**, Wisconsin, v. 152, p. 73-84, Oct. 2001.
- CHAVE, J.; RIÉRA, B.; DUBOIS, M. Estimation of biomass in a neotropical forest of French Guiana: spatial and temporal variability. **Journal of Tropical Ecology**, Winchelsea, v. 17, p. 79- 96, Jan. 2001.

CHAVE, J.; ANDALO, C.; BROWN, S.; CAIRNS, M.; CHAMBERS, J.C.; EAMUS, D., FÖLSTER, H.; FROMARD, F.; HIGUCHI, N.; KIRA, T.; LESCURE, J.; NELSON, B.W.; OGAWA, H.; PUIG, H.; RIÉRA, B.; YAMAKURA, T. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. **Oecologia**, Heidelberg, v. 145, n. 1, p. 87-99, June 2005.

CUNIA, T. Error of forest inventory estimates: its main components. In: WORKSHOP ON TREE BIOMASS REGRESSION FUNCTIONS AND THEIR CONTRIBUTION TO THE ERROR OF FOREST INVENTORY ESTIMATES, 1996, Syracuse. **Estimating tree biomass regressions and their error: proceedings...** Syracuse: Society of American Foresters, 1996. p. 1-13.

DAVIES, S.L.; NEATH, A.A.; CAVANAUGH, J.E. Cross validation model selection criteria for linear regression based on the Kullback-Leibler discrepancy. **Statistical Methodology**, Amsterdam, v. 2, n. 4, p. 249-266, Dec. 2005.

HIGUCHI, N.; CARVALHO Jr.; J.A. Fitomassa e conteúdo e carbono de espécies arbóreas da Amazônia. In: SEMINÁRIO EMISSÃO X SEQUESTRO DE CO₂: 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.

HILBORN, R.; MANGEL, M. **The ecological detective: confronting models with data.** Princeton: Princeton University Press, 1997. 330 p.

KOEHLER, H.S.; WATZLAWICK, L.F.; KIRCHNER, F.F. Fontes e níveis de erros nas estimativas do potencial de fixação de carbono. In: SANQUETA, C.R.; WATZLAWICK, L. F. ; BALBINOT, R.; ZILLIOTO, M. A. B.; GOMES, S. F. (Ed.). **As florestas e o carbono.** Curitiba: FUPEF/ Imprensa da UFPR. 2002. p. 251-264.

LAPPI, J. Calibration of height and volume equations with random parameters. **Forest Science**, Bethesda, v. 37, n. 3, p. 781-801, 1991.

NOGUEIRA, E.M.; FEARNside, P.M.; NELSON, B.; FRANÇA, M.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**, Winsconsin, n. 248, p. 119-135, Sept. 2007

SANTOS, J. **Análise de modelos de regressão para estimar a fitomassa da floresta tropical úmida de terra-firme da Amazônia brasileira.** 1996. 121 p. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa 1996.

SOTOMAYO, J.F.M.; VISMARA, E.S.; COUTO, H.T.Z.; BATISTA, J.L.F. Prediction models and estimation bias of forest aboveground biomass: the case of Atlantic Forest. In: INTERNATIONAL BIOMETRIC CONFERENCE, 25., 2010, Florianópolis. **Anais...** Florianópolis: IBS, 2010. p. 117-118.

STONE, M. Cross-validatory choice and assessment of statistical predictions. **Journal of the Royal Statistical Society**. Series B (Methodological), London, v. 36, n. 2, p. 111-147, 1974.

_____. An asymptotic equivalence of choice of model by cross-validation and Akaike's Criterion. **Journal of the Royal Statistical Society**. Series B (Methodological), London, v. 39, n. 1, p. 44-47, 1977.

ROYALL, R.M. **Statistical evidence: a likelihood paradigm**. London: Chapman and Hall. 2000. 191 p.

VANCLAY, J.K. **Modelling forest growth and yield: applications to mixed tropical Forests**. Wallingford: CAB International, 1994. 212 p.

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.; HIGUCHI, 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**, Campinas, v. 8, n. 2, p. 21-29, jun. 2008.

VISMARA, E.S. **Mensuração da biomassa e seleção de modelos para construção de equações de biomassa**. 2009. 102 p. Dissertação (Mestrado em Recursos Florestais) – Departamento de Ciências Florestais, ESALQ, Universidade de São Paulo, Piracicaba, 2009.

VISMARA, E.S.; BATISTA, J.L.F.; NICOLETTI, M.F.; CASTRO, T. N.; COUTO, H.T.Z. Análise da aplicação de equações de biomassa e carbono para espécies nativas da mata atlântica na região de Linhares, ES. In: JORNADAS TÉCNICAS FORESTALES Y AMBIENTALES, 14., 2010, Misiones. **Anais...** Misiones: Facultad de Ciencias Forestales, 2010. p. 92-112.

ZIANIS, D.; RADOGLU, K. Comparison between empirical and theoretical biomass allometric models and statistical implications in stem volume predictions. **Forestry**, Oxford v. 79, p. 477-487, 2006.

2 CONSTRUÇÃO E AVALIAÇÃO DE MODELOS EMPÍRICOS DE BIOMASSA LENHOSA, A PARTIR DA INCLUSÃO DE INFORMAÇÕES DENSIDADE BÁSICA

Resumo

Neste trabalho avaliou-se o desempenho, quanto ao seu ajuste e à aplicação, diversas equações de predição da biomassa lenhosa que incluíam algum nível de informação a cerca da densidade básica da árvore. Para tanto, partiu-se de uma amostra destrutiva de 200 indivíduos de dez espécies arbóreas distintas, originárias do bioma Atlântico. Foram testados três modelos teóricos comumente usados na literatura para predição de volume e biomassa e a esses foram acrescentados os preditores: medida de resistência a penetração no lenho, densidade básica de literatura, a densidade básica da espécie e da árvore individual. Para a avaliação os modelos ajustados foram simuladas três situações preditivas distintas. Os resultados demonstram que aplicar o modelo em situações distinta a da amostra de ajuste gera predições enviesadas. Esse viés é, no entanto, reduzido coma a entrada dos preditores.

Palavras-chave: Estimadores de máxima verossimilhança; Modelos de predição da biomassa; Seleção de modelos

Abstract

In this study we evaluated the performance, regarding its goodness of fit and its application, different biomass prediction models that included some level of information about the basic density of the tree. To this end, we started from a destructive sample of 200 trees from 10 different Atlantic forest species. Three different theoretical models commonly used in the literature to predict volume and biomass were tested and the following predictors were added: a measure of resistance to penetration in the wood, the basic density of literature, the basic density of tree species and individual basic density. To evaluate the fitted models we used three different simulated predictive situations. The results show that the model applied in situations distinct from the fitting sample had biased predictions. This bias is, however, reduced with the input of referred predictors.

Keywords: Maximum likelihood estimation; Biomass prediction models; Model selection

2.1 Introdução

Os inventários florestais têm sido usados como ponto de partida para estimar a biomassa aérea nas florestas tropicais no Brasil (BROWN et al., 1989). Os dados dendrométricas desses inventários, associados às estimativas de densidade básica média das

espécies, são usados então para converter o volume obtido por um fator de forma em biomassa aérea (SEGURA, 2005).

O uso dessa metodologia, apesar da simplicidade, pode, segundo Fearnside (1997) e Nogueira et al. (2007), levar a estimativas muito discrepantes em relação a biomassa verdadeira. Essa discrepância, segundo os autores, se deve aos erros associados à obtenção das variáveis dendrométricas e ao erro associado à aplicação da densidade básica em nível de espécie. O 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 nos diferentes locais.

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 à altura do peito (DAP) e altura total ou comercial, são medidas e associadas à biomassa aérea, determinada por métodos diretos (gravimetria) e indiretos (volumetria e determinação de densidade básica), através de técnicas de regressão (ARAUJO et al., 1996; HIGUCHI et al., 1998; NELSON et al., 1998; SANTOS et al., 2001; CHAMBERS, 2001; CHAVE, 2001, 2005; BURGUER, 2005, 2008, 2010).

Sob o ponto de vista da aplicação dessas equações acredita-se que, se construídas para uma determinada região, elas possam ser aplicadas aos dados de inventários florestais com desempenho superior aos da técnica de aplicação de estimativas de densidade básica na conversão de volume em biomassa. No entanto, nas florestas tropicais o estoque de biomassa varia espacialmente em função da composição florística, da variação intra e interespecífica de densidades e de formas (SAATCHI et al., 2007). Essa variação é causada, na maioria dos casos, por gradientes de fertilidade de solo, declividade do terreno, altitude e perturbação antrópica (SAMPAIO et al., 2010) e segundo, Alves et al. (2010), podem ser observadas em distâncias bastante curtas dentro da floresta.

Ao se considerar a natureza eminentemente empírica dessas equações, demonstrada no trabalho de Zianis e Radoglou (2006), onde não se conseguiu encontrar, partir de dados empíricos uma relação alométrica fixa entre DAP, altura e biomassa aérea de diferentes espécies arbóreas na Europa, é bem provável que esses modelos tenham um comportamento bastante enviesado em termos de predição em larga escala espacial.

Isso ocorre, pois, além da relação hipsométrica variar muito nas espécies tropicais, sem a informação de densidade básica ela se torna muito frágil em predizer a relação entre volume e biomassa observada entre as espécies que constituem esse bioma (VISMARA, 2009). A predição equivocada pode gerar erros sistemáticos positivos, quando a biomassa

predita é inferior à biomassa real ou de negativos, quando a biomassa predita é inferior a biomassa real. Esse erro sistemático ou viés, no entanto, é de difícil medição e comprovação, já que, dados de parcelas de inventário florestal em que a biomassa foi obtida diretamente são muito raros, dificultando a correta avaliação desses modelos.

Modelos de predição podem ser avaliados sob o ponto de vista estatístico, no que se refere qualidade do ajuste, ou sob o ponto de vista do seu desempenho de aplicação no campo (PRISLEY; MORTIMER, 2004). Equações empíricas de volume e biomassa devem ser avaliadas, principalmente, sob o ponto de vista do seu desempenho de predição, razão pela qual essas foram construídas (BATTAGLIA; SANDS, 1998). Diante da escassez de dados que permitam a avaliação adequada desses modelos pode-se lançar mão de procedimentos de simulação a fim de verificar a magnitude do erro de predição em escala espacial (NOGUEIRA, 2008b). Esses processos de simulação, apesar de distantes da realidade complexa da estrutura da floresta, podem fornecer informações importantes que contribuam para o processo de avaliação dos modelos para além da qualidade do ajuste.

Neste trabalho, trabalhou-se com a ideia de que a entrada de alguma informação de densidade básica das árvores como variável preditora do modelo possa permitir, de certa forma, uma melhora no ajuste e no desempenho do modelo em realizar predições de biomassa aérea de árvores em ambientes tropicais. O objetivo aqui é avaliar o ajuste e o desempenho de equações de predição da biomassa lenhosa com a inclusão de algum nível de informação a cerca da densidade básica da árvore, a partir de uma amostra destrutiva de 200 indivíduos de dez espécies arbóreas distintas, originárias do bioma Atlântico.

Para isso, foram testados três modelos teóricos comumente usados na literatura para predição de volume e biomassa e a esses foram acrescentadas como variáveis predictoras: uma medida de resistência à penetração no lenho, a densidade básica de literatura, a densidade básica da espécie e da árvore individual determinados em laboratório. Para realizar essa avaliação os modelos ajustados foram utilizados na predição da biomassa em três situações distintas, que foram criadas a fim de simular uma situação real onde ocorram variações no estoque de biomassa aérea entre unidades amostrais.

2.2 Material e Métodos

2.2.1 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. As árvores foram coletadas numa área de reflorestamento implantada na década de setenta e manejada intensivamente com o intuito de se verificar o potencial das espécies em plantios comerciais.

2.2.2 Escolha das espécies

A escolha das espécies para realização deste estudo seguiu dois critérios que juntos buscavam certa generalidade quanto à aplicação do modelo obtido 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 em função das frequências por classe de altura é apresentada na Tabela 2.1.

Tabela 2.1 – Distribuição de frequências das classes de altura pelas classes de diâmetro das árvores amostra estudadas.

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

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. A distribuição dos diâmetros em relação as alturas totais na amostra coletada demonstra a estrutura característica de povoamentos multiespecíficos, onde, a partir das classes intermediarias de altura, indivíduos de classes de diâmetros muito distintas, ocorrem numa mesma classe de altura (Tabela 2.1).

2.2.3 Determinação da biomassa lenhosa

A biomassa lenhosa foi determinada para os diferentes compartimentos das árvores e esses compartimentos foram definidos de acordo com a literatura (ARAUJO et al., 1996; HIGUCHI et al., 1998; SANQUETTA, 2002; SOARES et al., 2002; SEGURA, 2005).

Copa

Cada árvore foi desfolhada e, em seguida, desganhada e seus ramos e galhos, inclusive a ponta da árvore com diâmetro menor que 5 cm, foram pesados no campo com

célula de carga com capacidade máxima de 200 kg e precisão na casa de centenas de gramas. No conjunto de ramos e galhos foram retirados fragmentos localizados em diferentes posições da copa da árvore, os quais foram misturados, formando-se uma amostra composta.

Desta amostra, foi retirada uma sub amostra de aproximadamente 500g de galhos e 500g de ramos, para determinação da sua massa seca, após secagem em estufa de circulação forçada de ar a 103°C. A partir da razão entre massa verde (Kg) e massa seca (Kg) das amostras multiplicada por 100, foi obtido o teor de umidade dos ramos e galhos das árvores em porcentagem (%). A biomassa de ramos e galhos foi obtida de acordo com a seguinte expressão:

$$bc_i = m_i(1 - u_i/100), \quad (0.1)$$

onde i é o índice que se refere às árvores; bc_i é a biomassa de ramos ou galhos (kg); m_i é a massa verde de ramos ou galhos (kg); e u_i é o teor de umidade de ramos ou galhos (%).

Tronco

O tronco da árvore foi seccionado e pesado com auxílio da célula de carga. 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 no campo para determinação da sua massa verde. Após secagem em estufa de circulação forçada de ar a 103°C foi determinada sua massa seca. A partir da razão entre massa verde (Kg) e massa seca (Kg) dos discos multiplicada por 100, foi obtido o teor de umidade do tronco das árvores em porcentagem (%). A biomassa do tronco das árvores, com casca foi obtida de acordo com a seguinte expressão:

$$bt_i = m_i(1 - u_i/100), \quad (0.2)$$

onde, i é o índice que se refere às árvores; bt_i 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 (%).

Além da determinação do teor de umidade os discos foram utilizados na determinação 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) e a densidade básica das árvores foram obtidas pela média aritmética das densidades básicas dos cinco discos.

2.2.4 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) com clinômetro eletrônico Vertex III da empresa Haglöf.

Além das medidas tradicionalmente tomadas para construção de modelos de predição de volume e biomassa, como DAP e altura total foram coletadas informações acerca da densidade básica da madeira do lenho em quatro níveis distintos. Esses níveis foram definidos pela qualidade da informação que essas variáveis fornecem em identificar a densidade básica verdadeira da árvore e, em ordem crescente de informação, são: uma medida de resistência à penetração no lenho da árvore i da espécie j (p_{ij}) tomada na altura do DAP em mm, a densidade básica da espécie j informada pela literatura (db_{ij}) em g/cm^3 , a densidade básica de uma determinada espécie j em g/cm^3 , obtida pela média aritmética das densidades básicas dos i indivíduos dessa mesma espécie em g/cm^3 e determinada em laboratório (db_j), e por fim, a densidade básica de cada árvore i em g/cm^3 para todas as j espécies, determinada em laboratório (db_{ij}).

2.2.5 Construção dos modelos de predição de biomassa

Segundo Parresol (1999) os modelos de predição de biomassa apresentam, geralmente, erros heteroscedásticos, ou seja, a variância dos erros 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 o problema da heteroscedasticidade seja corrigido, a transformação resulta em duas sérias limitações. Primeiramente, certas propriedades estatísticas das estimativas geradas na escala transformada, como ausência de viés e variância constante, não são observadas quando se

realiza a correção para escala original a partir da aplicação do anti-logaritmo Parresol (1999). Em segundo lugar, vários dos critérios de seleção de modelos não são comparáveis quando estão em escalas distintas, como por exemplo, o erro padrão da estimativa e o Critério de informação de Akaike (AIC).

A abordagem de modelagem nesse trabalho consistiu em adaptar os modelos ao problema da heteroscedasticidade. Desta forma, neste trabalho foi considerada uma classe de modelos, designada de modelos Gaussianos heteroscedásticos, onde a biomassa é modelada como uma distribuição Gaussiana com a seguinte função de densidade probabilística:

$$f(y) = \frac{1}{\sigma(x_1, \dots, x_k)\sqrt{2\pi}} \exp \left\{ -\frac{[y - \mu(x_1, \dots, x_k)]^2}{2(\sigma(x_1, \dots, x_k))^2} \right\}, -\infty < y < \infty, \quad (0.3)$$

onde y é a biomassa do lenho, $\mu(x_1, \dots, x_k)$ 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 ($\sigma(\cdot)$).

Para este trabalho foram considerados três modelos teóricos, tradicionalmente utilizados no meio florestal para predição de volume de árvores individuais, como expressões da média do modelo da classe gaussiana.

A esses três modelos teóricos foram acrescentadas, na forma de variáveis preditoras, informações acerca da densidade básica da madeira do lenho (I_{ij}), descritas no item 2.2.4. Os diferentes modelos da classe gaussiana heteroscedastica, testados neste estudo e que se distinguem pela expressão que modela a média, são apresentados na Tabela 2.2.

Tabela 2.2 - Expressões para modelagem da média (b_{ij}) nos modelos Gaussianos heteroscedásticos. Onde b_{ij} -- Biomassa; d_{ij} -- DAP; h_{ij} -- altura total; I_{ij} -- Informação de densidade básica que pode ser: p_{ij} -- Medida de penetração; dbl_j -- Densidade básica da espécie de literatura; db_j -- Densidade básica da espécie determinada em laboratório e db_{ij} -- Densidade básica da árvore individual determinada em laboratório. O índice i se refere às árvores e o índice j se refere à espécie.

Denominação do Modelo	Forma da Expressão da Média	
	Tradicional	Com Informação de Densidade Básica (I)
Husch	$\beta_0 d_{ij}^{\beta_1}$	$\beta_0 d_{ij}^{\beta_1} I_{ij}^{\beta_2}$
Spurr	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} I_{ij}^{\beta_2}$
Schumacher-Hall	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} I_{ij}^{\beta_3}$

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

$$\sigma(d) = \alpha_0 d \cdot \alpha_1 \quad (0.4)$$

As diferenças na 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. Os parâmetros das funções de μ e σ foram obtidos inicialmente pelo método de mínimos quadrados. Isso foi realizado para se obter valores iniciais e para garantir a convergência do processo iterativo para um mínimo global da função de verossimilhança. O procedimento de ajuste foi implementado no pacote “stats4” do software estatístico “R”, utilizando a função de ajuste “mle” (maximum likelihood estimation), que utiliza o algoritmo “optim” para otimização numérica.

2.2.6 Avaliação dos modelos testados

A avaliação dos modelos testados para predição da biomassa do lenho foi realizada em duas etapas. Primeiramente se realizou uma avaliação estatística dos modelos quanto à qualidade de seu ajuste aos dados, ou seja, a sua adequação em explicar a variação da

variável reposta (biomassa lenhosa) tomada na amostra. Num segundo momento procedeu-se a avaliação do desempenho de predição do modelo numa situação em que parcelas foram geradas a partir de procedimentos de simulação. Para todos os modelos ajustados a normalidade dos resíduos foi verificada através de uma inspeção visual do gráfico quantil-quantil. Além disso, foram graficados os perfis de log-verossimilhança negativa dos parâmetros ajustados de todos os modelos, a fim de verificar a qualidade do processo de otimização, bem como a incerteza acerca das estimativas dos parâmetros.

Avaliação quanto a qualidade do ajuste

Segundo Parresol (1999) existem vários critérios de comparação de modelos (Índices de ajuste) que são comumente utilizados na comparação da qualidade do ajuste de modelos de predição de biomassa. Entre eles o autor cita o coeficiente de determinação (R^2), o Erro padrão residual (EPR), o Coeficiente de variação (CV), o índice de Furnival e o critério de informação de Akaike (AIC). Na abordagem da verossimilhança o critério natural de comparação de modelos é o (AIC), isso porque ele fornece uma medida relativa de plausibilidade entre os modelos concorrentes.

Este é descrito por (AKAIKE, 1974, 1985) como:

$$AIC = -2\log(L(\hat{\sigma}^2)) + 2p, \quad (0.5)$$

onde n é o tamanho da amostra, p é o número de parâmetros do modelo e $\hat{\sigma}^2$ é a estimativa de verossimilhança da variância. Segundo este critério quanto menor o valor de AIC obtido pelo modelo, mais plausível é o modelo e melhor sua qualidade para realizar predições de biomassa, considerando os dados coletados. Além disso, na análise comparativa, segundo Burham e Anderson (2002), segue-se a regra canônica de que modelos concorrentes com diferenças de AIC menores que dois são igualmente plausíveis.

Avaliação quanto ao desempenho preditivo

Nas florestas tropicais ocorrem variações na composição florística e na composição diamétrica em escalas espaciais reduzidas (Vieira et al., 2010). Sendo assim, um bom modelo, que possa ser aplicado em áreas relativamente grandes e nestes tipos de floresta,

deve apresentar um desempenho robusto tendo em vista essa variação. A fim de se avaliar o desempenho dos modelos considerados neste trabalho, foram montados três cenários hipotéticos onde todos os modelos foram usados para realizar predições de biomassa lenhosa. Estes cenários foram elaborados com intuito de simular situações distintas àquelas de construção do modelo.

A partir dos mesmos dados de ajuste foram criadas, através de processo de reamostragem "bootstrap" aplicado aos dados de ajuste, 5000 parcelas de 100 árvores para cada um dos três cenários considerados. No cenário 1, as parcelas foram criadas reamostrando-se as árvores com probabilidade diretamente proporcional a densidade básica das mesmas. No cenário 2, todas as árvores foram reamostradas com igual probabilidade e no cenário 3 as parcelas foram geradas reamostrando-se as árvores com probabilidade inversamente proporcional a sua densidade básica.

Com isso, pretendeu-se que o cenário 2 apresentasse uma composição florística e diamétrica e uma densidade básica média muito semelhantes as da amostra de ajuste. No cenário 1 o objetivo foi de criar parcelas com uma densidade básica média maior, acarretando numa composição florística de espécies de maior densidade. No cenário 3, por sua vez, o objetivo foi de compor parcelas com espécies de menor densidade básica, onde a densidade básica média fosse inferior a da amostra de ajuste.

Nos três cenários os modelos foram usados para realizar as predições da biomassa lenhosa. Esse processo gerou um erro de predição por árvore, calculado através da diferença entre a biomassa observada na parcela e a biomassa predita pelo modelo. A avaliação dos modelos considerados foi feita então em função da precisão destes em predizer a biomassa do lenho em cada cenário hipotético. A avaliação da dos modelos foi realizada através do cálculo das estimativas do viés, em escala absoluta (Kg) e em escala relativa (%), com seu respectivo desvio padrão, em cada cenário. Para melhor visualização do comportamento preditivo dos modelos foram gerados gráficos de distribuição dos erros médios de predição por parcela, em escala relativa, para cada um dos cenários simulados.

2.3 Resultados

2.3.1 Avaliação dos modelos testados quanto ao ajuste

O anexo I apresenta as estimativas dos coeficientes dos modelos ajustados neste trabalho com seus respectivos erros padrão. Todos os modelos ajustados apresentaram comportamento simétrico de distribuição dos resíduos e perfis de log-verossimilhança negativa que demonstraram comportamento adequado do processo de otimização que levou a obtenção dos parâmetros. Diante disso, os modelos se mostraram aptos a realizar previsões nos cenários que foram simulados para avaliação de seu desempenho.

Antes de iniciar a comparação entre os modelos, cabe ressaltar que, quando se compara o modelo teórico de Husch com os modelos teóricos de Spurr e Schumacher-Hall esta se avaliando as expressões da média (relação funcional) destes modelos, além da entrada de uma informação fundamental que é a altura total.

Analisando os valores de (AIC) do modelo teórico de Husch na Tabela 2.3 pode-se afirmar que este teve, de maneira geral, o pior ajuste dentre os modelos teóricos testados e a entrada das variáveis preditoras informativas da densidade básica foi bastante efetiva em melhorar seu ajuste aos dados coletados. Uma observação interessante acerca deste modelo é de que as expressões da média que incluíram as variáveis preditoras densidade básica de literatura (dbl_j) e densidade básica média não demonstraram diferenças relevantes entre si na explicação da biomassa lenhosa, já que a diferença nos valores de AIC foi inferior a duas unidades.

O modelo teórico de Schumacher-Hall, em contrapartida, foi o que apresentou dentre os concorrentes, os menores valores de AIC e que, portanto, melhor se ajustou aos dados. A única exceção é a expressão da média com inclusão da medida de penetração (p_{ij}) que não diferiu em plausibilidade em relação a expressão da média do modelo de Spurr, que inclui a mesma variável preditora. Da mesma forma, a inclusão das variáveis preditoras informativas da densidade básica foram muito eficientes em conferir plausibilidade ao modelo.

O modelo de Spurr, por sua vez, apresentou um ajuste aquém do modelo de Schumacher-Hall, mas muito superior ao modelo de Husch. Assim como este último as expressões que incluíram os preditores densidade básica de literatura (dbl_j) e densidade básica média (db_j) não diferiram em plausibilidade. Realizando uma análise geral, foi notável a importância da variável preditora altura total (h_{ij}) em conferir plausibilidade e

ajuste aos modelos de Schumacher-Hall e Spurr. Aliás, comparando esses dois modelos isoladamente em termos de (AIC) pode-se afirmar ainda que a entrada dos preditores informativos da densidade básica foi muito mais efetiva em conferir plausibilidade ao modelo de biomassa que o próprio modelo teórico em si. Isto porque ao se observar a Tabela 2.3 nota-se que as diferenças de AIC entre as expressões da média dentro do modelo teórico de Schumacher-Hall e Spurr são maiores que as diferenças de AIC entre os modelos teóricos num mesmo nível de informação de densidade básica.

Considerando apenas a entrada das informações de densidade básica, nos seus diferentes níveis, como variáveis preditoras observa-se que a densidade básica das árvores individuais (db_{ij}) foi mais efetiva enquanto que a medida de penetração (p_{ij}) foi a menos efetiva. As demais variáveis densidade básica de literatura (dbl_j) e densidade básica média da espécie (db_j) se colocaram num nível intermediário entre as duas e somente diferiram em agregar plausibilidade no modelo teórico de Schumacher-Hall.

A inclusão da medida de penetração apesar do pior desempenho entre as variáveis preditoras inseridas na análise melhorou muito a plausibilidade em relação aos modelos teóricos na sua forma tradicional, demonstrando enorme potencial quando informações mais acuradas não estão disponíveis.

Tabela 2.3 - Valores de AIC para os modelos de predição da biomassa lenhosa testados neste trabalho.

Relações Funcionais Teóricas					
Husch		Schumacher-Hall		Spurr	
Expressão da média	AIC	Expressão da média	AIC	Expressão da média	AIC
$\beta_0 d_{ij}^{\beta_1}$	1866,46	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	1833,45	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	1856,32
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	1793,12	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	1690,50	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	1692,23
$\beta_0 d_{ij}^{\beta_1} \overline{dbl}_j^{\beta_2}$	1754,71	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{dbl}_j^{\beta_3}$	1666,50	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{dbl}_j^{\beta_2}$	1680,21
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	1755,26	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	1655,85	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	1680,41
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	1723,84	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	1591,30	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	1628,56

2.3.2 Avaliação do desempenho preditivo dos modelos testados

Cenário 1: Parcelas geradas com árvores reamostradas com probabilidade diretamente proporcional à densidade básica

Considerando-se os valores do erro de predição apresentados na Tabela 2.4 e na figura 2.1 pode-se notar que a expressão da média tradicional do modelo de Husch subestima a biomassa do lenho das árvores, enquanto que as expressões que incluem alguma informação acerca da densidade tendem a superestimar. Nota-se também um ligeiro aumento da acurácia do modelo com a entrada das novas variáveis preditoras, não eliminando, porém, seu comportamento notadamente enviesado.

Quanto ao modelo de Schumacher-Hall, a expressão tradicional subestima a biomassa do lenho nas parcelas. Esse viés de subestimativa vai, no entanto, se reduzindo paulatinamente com a entrada dos preditores até atingir um valor próximo de zero quando da inclusão da informação da densidade básica em nível de árvore (Tabela 2.4 e Figura 2.2).

Quanto ao modelo teórico de Spurr, sua expressão tradicional apresenta viés de subestimativa duas vezes maior que a expressão tradicional do modelo de Schumacher-Hall (Tabela 2.4 e Figura 2.3). Esse viés perde força com a entrada dos preditores, se mantendo, no entanto, elevado mesmo com a inclusão da densidade básica em nível de árvore.

O viés pode ser visto como um componente do erro que indica desvios em relação à média zero. No cenário 1 somente o modelo de Schumacher apresentou redução do viés ou melhora na exatidão das predições com a inclusão das informações fornecidas pelos preditores. No entanto, com a entrada desses preditores sempre houve uma redução no desvio padrão do erro médio, indicando aumento de precisão do modelo.

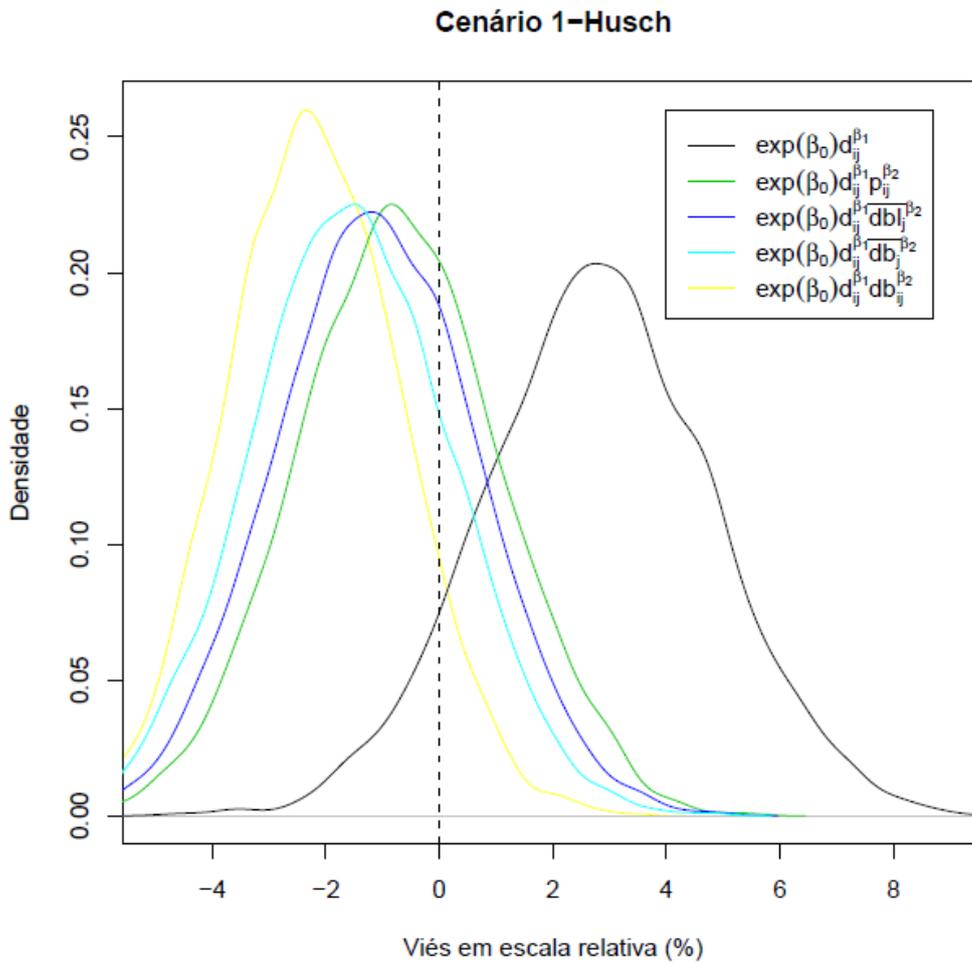


Figura 2.1 - Gráfico de distribuição do erro de predição médio, ou viés, por parcela das diferentes expressões da média do modelo de Husch quando aplicado ao cenário 1, onde as árvores foram amostradas com probabilidade diretamente proporcional a sua densidade básica.

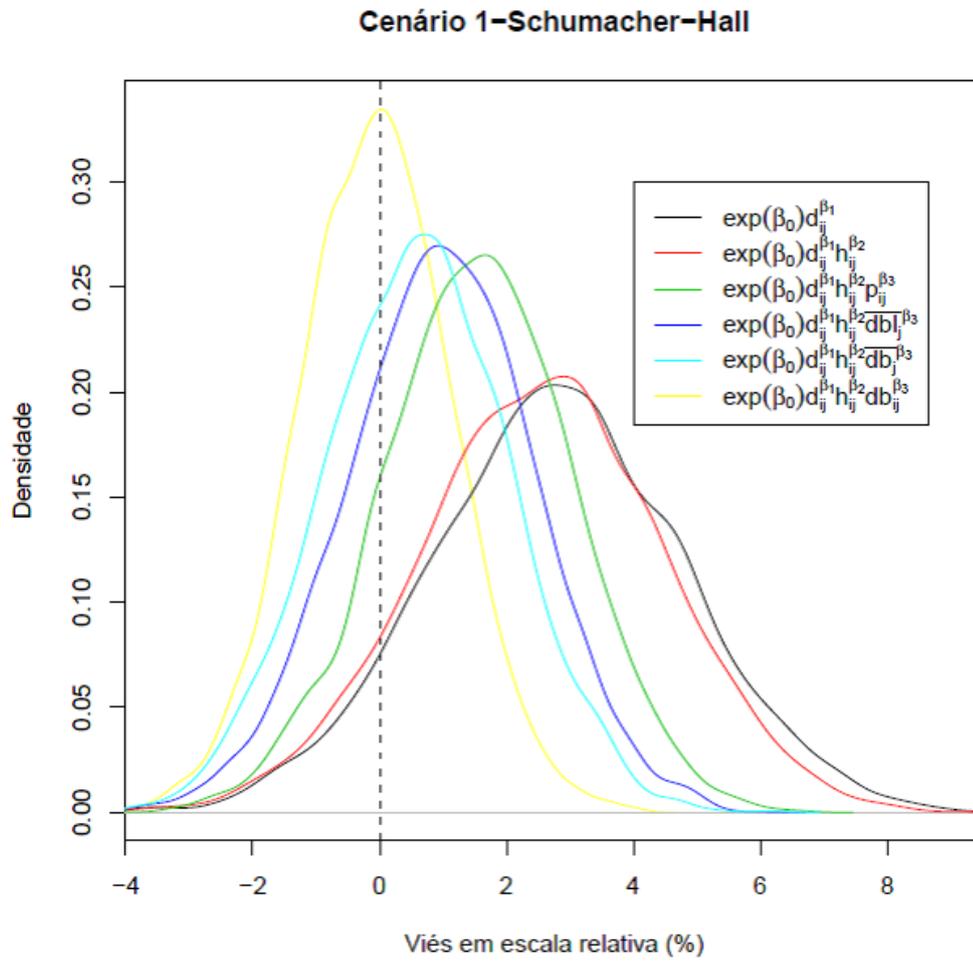


Figura 2.2 - Gráfico de distribuição do erro de previsão médio, ou viés, por parcela das diferentes expressões da média do modelo de Schumacher-Hall quando aplicado ao cenário 1, onde as árvores foram amostradas com probabilidade diretamente proporcional a sua densidade básica.

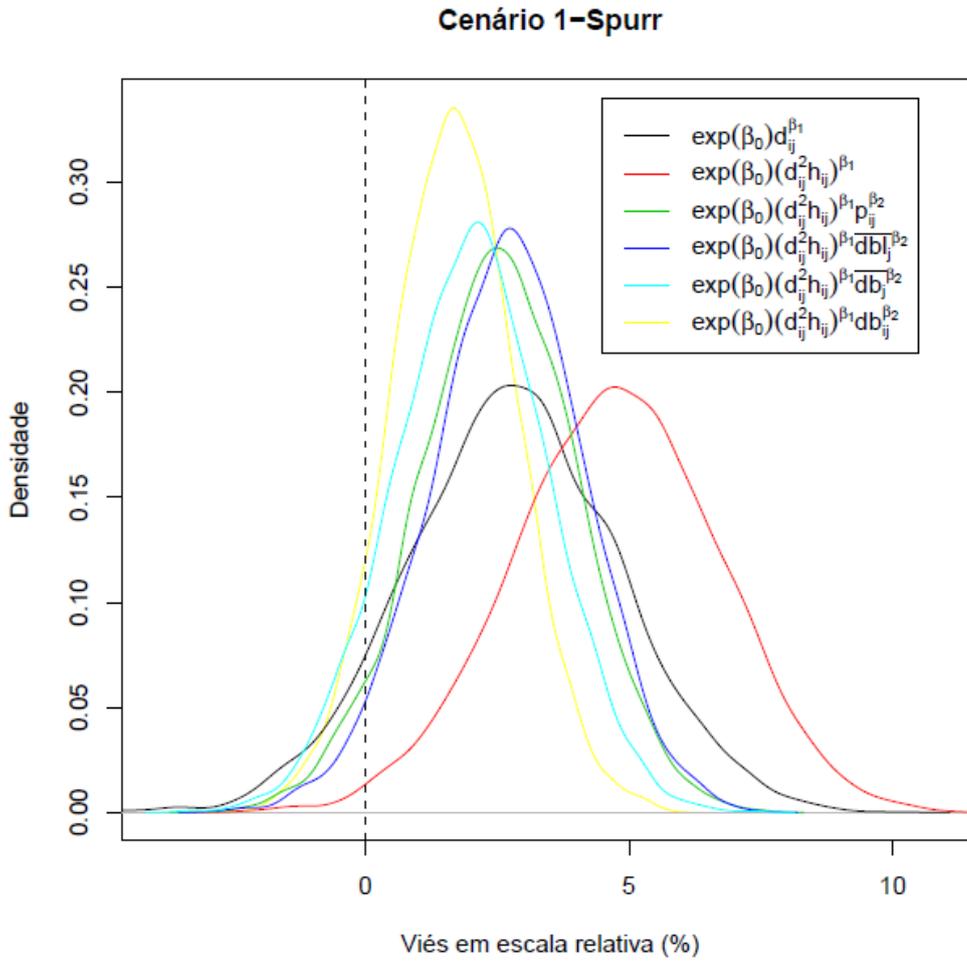


Figura 2.3 - Gráfico de distribuição do erro de predição médio, ou viés, por parcela das diferentes expressões da média do modelo de Spurr quando aplicado ao cenário 1, onde as árvores foram amostradas com probabilidade diretamente proporcional a sua densidade básica.

Tabela 2.4 - Valores do erro de predição médio (Viés) dos modelos de predição do biomassa do lenho quando aplicados às 5000 parcelas do Cenário 1 (árvores amostradas com probabilidade diretamente proporcional à sua densidade básica).

Relações Funcionais Teóricas					
Husch		Schumacher-Hall		Spurr	
<i>Viés e desvio padrão (σ) em escala absoluta (Kg)</i>					
Expressão da média	Viés e σ	Expressão da média	Viés e σ	Expressão da média	Viés e σ
$\beta_0 d_{ij}^{\beta_1}$	6,17 (4,41)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	5,63 (4,25)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	10,47 (4,40)
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	-2,46 (3,87)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	3,44 (3,32)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	5,58 (3,38)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-1,56 (3,70)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	2,13 (3,25)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	5,95 (3,29)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-3,43 (3,86)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	1,30 (3,22)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	4,42 (3,25)
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	-4,83 (3,40)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	-0,10 (2,62)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	3,72 (2,66)
<i>Viés e desvio padrão em escala relativa (%)</i>					
Expressão da média	Viés e σ	Expressão da média	Viés e σ	Expressão da média	Viés e σ
$\beta_0 d_{ij}^{\beta_1}$	2,79 (2,20)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	2,54 (1,92)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	4,72 (1,95)
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	-0,71 (1,78)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	1,54 (1,49)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	2,51 (1,50)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-1,11 (1,76)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	0,95 (1,48)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	2,68 (1,47)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-1,55 (1,75)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	0,58 (1,46)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	1,99 (1,46)
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	-2,18 (1,54)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	-0,05 (1,19)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	1,67 (1,19)

Cenário 2: Parcelas geradas com árvores reamostradas com igual probabilidade

No cenário 2, o modelo teórico de Husch superestima a biomassa lenhosa e a inclusão das informações de densidade básica, na forma de variáveis preditoras, não foi efetiva em corrigir esse viés. O modelo teórico neste caso se mostrou intrinsecamente enviesado para o conjunto de dados deste estudo. No caso do modelo Schumacher-Hall, exceto para a expressão tradicional que demonstrou ligeira tendência à superestimativa, nenhuma outra expressão da média apresentou um viés importante (Tabela 2.5 e Figuras 2.4 e 2.5). Além disso, pode-se observar um aumento da precisão com a inclusão da informação relativa à densidade básica (I) nos seus diferentes níveis.

Com relação ao modelo de Spurr, pode-se dizer que seu comportamento foi semelhante ao de Husch, mas com viés de subestimativa. Este modelo teórico apresentou também um viés intrínseco nas predições de biomassa lenhosa no cenário 2 (Tabela 2.5 e Figura 2.6). Como no cenário 1, a entrada de novos preditores sempre reduziu o desvio padrão da estimativa do erro médio de predição indicando melhora na precisão do modelo.

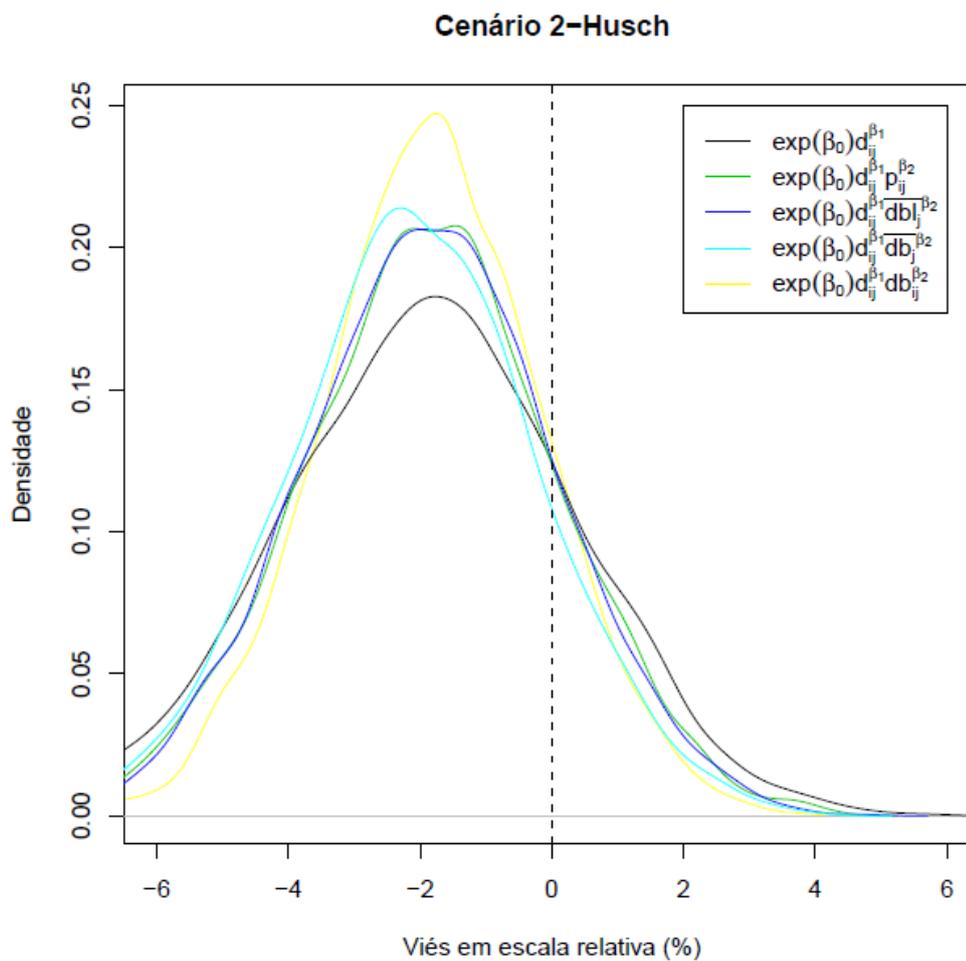


Figura 2.4 - Gráfico de distribuição do erro de previsão médio, ou viés, por parcela das diferentes expressões da média do modelo de Husch quando aplicado ao cenário 2, onde as árvores foram amostradas com igual probabilidade.

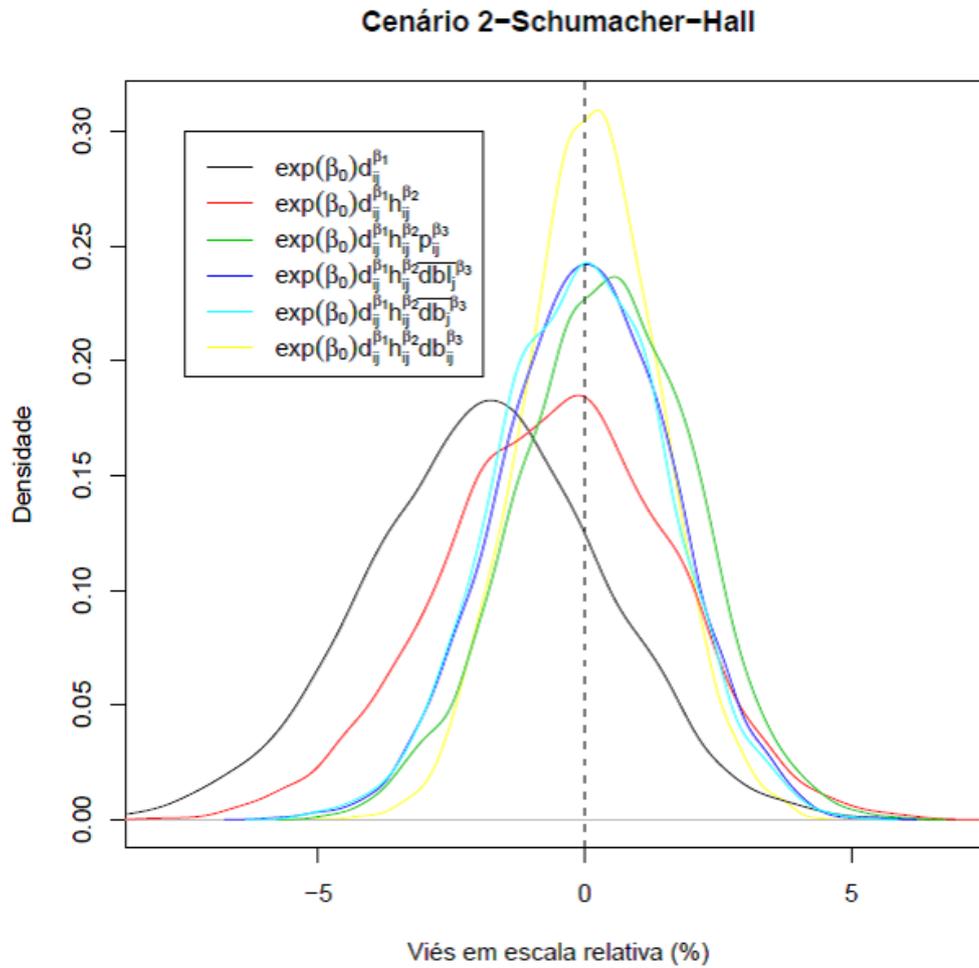


Figura 2.5 - Gráfico de distribuição do erro de predição médio, ou viés, por parcela das diferentes expressões da média do modelo de Schumacher-Hall quando aplicado ao cenário 2, onde as árvores foram amostradas com igual probabilidade.

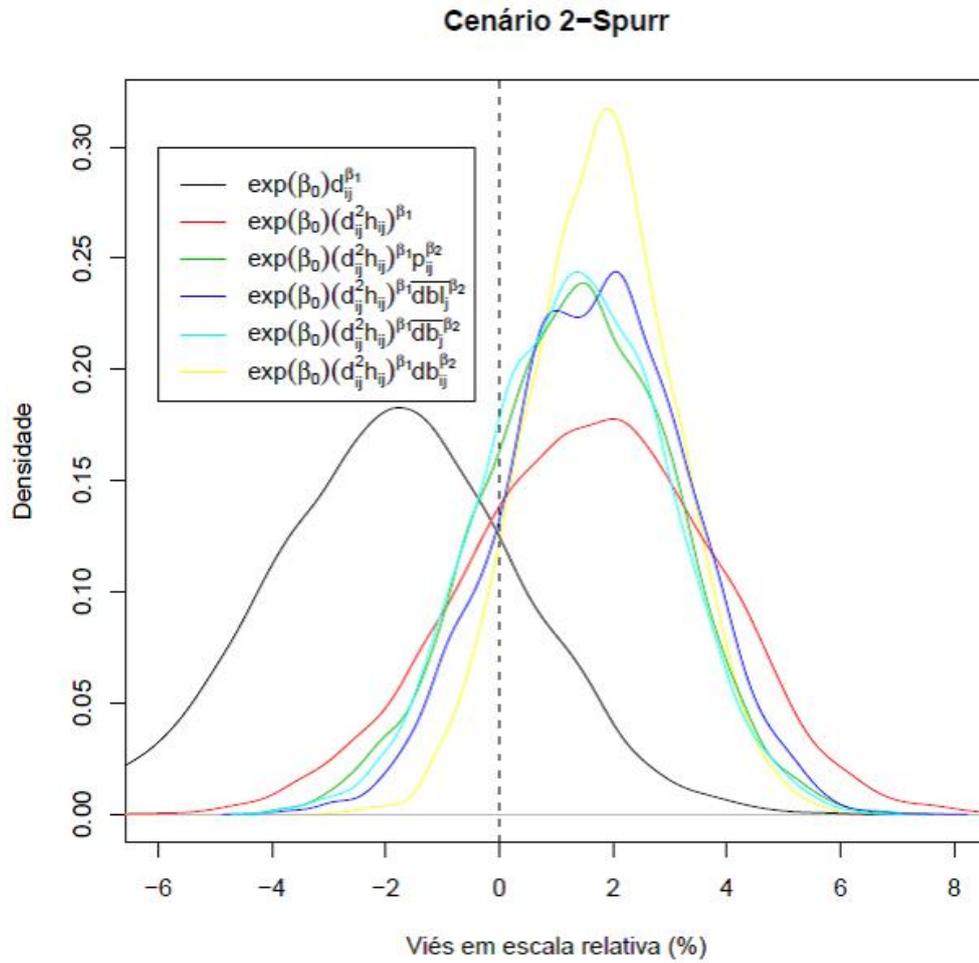


Figura 2.6 - Gráfico de distribuição do erro de predição médio, ou viés, por parcela das diferentes expressões da média do modelo de Spurr quando aplicado ao cenário 2, onde as árvores foram amostradas com igual probabilidade.

Tabela 2.5 - Valores do erro de predição médio (Viés) dos modelos de predição do biomassa do lenho quando aplicados às 5000 parcelas do Cenário 2 (árvores amostradas com igual probabilidade).

Relações Funcionais Teóricas					
Husch		Schumacher-Hall		Spurr	
<i>Viés e desvio padrão(σ) em escala absoluta (Kg)</i>					
Expressão da média	Viés e σ	Expressão da média	Viés e σ	Expressão da média	Viés e σ
$\beta_0 d_{ij}^{\beta_1}$	-3,82 (4,55)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	-1,07 (4,36)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	3,28 (4,51)
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	-3,85 (3,90)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	0,84 (3,38)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	2,86 (3,44)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_{ij}^{\beta_2}$	-3,80 (3,97)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_{ij}^{\beta_3}$	0,03 (3,27)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_{ij}^{\beta_2}$	3,48 (3,31)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-4,32 (3,89)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	-0,10 (3,24)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	2,78 (3,29)
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	-3,74 (3,41)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	0,25 (2,62)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	3,71 (2,67)
<i>Viés e desvio padrão em escala relativa (%)</i>					
Expressão da média	Viés e σ	Expressão da média	Viés e σ	Expressão da média	Viés e σ
$\beta_0 d_{ij}^{\beta_1}$	-1,87 (2,23)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	-0,55 (2,14)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	1,56 (2,18)
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	-1,86 (1,95)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	0,38 (1,65)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	1,36 (1,66)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_{ij}^{\beta_2}$	-1,87 (1,91)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_{ij}^{\beta_3}$	0,01 (1,60)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_{ij}^{\beta_2}$	1,67 (1,60)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-2,10 (1,90)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	-0,06 (1,59)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	1,33 (1,59)
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	-1,81 (1,66)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	0,119 (1,28)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	1,79 (1,28)

Cenário 3: Parcelas geradas com árvores reamostradas com probabilidade inversamente proporcional à densidade básica

Neste cenário (Tabela 2.6 e Figuras 2.7, 2.8 e 2.9) o modelo teórico de Husch apresentou um viés de superestimativa da biomassa. Esse viés vai sendo reduzido a partir da inclusão das variáveis preditoras informativas de densidade básica na expressão que modela a média, sem, no entanto, anular-se completamente. O modelo de Spurr, por sua vez, também apresentou tendência sistemática de superestimativa na expressão tradicional, mas em menor intensidade que os demais modelos teóricos. Com a entrada dos preditores o viés foi reduzido e o modelo passou a apresentar uma tendência sistemática de superestimativa, mas num intensidade menor. Curiosamente, neste cenário, o modelo de Spurr a ordem hierárquica dos níveis de informação de densidade se alterou no que se refere à correção do viés, tendo como a expressão da média que incluiu a medida de penetração como a mais eficiente em realizar predição exatas.

O modelo de Schumacher-Hall apresentou desempenho muito semelhante ao do cenário 1, apenas com tendência de superestimativa ao invés de subestimativa. Da mesma forma a entrada dos preditores informativos da densidade básica foi muito eficiente em reduzir o viés, tendo a expressão da média que incluiu a densidade básica em nível de árvore um comportamento bastante preciso. Assim como nos cenários 1 e 2, a inclusão dos preditores informativos da densidade básica sempre melhorou a precisão do modelo.

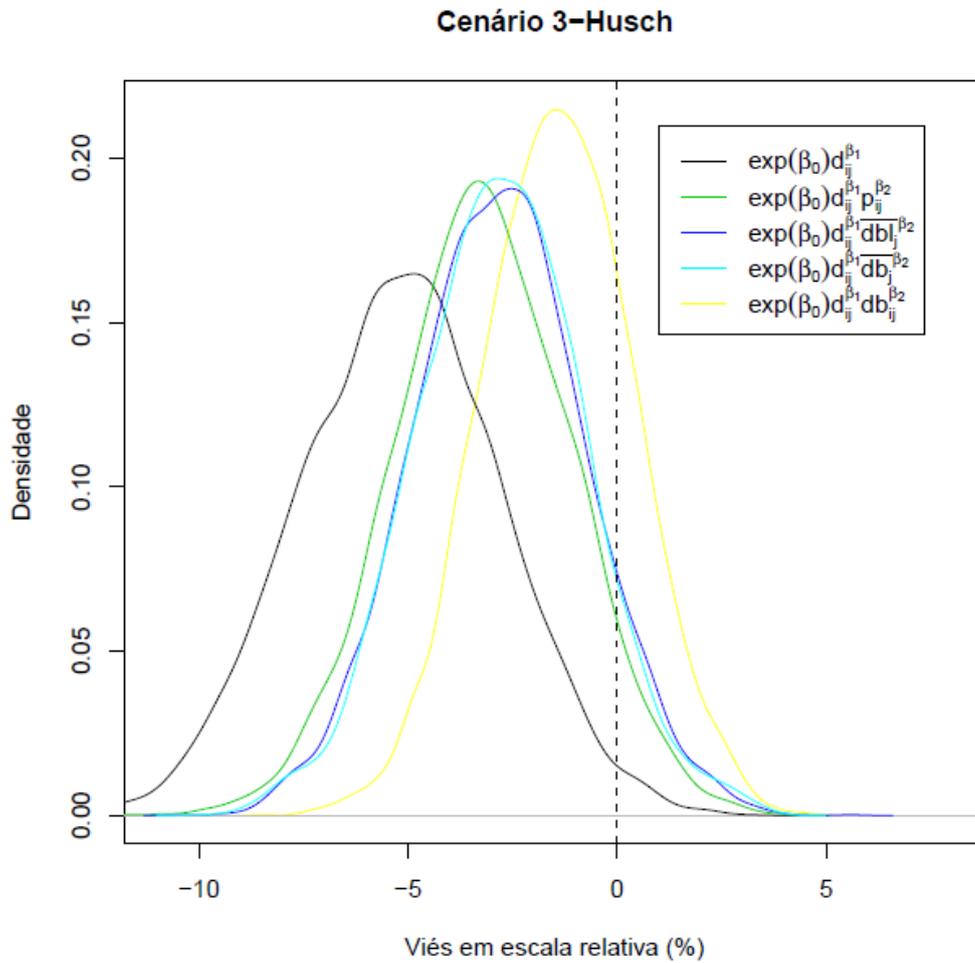


Figura 2.7 - Gráfico de distribuição do erro de predição médio, ou viés, por parcela das diferentes expressões da média do modelo de Husch quando aplicado ao cenário 3, onde as árvores foram amostradas com probabilidade inversamente proporcional a sua densidade básica.

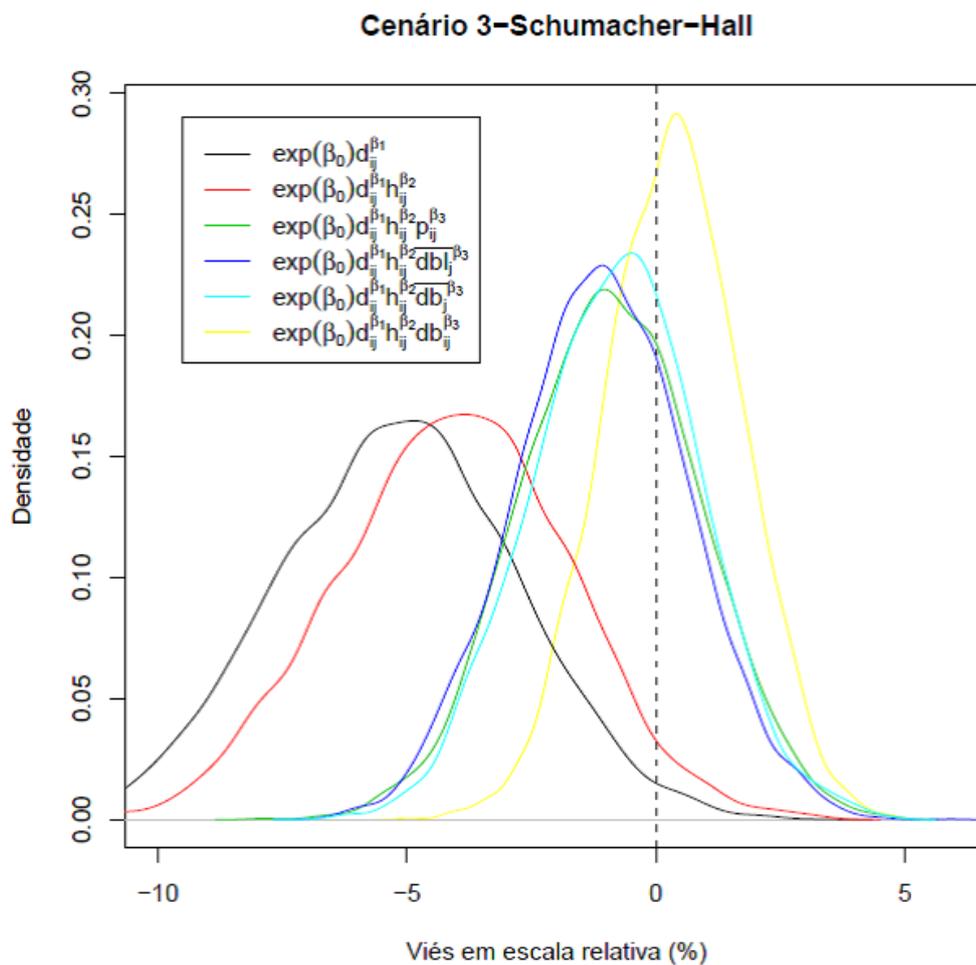


Figura 2.8 - Gráfico de distribuição do erro de previsão médio, ou viés, por parcela das diferentes expressões da média do modelo de Schumacher-Hall quando aplicado ao cenário 3, onde as árvores foram amostradas com probabilidade inversamente proporcional a sua densidade básica.

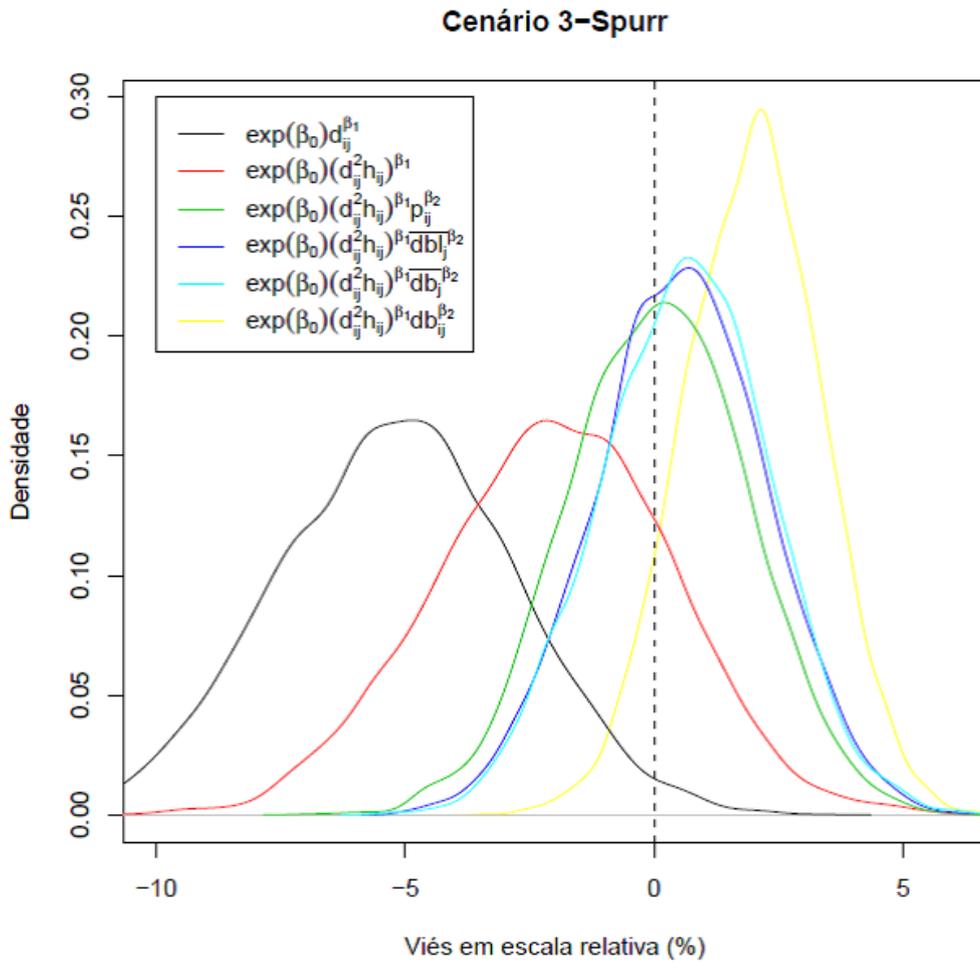


Figura 2.9 - Gráfico de distribuição do erro de predição médio, ou viés, por parcela das diferentes expressões da média do modelo de Spurr quando aplicado ao cenário 3, onde as árvores foram amostradas com probabilidade inversamente proporcional a sua densidade básica.

Tabela 2.6 - Valores do erro de predição médio (viés) dos modelos de predição do biomassa do lenho quando aplicados às 5000 parcelas do Cenário 3 (árvores amostradas com probabilidade inversamente proporcional à sua densidade básica).

Relações Funcionais Teóricas					
Husch		Schumacher-Hall		Spurr	
<i>Viés e desvio padrão(σ) em escala absoluta (Kg)</i>					
Expressão da média	Viés e σ	Expressão da média	Viés e σ	Expressão da média	Viés e σ
$\beta_0 d_{ij}^{\beta_1}$	-9,8 (4,45)	-	-	-	-
-	-	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	-7,66 (4,25)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	-3,76 (4,40)
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	-6,08 (3,94)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	-1,70 (3,33)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	0,19 (3,40)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_{ij}^{\beta_2}$	-5,33 (3,88)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_{ij}^{\beta_3}$	-2,05 (3,23)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_{ij}^{\beta_2}$	1,06 (3,27)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-5,29 (3,85)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	-1,49 (3,19)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	1,19 (3,24)
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	-2,72 (3,41)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	0,63 (3,20)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	3,74 (2,67)
<i>Viés e desvio padrão em escala relativa (%)</i>					
Expressão da média	Viés e σ	Expressão da média	Viés e σ	Expressão da média	Viés e σ
$\beta_0 d_{ij}^{\beta_1}$	-5,20 (2,40)	-	-	-	-
-	-	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	-4,08 (2,32)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	-2,04 (2,38)
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	-3,22 (2,10)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	-0,93 (1,78)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	0,06 (1,79)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_{ij}^{\beta_2}$	-2,81 (2,05)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_{ij}^{\beta_3}$	-1,10 (1,72)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_{ij}^{\beta_2}$	0,53 (1,72)
$\beta_0 d_{ij}^{\beta_1} \overline{db}_j^{\beta_2}$	-2,79 (2,03)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db}_j^{\beta_3}$	-0,81 (1,70)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db}_j^{\beta_2}$	0,60 (1,70)
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	-1,42 (1,79)	$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	0,32 (1,39)	$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	1,95 (1,38)

2.4 Discussão

De maneira geral o Modelo teórico de Husch apresentou um viés de superestimativa em todos os cenários, exceto no cenário 1 onde a expressão da média tradicional apresentou viés de subestimativa. Além disso, as informações de densidade básica foram menos efetivas em corrigir o viés que nos modelos teóricos de Schumacher-Hall e Spurr. Desta forma, pode-se concluir que este modelo tem comportamento intrinsecamente enviesado na predição dos dados em questão. Esse viés natural do modelo teórico dificulta qualquer observação acerca da inclusão das informações de densidade básica.

Quanto ao modelo de Spurr, este também apresentou desvios sistemáticos, mas de subestimativa. O viés intrínseco no modelo de Husch pode ser explicado pela ausência da informação de altura das árvores nas expressões da média. A entrada dos preditores informativos da densidade básica não foi efetiva em corrigir o viés sem a informação da altura total das árvores. No caso do modelo de Spurr, a forma como essa informação é inserida no modelo teórico, alterando a relação funcional, é que parece ter inserido o viés sistemático. Por incluir a altura o modelo de Spurr, apesar do viés advindo da relação funcional, respondeu melhor que o de Husch à entrada das informações de densidade básica na redução do viés.

O modelo de Schumacher-Hall apresentou viés de subestimativa nas predições do cenário 1, de superestimativa no cenário 3 e praticamente nenhum viés no cenário 2. Esse comportamento enviesado nos cenários 1 e 3 se deve ao fato das parcelas nestes cenários terem características distintas, em termos de composição florística, aos dados de construção do modelo. No entanto, nestes cenários distintos a inclusão das informações de densidade básica permitiu ao modelo um comportamento mais preciso em termos de predições.

Apesar do viés intrínseco dos modelos de Husch e Spurr e da baixa resposta à entrada de informações de densidade na redução do viés nestes dois modelos, cabe ressaltar que em todos os três modelos e cenários observou-se uma redução no desvio padrão das estimativas do viés. Isso implica que a inclusão das variáveis predictoras informativas sempre melhorou a precisão da predição dos modelos.

A avaliação do ajuste realizada na seção 2.3.1 já indicava, de certa forma, o desempenho dos modelos em termos das predições. A análise do ajuste só não deixava tão clara a diferença de desempenho entres os modelos teóricos de Schumacher-Hall e Spurr. Esses resultado ressalta a importância da realização de análise de desempenho de aplicação

dos modelos em termos preditivos, mesmo em situações simuladas como as do presente trabalho.

Levando em consideração os resultados de ajuste e desempenho das expressões da média dentre dos modelos teóricos, pode-se sugerir, considerando-se principalmente o modelo sem viés intrínseco, a seguinte ordem crescente de eficiência das variáveis predictoras em reduzir o viés sistemático do modelo em situações preditivas: Medida de penetração no lenho, densidade básica da espécie retirada da literatura, densidade básica média determinada em laboratório e densidade básica determinada em laboratório em nível de árvore.

Refletindo sobre essa ordem hierárquica, tem-se que considerar que a densidade básica média (db_j) e a densidade básica em nível de árvore (db_{ij}) são difíceis de obter sem amostras destrutivas e, portanto, são de difícil aplicação em inventários tradicionais. A densidade básica da literatura (dbl_j) pode ser encontrada para várias espécies arbóreas tropicais, mas não para todas elas, além de exigir a identificação das espécies no campo no processo de construção e aplicação do modelo nos inventários.

A medida de penetração, por sua vez, é uma medida de fácil obtenção no campo e que, segundo os resultados deste trabalho, melhora muito a qualidade do ajuste e das predições dos modelos, principalmente em situações de variação espacial no estoque de biomassa. Há de se considerar que o viés das predições do modelo que incluiu essa medida como variável preditora ficou ao redor de 2%, que apesar de sistemático é bastante aceitável quando comparado ao desempenho dos modelos tradicionais.

No cenário 2 a utilização desses preditores foi pouco relevante, já que, a amostra de predição é semelhante à de ajuste ou de construção dos modelos. No entanto, essa situação é pouco realista, já que se sabe que a composição das parcelas de inventário não segue às da amostra de ajuste. As situações criadas nos cenários 1 e 3 são mais realistas do ponto de vista da aplicação do modelo permitindo uma melhor análise do desempenho destes em situações em que ocorrem variações espaciais no estoque de biomassa em curtas distâncias. Desta forma, se recomenda fortemente a utilização de informações de densidade básica das árvores na construção de modelos preditivos de biomassa, a fim de melhorar a robustez e a precisão desses na aplicação em inventários de biomassa.

A avaliação conjunta de ajuste e desempenho preditivo, realizada neste trabalho, demonstra o caráter eminentemente empírico desses modelos, já observado por Zianis e Radoglou (2006). Isto porque no processo de construção dos modelos, a robustez na aplicação, foi obtida a partir da entrada de informações retiradas da população como altura e

a densidade básica nos seus diferentes níveis, e não de parâmetros fixos de alguma relação alométrica.

2.5 Conclusões

A análise dos modelos de predição de biomassa, no que diz respeito ao ajuste e aplicação, realizada de maneira integrada, demonstrou ser muito importante para avaliação da precisão e robustez em situações de campo. Essa avaliação integrada permitiu verificar o comportamento intrinsecamente enviesado dos modelos teóricos de Husch e Spurr, além de verificar o pouco efeito da entrada das informações de densidade básica nestes modelos.

O modelo de Schumacher-Hall foi o único entre os testados a apresentar um comportamento sem viés intrínseco e onde, os preditores informativos de densidade básica, foram efetivos em reduzir o viés nas situações de aplicação simuladas e distintas às da amostra de ajuste. Dentre esses preditores pôde-se observar certa hierarquia em fornecer informação de densidade básica e, portanto, em reduzir o viés e aumentar a precisão das predições. Essa hierarquia foi em ordem crescente de eficiência: Medida de penetração no lenho, densidade básica da espécie retirada da literatura, densidade básica média determinada em laboratório e densidade básica determinada em laboratório em nível de árvore.

Como as situações de campo são distintas às de construção dos modelos, o uso da informação de densidade básica como variável preditora, em qualquer um dos níveis usados neste trabalho, é altamente recomendada. Isso porque, na Floresta Tropical se observa um gradiente espacial de variação no estoque de biomassa bastante relevante.

Dentre os preditores usados neste trabalho, a medida de penetração no lenho se destaca pelo seu desempenho em melhorar a acurácia das predições e pela sua facilidade de obtenção no campo.

Os resultados deste trabalho demonstraram, ainda, a natureza eminentemente empírica dos modelos de predição de biomassa, já que, evidenciaram a importância do processo de construção dos modelos, representada pelo tipo de dados utilizados na construção, refletindo na capacidade preditiva dos mesmos.

Referências

ALVES, L.F.; VIEIRA, S.A.; SCARANELLO, M.A.; CAMARGO, P.B.; SANTOS, F.A.M.; JOLY, C.A.; MARTINELLI, L.A. Forest structure and live aboveground biomass variation along an elevational gradient of tropical Atlantic moist forest (Brazil). **Forest Ecology and Management**, Winsconsin, v. 260, n. 5, p. 679-691, July 2010.

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.

BATTAGLIA, M.; SANDS, P.J. Process-based forest productivity models and their application in forest management. **Forest Ecology and Management**, Wisconsin, v. 102, n. 1, p. 13-32, Mar. 1998.

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

BURGER, D.M. **Modelos alométricos para a estimativa da fitomassa de Mata Atlântica na Serra do Mar, SP**. 2005. 112 p. Tese (Doutorado em Ecologia) – Instituto de Biociências, Universidade de São Paulo, São Paulo, 2005.

BURGER, D.M.; DELITTI, W.B.C. Allometric models for estimating the phytomass of a secondary Atlantic Forest area of southeastern Brazil. **Biota Neotropica**, Campinas, v. 8, n. 4, p. 131-136, Oct./Dec. 2008.

_____. Modelos preditores da fitomassa aérea da Floresta Baixa de Restinga. **Revista Brasileira de Botânica**, São Paulo, v. 33, n.1, p. 143-153, 2010.

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

CHAMBERS, J.Q.; SANTOS, J.; RIBEIRO, R.J.; HIGUCHI, N. Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. **Forest Ecology and Management**, Wisconsin, v. 152, p. 73-84, Oct. 2001.

CHAVE, J.; RIÉRA, B.; DUBOIS, M. Estimation of biomass in a neotropical forest of French Guiana: spatial and temporal variability. **Journal of Tropical Ecology**, Winchelsea, v. 17, p. 79- 96, Jan. 2001.

CHAVE, J.; ANDALO, C.; BROWN, S.; CAIRNS, M.; CHAMBERS, J.C.; EAMUS, D., FÖLSTER, H.; FROMARD, F.; HIGUCHI, N.; KIRA, T.; LESCURE, J.; NELSON, B.W.; OGAWA, H.; PUIG, H.; RIÉRA, B.; YAMAKURA, T. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. **Oecologia**, Heidelberg, v. 145, n. 1, p. 87-99, June 2005.

FEARNSIDE, P.M. Wood density for estimating forest biomass in Brazilian Amazonia. **Forest Ecology and Management**, Winsconsin, v. 90, n. 1, p. 59-87, Jan. 1997.

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.

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

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**, Wisconsin, v. 117, n. 1/3, p. 149-167, May 1998.

NOGUEIRA, E.M.; FEARNSTIDE, 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, Sept. 2007

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

PRISLEY, S.P.; MORTIMER, J.P. A synthesis of literature on evaluation of models for policy applications, with implications for forest carbon accounting, **Forest Ecology and Management**, Wisconsin, v. 198, n. 1/3, p. 89-103, Aug. 2004.

SAATCHI, S.S.; HOUGHTON, R.A.; DOS SANTOS ALVALÁ, R.C.; SOARES, J.V.; YU, Y. Distribution of aboveground live biomass in the Amazon basin. **Global Change Biology**, New York, v. 13, p. 816–837, 2007.

SAMPAIO, E.; GASSON, P.; BACARAT, A.; CUTLER, D.; PAREYN, F.; LIMA, K.C. Tree biomass estimation in regenerating areas of tropical dry vegetation in northeast Brazil, **Forest Ecology and Management**, Wisconsin, v. 259, n. 6, p. 1135-1140, Mar. 2010.

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**: FUPEF/ Imprensa da UFPR, 2002. cap. 6, p. 119-140.

SANTOS, J.; PAULA NETO, F.; HIGUCHI, 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./dez. 2001.

SEGURA, M. Allometric models for tree volume and total aboveground biomass in a tropical humid forest in Costa Rica. **Biotropica**, Malden, v. 37, n. 1, p. 2-8, Mar. 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./out. 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.; HIGUCHI, 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**, Campinas, v. 8, n. 2, p. 21-29, June 2008.

VISMARA, E.S. **Mensuração da biomassa e seleção de modelos para construção de equações de biomassa**. 2009. 102 p. Dissertação (Mestrado em Recursos Florestais) –Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2009.

ZIANIS, D.; RADOGLU, K. Comparison between empirical and theoretical biomass allometric models and statistical implications in stem volume predictions. **Forestry**, Oxford, v. 79, p. 477–487, 2006.

3 APLICAÇÃO DE MODELOS LINEARES DE EFEITO MISTO E CALIBRAÇÃO NA PREDIÇÃO DE VOLUME EM PLANTIOS REGIONAIS DE *EUCALYPTUS GRANDIS*

Resumo

Este trabalho apresenta aplicações da calibração do modelo linear de efeito misto na predição do volume em plantios de *Eucalyptus grandis* em primeira e segunda rotação, situados em diferentes fazendas de uma mesma região do estado de São Paulo, Brasil. Partiu-se do modelo de Schumacher e Hall, em sua forma linearizada, para o desenvolvimento de um modelo de efeitos mistos que considerou alguns de seus parâmetros como sendo aleatórios ao longo das diferentes fazendas. A calibração foi realizada em nível de fazenda partindo-se de um pequeno número de árvores-amostra. A abordagem foi desenvolvida para modelos univariados de primeira rotação e bivariados de duas rotações). Os resultados mostram que o procedimento de calibração de modelos de efeitos mistos fornece predições mais confiáveis do que os modelos tradicionais de efeitos fixos em ambas as rotações, além de indicar uma redução nos esforços de medição nos plantios de segunda rotação.

Palavras-chave: Modelos de efeitos mistos; Modelos de predição de volume; Calibração; BLUP

Abstract

This work presents applications of the linear mixed model calibration to predict volumes of *Eucalyptus grandis* plantations on first and second rotation located in different farms of the same region in Sao Paulo State, Brazil. We started with the Schumacher and Hall equation in its linearized form to develop our mixed effects model. Some parameters were considered as random among the different farms and the calibration were made at farm level using a small number of sample trees. The approaches were developed for univariate models of the first rotation and bivariate models of the two rotations. The results show that the mixed models and calibration provides more reliable predictions than traditional fixed effects models in both rotations, reducing the measurement efforts on second rotation crops.

Keywords: Mixed-effects models; Volume models; Calibration; BLUP

3.1 Introdução

Em biometria florestal, modelos de predição são amplamente utilizados em situações em que a obtenção das variáveis dendrométricas é de alto custo ou quando envolve a utilização de amostras destrutivas. Tais modelos são muito usados na predição do volume de árvores individuais dentro do contexto do inventário florestal (HUSCH et al., 2003).

As plantações de eucalipto consistem num dos sistemas florestais mais eficientes em termos de produção de madeira, com cerca de 5 milhões de hectares, intensivamente

manejados no Brasil. A maioria dessas plantações pertencem a empresas de celulose e papel, que possuem extensas áreas espalhadas por diferentes regiões. Muitas espécies plantadas por essas empresas brotam após o corte raso. Portanto, é possível gerar múltiplas rotações do mesmo ciclo de produção reduzindo os custos de cultivo. Em comparação com a primeira rotação, os volumes e formas das árvores nas subsequentes rotações podem ser extremamente variáveis (DIAZ-BALTEIRO; RODRIGUES, 2006). Sendo assim, para se obter boas predições de volume, as empresas necessitam de modelos que possam ser aplicados numa escala regional e em plantios de todas as rotações.

Um modelo de predição de volume regional muitas vezes exige um grande e representativo conjunto de dados para a etapa de construção e ajuste. Embora este tipo de conjunto de dados tenha uma estrutura hierárquica natural, representada pelos diferentes municípios, talhões, parcelas, etc, os modelos de regressão tradicionais não levam em conta a possível variação dos parâmetros, entre os diferentes grupos hierárquicos. Os modelos de efeitos mistos, por outro lado, podem lidar com essa variação, considerando alguns dos parâmetros dos modelos como sendo aleatórios (PINHEIRO; BATES, 2000).

Na abordagem de efeitos mistos, os parâmetros do modelo podem variar entre os grupos e, com algumas poucas árvores-amostra, o modelo pode ser calibrado para um determinado grupo afim melhorar a qualidade das predições dentro do grupo em questão. A idéia da calibração de modelos de efeitos mistos para um determinado grupo foi aplicada com sucesso na área florestal em diversos tipos de modelos preditivos, tais como: modelos de volume e funções de afilamento (LAPPI 1986; EERIKÄINEN, 2001), modelos de altura em função do diâmetro (LAPPI, 1997; CALAMA; MONTERO 2004; MEHTÄTALO, 2005a, DORADO et al., 2006), modelos de altura dominante (LAPPI; BAILEY, 1988; CALEGARIO et al., 2005; PESONEN et al., 2009) e modelos de distribuição de diâmetro (MEHTÄTALO, 2005b; MEHTÄTALO et al., 2010). A maioria desses trabalhos, no entanto, aplica apenas a abordagem univariada nos procedimentos de modelagem.

Aplicação da abordagem do modelo de efeitos mistos para o caso multivariado com diversas variáveis respostas foi apresentado por Lappi (1991) (ver também LAPPI et al., 2006). Desde então, a abordagem multivariada foi aplicada a modelos de crescimento e produção a partir da utilização de sistemas de equações compatíveis (FANG et al., 2001; WANG et al., 2011) e melhorando estimativas do inventário baseado em escaneamento a laser aerotransportado (MALTAMO et al., 2012) (no prelo).

A hipótese do presente trabalho é a de que a aplicação de um sistema multivariado constituído por equações de ambas as rotações possa ser útil para melhorar as predições de

volume nas florestas de Eucalipto em primeira e em segunda rotação. Além disso, acredita-se que esta abordagem possa ser usada para reduzir os esforços de medição nas predições de volume em ambas as rotações.

O objetivo deste trabalho é apresentar aplicações da calibração do modelo de efeito misto na predição de volume de árvores individuais num contexto de um inventário florestal regional. A abordagem foi desenvolvida para modelos univariados de uma rotação e bivariados de duas rotações, com aplicações práticas em quaisquer plantios florestais de produção. A abordagem é demonstrada para florestas de *Eucalyptus grandis* em primeira e segunda rotação localizadas em diferentes municípios de uma mesma região do estado de São Paulo, Brasil.

3.2 Material

O conjunto de dados é proveniente de plantios de *Eucalyptus grandis* em primeira e segunda rotação localizados em quatro municípios da região sudeste do Estado de São Paulo, Brasil. As plantações estão distribuídas em diferentes unidades de produção, chamadas fazendas. A partir destas fazendas, árvores foram amostradas e medições de diâmetro à altura do peito (DAP), altura total e volume total foram feitas em cada fuste. O volume total foi obtido por análise de tronco a partir da aplicação da fórmula de Smalian, e com os diâmetros medidos a cada metro ao longo do fuste.

Tabela 3.1 - Distribuição dos fustes nas diferentes fazendas, municípios e rotações. * Indica as fazendas que têm povoamentos em primeira e segunda rotação.

Município	Fazendas	Primeira Rotação	Segunda Rotação
		Número de fustes	
Bofete	6*	256	87
	9*	34	33
	25*	213	78
	35*	124	104
Botucatu	28*	85	20
	36*	68	17
	42	87	-
Itatinga	16	31	-
	17	70	-
	20	35	-
	23*	24	35
Salto	1*	110	69
	2	113	-
	3	50	-
	4	131	-
	5	70	-
Total		1501	443
Total*		921*	443*

As fazendas possuem tamanhos diferentes e o número de fustes amostrados em cada fazenda foi proporcional à área da mesma. Este procedimento produziu uma distribuição desequilibrada da amostra entre as fazendas, como pode ser visto na Tabela 3.1. A amplitude de variação das variáveis que compõem o conjunto de dados pode ser vista na Tabela 3.2.

Ao todo, 16 fazendas foram utilizadas, sendo 16 fazendas com 1.501 fustes em primeira rotação e 8 fazendas com 443 fustes em segunda rotação. No entanto, considerando apenas as fazendas que continham plantios em ambas as rotações, esse total é reduzido para 8 fazendas com 921 fustes em primeira rotação e com 443 fustes em segunda rotação.

Tabela 3.2 – Algumas informações sobre o conjunto de dados, onde * indica fazendas com fustes em primeira e segunda rotação

Fazenda	Volume total (dm ³)			DAP (cm)			Altura total (m)		
	Min.	Média	Max.	Min.	Média	Max.	Min.	Média	Max.
<i>Primeira Rotação</i>									
Bofete-6*	8,9	104,8	759,2	4,5	11,5	26,7	9,7	18,1	32,9
Bofete-9*	15,7	183,5	588,6	5,4	13,9	25,5	10,9	20,6	31,2
Bofete-25*	2,6	253,8	2012,1	3,8	13,4	38,2	7,1	19,5	44,9
Bofete-35*	5,3	60,7	229,2	3,8	9,7	17,5	7,9	14,6	21,4
Botucatu-28*	6,0	169,3	568,4	4,1	13,0	23,2	8,8	21,0	30,9
Botucatu-36*	26,5	159,2	484,7	7,0	13,8	22,9	14,1	21,9	29,0
Botucatu-42	4,6	43,0	93,6	3,8	9,3	13,0	8,0	12,6	15,9
Itatinga-16	6,4	111,2	530,8	4,1	11,6	24,2	8,3	18,9	27,2
Itatinga-17	5,0	33,9	84,1	4,4	8,8	13,0	5,4	11,4	15,6
Itatinga-20	5,5	148,3	653,3	4,1	12,2	22,9	8,7	18,7	33,5
Itatinga-23*	20,7	148	281,6	6,4	13,1	18,8	13,8	22,8	29,0
Salto-1*	11,7	94,3	270,3	5,4	11,9	18,5	9,5	17,5	23,5
Salto-2	8,0	59,0	118,6	4,5	10,1	13,4	9,4	15,7	19,8
Salto-3	8,1	72,7	158,7	4,5	10,1	14,6	9,6	16,3	21,8
Salto-4	5,7	80,0	226,2	3,8	10,8	16,5	8,6	16,4	23,4
Salto-5	12,1	61,4	127,4	5,4	10,4	14,6	9,9	14,8	18,7
Todas	4,6	116,1	2012,1	3,8	11,4	38,2	5,4	17,3	44,9
Todas*	4,7	145,7	2012,1	3,8	12,1	38,2	7,1	18,6	44,9
<i>Segunda Rotação</i>									
Bofete-6*	4,2	21,6	66,6	3,8	7,2	12,4	5,5	9,6	13,4
Bofete-9*	3,2	14,6	46,6	3,8	6,2	12,1	4,9	8,3	10,9
Bofete-25*	3,9	53,2	135,6	3,8	9,3	14,6	6,3	15,0	18,9
Bofete-35*	2,8	14,9	53,5	3,8	6,1	10,8	4,9	8,7	13,1
Botucatu-28*	6,0	27,3	52,5	3,8	7,6	10,8	6,1	10,7	13,6
Botucatu-36*	1,2	17,9	38,6	4,1	7,6	11,5	6,3	9,3	12,2
Itatinga-23*	4,8	22,3	57,1	3,8	7,3	11,8	6,8	10	13,3
Salto-1*	5,8	70,7	316,2	4,1	10,0	20,5	6,6	13,9	24,2
Todas	1,2	32,6	316,2	3,8	7,7	20,1	4,9	10,9	24,2

3.3 Métodos

3.3.1 Os modelos de volume

Para o desenvolvimento do modelo, partiu-se da função não linear proposta por Schumacher e Hall (1933), que descreve o volume de árvores individuais em função do seu diâmetro à altura do peito (dap) e da sua altura (ht):

$$v = \beta_0 dap^{\beta_1} ht^{\beta_2} \quad (3.1)$$

No entanto, como a predição de efeitos aleatórios é mais conveniente utilizando os modelos lineares, o modelo proposto foi linearizado por transformação logarítmica. Assim, o modelo para predição do volume da i -ésima árvore é

$$\ln v_i = \beta_0 + \beta_1 \ln dap_i + \beta_2 \ln ht_i + \epsilon_i, \quad (3.2)$$

onde v_i é o volume total da árvore (litros); dap_i é o diâmetro à altura do peito (cm); ht_i é a altura total da árvore (m), β_0 , β_1 , β_2 são os parâmetros do modelo e, ϵ_i é o erro aleatório residual, $\epsilon_i \sim NID(0, \sigma^2)$. Este é o modelo linear de referência que foi utilizado neste trabalho para se obter as predições sobre a média marginal. No entanto, este modelo não se justifica como modelo final, pois não leva em conta a estrutura hierárquica do conjunto de dados causada pela distribuição das árvores nas diferentes fazendas.

A partir do modelo marginal (eq. 3.2), assume-se que a sua forma funcional é fixa, mas que alguns dos seus parâmetros podem variar de acordo com as diferentes fazendas. Com isso, ajustou-se um modelo linear de efeitos mistos para a predição do volume. Para desenvolver tal modelo, testaram-se várias combinações de parâmetros para compor a parte aleatória do modelo de efeitos mistos. A seleção do modelo de efeitos mistos final, entre os candidatos foi feita através do teste de razão de verossimilhança e dos valores do critério de informação de Akaike (AIC). Assim, o modelo linear de efeitos mistos para a árvore i da fazenda k tem a seguinte forma:

$$\ln v_{ki} = \beta_0 + \beta_1 \ln dap_{ki} + \beta_2 \ln ht_{ki} + \alpha_{0k} + \alpha_{2k} \ln ht_{ki} + \epsilon_{ki}, \quad (3.3)$$

onde α_{0k} e α_{2k} são os normalmente distribuídos efeitos aleatórios em nível de fazenda.

Para uma apresentação formal da predição de efeito aleatório, precisou-se reformular o modelo, apresentando-o na forma matricial para uma única fazenda. Assim, o modelo para a fazenda k pode ser representado como:

$$\ln v_k = \mu_k + Z_k \alpha_k + \epsilon_k, \quad (3.4)$$

onde o vetor $\ln v_k$ inclui os volumes das árvores observadas na fazenda k , μ_k inclui as predições com base na parte fixa do modelo de efeitos aleatórios, fornecendo o valor esperado do vetor $\ln v_k$, Z_k é a matriz que define a parte aleatória, $\alpha_k = (\alpha_{0k}, \alpha_{2k})'$ inclui os efeitos aleatórios, grupo-específicos, que têm distribuição normal com média 0 e matriz de variância e covariância matriz $D_{2 \times 2}$, definida por

$$D = \text{var}(\alpha_k) = \text{var} \begin{bmatrix} \alpha_{0k} \\ \alpha_{2k} \end{bmatrix} = \begin{bmatrix} \text{var}(\alpha_{0k}) & \text{cov}(\alpha_{0k}, \alpha_{2k}) \\ \text{cov}(\alpha_{2k}, \alpha_{0k}) & \text{var}(\alpha_{2k}) \end{bmatrix},$$

e ϵ_k é o erro aleatório residual também normalmente distribuído com média 0 e variância $\text{var}(\epsilon_k) = R = \sigma^2 I$.

A matriz D , bem como a estrutura de R é comum à todas as fazendas, e são estimadas na fase de ajuste do modelo. No entanto, os valores dos vetores de efeitos aleatórios são preditos posteriormente usando o melhor preditor linear não enviesado (BLUP).

3.3.2 O modelo multivariado

Como se tem duas rotações, o modelo 3.3 foi ajustado para as k fazendas, nas duas rotações, resultando no seguinte sistema de equações:

$$\begin{aligned} \ln v_{fki} &= \beta_0 + \beta_1 \ln dap_{fki} + \beta_2 \ln ht_{fki} + \alpha_{0k} + \alpha_{2k} \ln ht_{fki} + \epsilon_{fki}, \\ \ln v_{ski} &= \beta_0 + \beta_1 \ln dap_{ski} + \beta_2 \ln ht_{ski} + \alpha_{0k} + \alpha_{2k} \ln ht_{ski} + \epsilon_{ski} \end{aligned} \quad (3.5)$$

Onde o índice f indica primeira rotação e o índice s indica a segunda rotação. O modelo 3.5 tem uma estrutura de um sistema de modelos aparentemente não relacionados. Portanto, as estimativas de correlação entre os efeitos aleatórios e os termos de erro dentro das k fazendas, em ambas as rotações, vêm de um procedimento de estimação simultânea (FIEUWS; VERBEKE, 2006). Para isso foi empregada uma abordagem em que uma "resposta" foi adicionada aos dados na forma de um nível hierárquico adicional. Especificando as matrizes do modelo, bem como as estruturas de variância-covariância, separadamente para cada nível das respostas, torna possível que o ajuste do modelo de efeitos mistos multivariado fosse realizado por procedimentos de ajuste univariados (GOLDSTEIN, 2011).

Para formular o preditor de efeitos aleatórios, o sistema 3.5 pode ser escrito, para um determinado grupo k , numa forma matricial, como

$$\begin{aligned}\ln v_{fk} &= \mu_{fk} + Z_{fk}\alpha_{fk} + \epsilon_{fk}, \\ \ln v_{sk} &= \mu_{sk} + Z_{sk}\alpha_{sk} + \epsilon_{sk},\end{aligned}\tag{3.6}$$

Então a nova matriz $D_{4 \times 4} = \text{var}(\alpha_k)$ se torna

$$\begin{aligned}D &= \text{var} \begin{bmatrix} \alpha_{f0k} \\ \alpha_{f2k} \\ \alpha_{s0k} \\ \alpha_{s2k} \end{bmatrix} \\ &= \begin{bmatrix} \text{var}(\alpha_{f0k}) & \text{cov}(\alpha_{f0k}, \alpha_{f2k}) & \text{cov}(\alpha_{f0k}, \alpha_{s0k}) & \text{cov}(\alpha_{f0k}, \alpha_{s2k}) \\ \text{cov}(\alpha_{f0k}, \alpha_{f2k}) & \text{var}(\alpha_{f2k}) & \text{cov}(\alpha_{f2k}, \alpha_{s0k}) & \text{cov}(\alpha_{f2k}, \alpha_{s2k}) \\ \text{cov}(\alpha_{s0k}, \alpha_{f0k}) & \text{cov}(\alpha_{s0k}, \alpha_{f2k}) & \text{var}(\alpha_{s0k}) & \text{cov}(\alpha_{s0k}, \alpha_{s2k}) \\ \text{cov}(\alpha_{s2k}, \alpha_{f0k}) & \text{cov}(\alpha_{s2k}, \alpha_{f2k}) & \text{cov}(\alpha_{s0k}, \alpha_{s2k}) & \text{var}(\alpha_{s2k}) \end{bmatrix}\end{aligned}$$

Para utilização posterior particiona-se a matriz, tal qual mostrado a seguir:

$$\begin{aligned}D &= \begin{bmatrix} D_1 & D_{12} \\ D_{21} & D_2 \end{bmatrix} \\ &= [C \ H].\end{aligned}\tag{3.7}$$

3.3.3 Calibração dos modelos de efeitos mistos para as predições condicionais

3.3.3.1 A melhor predição linear, para o caso geral.

Para prever os efeitos aleatórios, foi usado o melhor preditor linear não enviesado (BLUP). No caso geral, considerando um vetor não observado b_1 e um vetor observado b_2 , o BLUP para b_1 é:

$$\widehat{b}_1 = \mu_1 + V_{12}V_2^{-1}(b_2 - \mu_2), \quad (3.8)$$

com variância

$$\text{var}(\widehat{b}_1 - b_1) = V_1 + V_{12}V_2^{-1}V_{12}', \quad (3.9)$$

onde \widehat{b}_1 é a predição do vetor não observado, μ_1 e μ_2 são as médias de b_1 e b_2 , V_1 e V_2 são as variâncias de b_1 e b_2 e V_{12} é a covariância entre V_1 e V_2 . Neste estudo, a forma geral do BLUP foi adaptada para as três situações apresentadas a seguir, que são posteriormente referenciadas como abordagens de 1 a 3.

3.3.3.2 Calibração do modelo de primeira rotação com árvores-amostra de primeira rotação (Abordagem 1)

Na primeira abordagem foi aplicado o BLUP para o modelo do caso univariado (equações 3 e 4). Essa abordagem é utilizada em diversos livros texto de modelos lineares de efeitos mistos para prever os efeitos aleatórios de dados de modelagem (por exemplo, McCULLOCH; SEARLE, 2001). Neste caso, o vetor observado b_1 inclui os volumes das árvores amostra e o vetor não observado b_2 inclui os efeitos aleatórios para a fazenda em questão. Os volumes observados das árvores amostra da fazenda k em primeira rotação são escritos no vetor $\ln v_k$ e as predições correspondentes que utilizam apenas a parte fixa do modelo (isto é, os diâmetros e alturas medidas) no vetor $\widehat{\mu}_k$.

O BLUP para os efeitos aleatórios se torna:

$$\widehat{\alpha}_k = 0 + DZ'_k(Z_k DZ'_k + R)^{-1}(\ln v_k - \widehat{\mu}_k), \quad (3.10)$$

e a variância da predição de $\widehat{\alpha}_k$ é

$$\text{var}(\widehat{\alpha}_k - \alpha_k) = D - DZ'_k(Z_k DZ'_k + R)^{-1}Z_k D \quad (3.11)$$

onde D , Z_k e R são matrizes relacionadas com o modelo univariado, tal como descrito na seção 3.3.3.1.

3.3.3.3 Calibração do modelo de primeira rotação com árvores-amostra de segunda rotação (Abordagem 2)

Na segunda abordagem, a pressuposição é a de que, ainda, apenas um modelo para a primeira rotação está disponível. A forma do BLUP é a mesma da primeira abordagem. No entanto, as árvores-amostra agora são selecionadas nos povoamentos de segunda rotação. Do ponto de vista da modelagem, um pressuposto implícito por trás desta segunda abordagem é de que as árvores da primeira e da segunda rotação são geradas pelo mesmo processo aleatório. Como foi dito na introdução, isto, aparentemente, não é verdadeiro. No entanto, esta abordagem busca avaliar o grau de imprecisão, comparada a decorrente economia de esforços de modelagem, quando se faz essa falsa suposição.

3.3.3.4 Calibração cruzada de ambos os modelos com árvores amostra de primeira rotação (Abordagem 3)

A terceira abordagem é uma aplicação do modelo bivariado (equações 3.5 e 3.6), (ver LAPPI et al., 2006, para mais detalhes). Esta abordagem assume que as árvores de diferentes rotações são geradas por processos diferentes. No entanto, os processos têm em comum o fato de que árvores das duas rotações terem crescido na mesma fazenda. Na predição dos efeitos aleatórios, essa semelhança é levada em conta através da utilização das correlações entre os modelos de primeira e segunda rotação.

O vetor de efeitos aleatórios α_k agora inclui os efeitos aleatórios da primeira e da segunda rotação, conforme definido no subitem 3.3.2. O BLUP de α_k usando os volumes observados das árvores-amostra de primeira rotação ($\ln v_k$) e as correspondentes predições da parte fixa μ_k , torna-se:

$$\widehat{\alpha}_k = 0 + CZ'_k(Z_k D_1 Z'_k + R_f)^{-1}(\ln v_k - \widehat{\mu}_k), \quad (3.12)$$

e a variância da predição de $\widehat{\alpha}_k$ é

$$\text{var}(\widehat{\alpha}_k - \alpha_k) = D - DZ'_k(Z_k DZ'_k + R_f)^{-1}Z_k D \quad (3.13)$$

Onde, D , C e D_1 são as matrizes definidas na eq. 3.7, Z_k é a matriz da parte aleatória do modelo de primeira rotação e R_f é a diagonal da matriz de variâncias residuais do modelo de primeira rotação para as árvores-amostra observadas.

3.4 Avaliação com dados empíricos.

Para avaliar as três abordagens de calibração, descritas anteriormente, foi implementado um procedimento de validação cruzada que foi aplicado ao conjunto de dados descrito anteriormente. Neste procedimento, dentre o conjunto de fazendas disponíveis, uma particular fazenda k é retirada por vez do conjunto de dados. A implementação consiste em três etapas: (i) Ajuste dos modelos das equações 3.2, 3.3 e 3.5 ao conjunto de dados, (ii) calibração do modelo usando um pequeno número de amostras de árvores selecionadas aleatoriamente da fazenda k , retirada do conjunto de dados e (iii) a predição da volume para as árvores da fazenda k que não foram utilizados no passo de calibração (ii) e avaliação das predições contra os volumes reais. A fim de cumprir o objetivo do presente trabalho, a aplicação das duas primeiras fases foi feita através das três diferentes abordagens de calibração apresentadas anteriormente. As próximas subseções apresentam essas fases de análise com mais detalhes.

3.4.1 Estimação

Com as árvores de todas as fazendas, exceto as da fazenda k retirada, um modelo linear de efeito fixo (modelo marginal) é ajustado por meio de mínimos quadrados ordinários. Através deste procedimento, as estimativas de $\beta_0, \beta_1, \beta_2$ para a eq. 3.2 são obtidas.

Com as mesmas árvores, o modelo linear de efeitos mistos foi ajustado utilizando

máxima verossimilhança restrita. Isso resulta nas estimativas dos parâmetros fixos (β_0 , β_1 e β_2), as variâncias e covariâncias dos efeitos aleatórios $\text{var}(\alpha_{0k}; \alpha_{2k})$, e na variância do resíduo (ϵ_{ki}) das equações 3.3. e 3.5.

3.4.2 Calibração

Na calibração, os efeitos aleatórios dos modelos de efeitos mistos foram preditos usando um pequeno número de árvores-amostra da fazenda k que não foi utilizada na estimação do modelo. As árvores da fazenda em questão foram classificadas em classes de tamanho de acordo com seu diâmetro, sendo amostradas uniformemente em todas as classes de diâmetro. Este procedimento foi repetido mil vezes, para evitar a influência dos erros de amostragem nos resultados finais. Os modelos se tornam calibrados, adicionando os valores preditos dos efeitos aleatórios às estimativas de parâmetros fixos da eq. 3.2.

3.4.3 Predição

Nesta fase, utilizaram-se os modelos marginais, a parte fixa dos modelos de efeitos mistos, e os modelos calibrados, para prever os volumes das árvores da fazenda k , que não foram utilizados na etapa de calibração. A fim de permitir a comparação entre os volumes preditos e volumes observados, os volumes preditos foram transformados de volta para a escala original, usando a abordagem de correção de viés proposta por (LAPPI, 2006):

$$b_{ei} = \frac{e^{[\ln v_{ki} + \sigma]} + e^{[\ln v_{ki} - \sigma]}}{2} \quad (3.14)$$

Onde σ é o desvio padrão do erro de predição de volume em escala logarítmica.

3.4.4 Implementação das abordagens de calibração

Devido a algumas limitações dos dados, as diferentes abordagens foram aplicadas em conjuntos de dados ligeiramente distintos. O conjunto de dados completo incluiu 16 fazendas com dados de primeira rotação e destas, apenas 8 fazendas também tinham dados de segunda

rotação. Portanto, avaliou-se a abordagem 1 com todas as 16 fazendas, enquanto que as abordagens 2 e 3 foram avaliadas apenas com as oito fazendas que possuíam dados de ambas as rotações. O número de árvores-amostra utilizadas na calibração variou de 1 a 20 no caso das abordagens 1 e 2, que utilizam árvores-amostra de primeira rotação. Na abordagem 3 o número máximo de árvores-amostra foi de 16, já que algumas das 8 fazendas utilizadas tinham apenas 17 árvores medidas.

3.4.5 Critérios de avaliação

A avaliação de cada abordagem foi feita através do erro quadrático médio (RMSE) e do viés, como segue:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (v_{ki} - \widehat{v}_{ki})^2}{n}} \quad (3.15)$$

$$VIÉS = \frac{\sum_{i=1}^n (v_{ki} - \widehat{v}_{ki})}{n} \quad (3.16)$$

O RMSE e o viés foram calculados usando todas as predições a fim de comparar a resposta em desempenho das três diferentes abordagens, com o aumento do número de árvores de amostra na calibração. Estes dois critérios foram também calculados na escala relativa, dividindo-os pelo valor do volume médio observado e multiplicando por 100.

3.5 Resultados

As estimativas dos parâmetros para os modelos marginais (eq. 3.2) e condicionais (eq. 3.3), ajustados ao conjunto de dados descrito na seção 3.2 são apresentadas nas Tabelas 3.3, 3.4 e 3.5. Mesmo que no procedimento de validação cruzada uma das fazendas sempre seja retirada, decidiu-se reportar aqui os modelos ajustados com os dados completos visando potencial uso futuro. Assim, as estimativas da Tabela 3.3 são baseadas em todas as 16 fazendas da primeira rotação e as estimativas da Tabela 3.4 são baseadas nas 8 fazendas coincidentes que tinham árvores em ambas as rotações. A Tabela 3.5 relata as correlações e

covariâncias entre os dois modelos aparentemente não relacionados reportados na Tabela 3.4.

Tabela 3.3 – Estimativas dos parâmetros do modelo marginal e do modelo condicional ajustados às 16 fazendas de primeira rotação.

Parâmetros	Estimativas	
	Modelo marginal	Modelo condicional
$\bar{\beta}_0$	-3,1774 (0,01632)	-3,0919 (0,04068)
$\bar{\beta}_1$	1,8283 (0,00984)	1,8325 (0,01019)
$\bar{\beta}_2$	1,0954 (0,01172)	1,0588 (0,01988)
$d.p.(\alpha_{0k})$	-	0,13503
$d.p.(\alpha_{2k})$	-	0,05609
$cor(\alpha_{0k}, \alpha_{2k})$	-	-0,991
$d.p.(\epsilon_{ki})$	0,06328	0,05857

Todas as hipóteses consideradas para estes modelos foram cumpridas e, por isso, estes são adequados para fins de aplicação propostos neste trabalho. A correlação entre os parâmetros aleatórios dos modelos condicionais de primeira e segunda rotação apresentada na Tabela 3.5, por volta de 0,5, confirma o potencial de utilização da calibração cruzada proposta na terceira abordagem.

Tabela 3.4 – Estimativas dos parâmetros do modelo marginal e do modelo condicional ajustados às oito fazendas de primeira e segunda rotação.

Parâmetros	Modelo marginal		Modelo condicional	
	Primeira rotação	Segunda rotação	Primeira rotação	Segunda rotação
$\bar{\beta}_0$	-3,1374 (0,0209)	-3,0043 (0,0426)	-3,1212 (0,0411)	-3,2376 (0,2909)
$\bar{\beta}_1$	1,8467 (0,0122)	1,7989 (0,0343)	1,8332 (0,0123)	1,8183 (0,0272)
$\bar{\beta}_2$	1,0691 (0,0152)	1,0545 (0,03644)	1,0740 (0,0205)	1,1267 (0,1145)
<i>d.p.</i> (α_{0k})	-	-	0,08516	0,81140
<i>d.p.</i> (α_{2k})	-	-	0,03283	0,30928
<i>cor</i> (α_{0k}, α_{2k})	-	-	-0,876	-0,978
<i>d.p.</i> (ϵ_{ki})	-	-	0,05671	0,08592

Tabela 3.5 – Estimativas de desvio padrão, correlações e variâncias e covariâncias do modelo aparentemente correlacionado ajustado às oito fazendas de primeira e segunda rotação.

	Correlações e Desvio padrão ¹				Variâncias e Covariâncias			
	Primeira rotação		Segunda rotação		Primeira rotação		Segunda rotação	
	α_{f0k}	α_{f2k}	α_{s0k}	α_{s2k}	α_{f0k}	α_{f2k}	α_{s0k}	α_{s2k}
α_{f0k}	0,086				0,0074			
α_{f2k}	-0,977	0,028			-0,0024	0,0008		
α_{s0k}	0,484	-0,292	0,823		0,0343	-0,0068	0,6787	
α_{s2k}	-0,469	0,275	-0,999	0,315	-0,0127	0,0024	-0,2595	0,099

1 O desvio padrão é mostrado na diagonal da matriz de correlações

Primeira abordagem

A parte fixa do modelo de efeitos mistos apresentou predições de volume menos precisas que o modelo marginal (Figura 3.1 A). O RMSE relativo das predições feitas com a

parte fixa do modelo de efeitos mistos foi de 9,3% enquanto que as feitas com o modelo marginal geraram um RMSE relativo de 8,4%. O valor de RMSE do modelo calibrado já é menor que o valor de RMSE do modelo marginal após 4 árvores amostradas. A redução do RMSE parece estabilizar, ao nível de 7,6%, após 17 ou 18 árvores amostradas.

O procedimento de calibração também foi eficaz para reduzir o viés (Figura 3.1 B). Depois de nove árvores amostradas, o modelo calibrado tornou-se menos enviesado do que o modelo marginal. O viés não foi completamente removido, no entanto, o viés remanescente, de cerca de 1% depois de 20 árvores, é bastante aceitável. Bem como observado no gráfico do RMSE, o viés da parte fixa do modelo condicional é maior do que o viés do modelo marginal, mostrando também uma menor precisão nas predições.

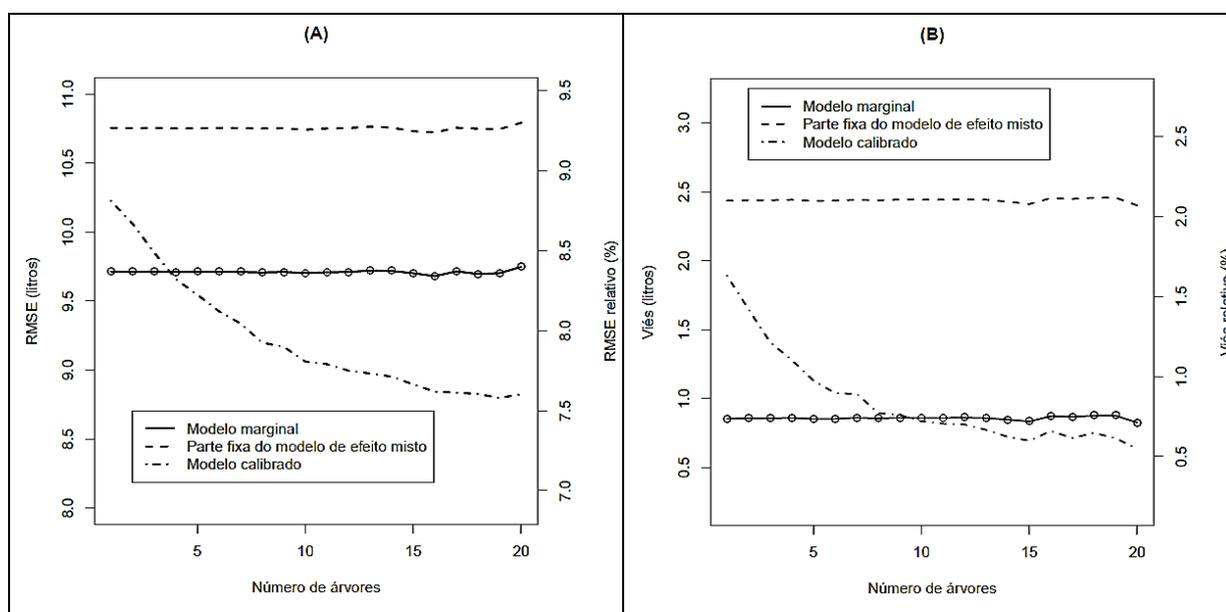


Figura 3.1 – Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE (A) e do Viés (B) na primeira abordagem.

A Figura 3.2 mostra os gráficos do RMSE para duas fazendas diferentes usadas na primeira abordagem. Na fazenda da Figura 3.2 A, pode-se verificar tendências semelhantes como as observadas na (Figura 3.1 A): a calibração foi muito eficaz na redução dos valores de RMSE e a diferença relativa entre os modelos calibrado e marginal apresentam magnitude semelhante. No entanto, os valores de RMSE de todos os três modelos nesta fazenda são muito elevados quando comparados com o resultado geral. Por outro lado, na fazenda da figura 3.2 B, mesmos os modelos sem calibração já se mostraram bastante precisos e o procedimento de calibração não melhorou essa precisão. O comportamento das curvas é muito diferente do resultado geral (Figura 3.1 A). Além disso, a diferença entre os três

métodos, neste segundo exemplo, é muito pequeno, tanto em escala absoluta quanto em escala relativa.

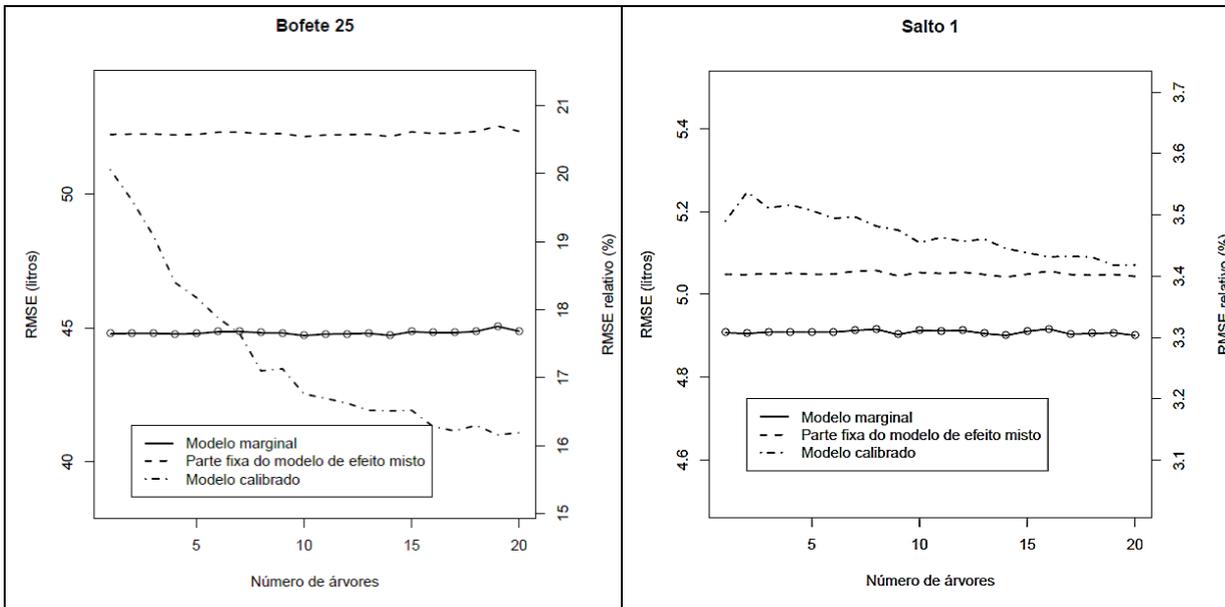


Figura 3.2 - Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE em duas fazendas usadas na primeira abordagem.

Segunda abordagem

A Figura 3.3 A mostra o gráfico do RMSE dos modelos ajustados na segunda abordagem. Neste gráfico, o modelo marginal se refere a um modelo ajustado aos dados de segunda rotação e pode ser visto como um limite superior para a precisão das predições da segunda abordagem. Semelhante à primeira abordagem, o RMSE do modelo calibrado tem uma tendência de redução com o aumento no número de árvores-amostra utilizadas no procedimento de calibração. A precisão se aproxima da precisão do modelo marginal de segunda rotação com o aumento do tamanho da amostra de calibração.

A Figura 3.3 B mostra o gráfico do viés dos modelos ajustados na segunda abordagem. O procedimento de calibração também foi eficaz para reduzir o viés. O comportamento do viés começa como uma ligeira tendência de superestimativa, que se altera para uma tendência de subestimativa, mas sempre com valores relativos muito baixos. Neste caso, tanto o modelo calibrado quanto a parte fixa do modelo de efeitos mistos foram sempre geraram predições menos enviesadas que o modelo marginal. Isto ocorreu porque o ganho do procedimento de calibração se inicia a partir da parte fixa do modelo de efeitos mistos e este já era menos enviesado que o modelo marginal. A inversão no comportamento do viés em comparação com a primeira abordagem pode ser explicada pelo fato de que o modelo

marginal e o modelo condicional foram ajustados, por esta abordagem, a partir conjuntos de dados diferentes.

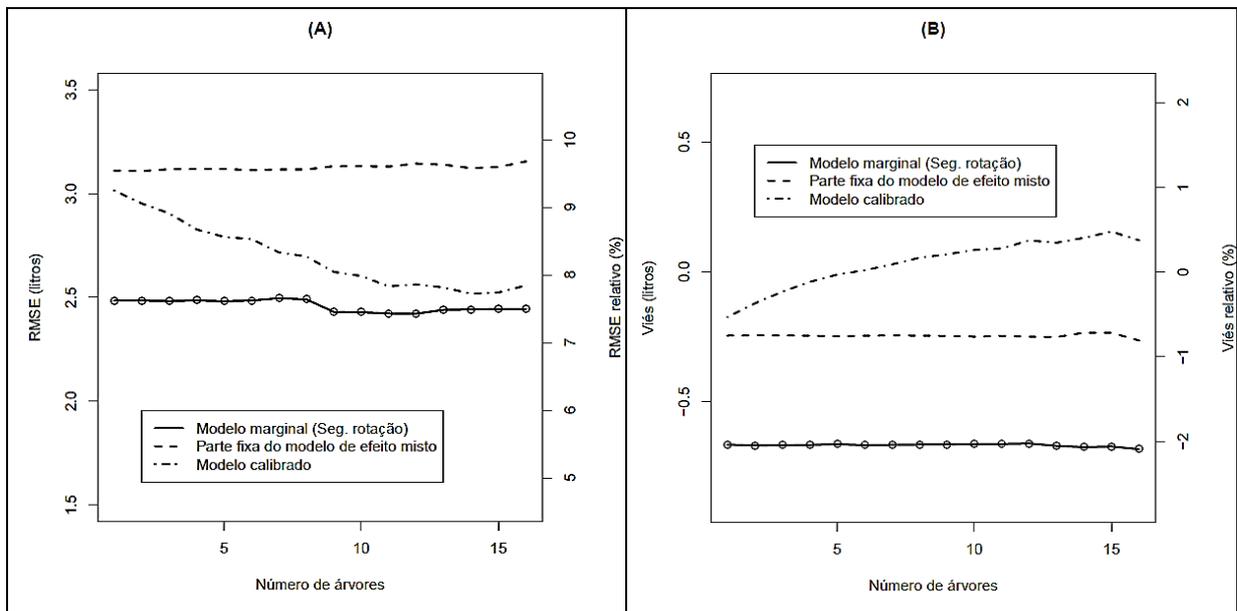


Figura 3.3 - Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE (A) e do Viés (B) na segunda abordagem.

As Figuras 3.4 A e 3.4 B trazem os gráficos do RMSE para duas diferentes fazendas utilizadas na segunda abordagem. Em 3.4 A, a calibração foi muito eficaz, reduzindo os valores do RMSE pela metade: de 20% para 10%. Mais uma vez, os valores do RMSE para todos os três modelos aplicados nesta fazenda são bastante elevados em comparação com o resultado geral. Em contrapartida, a fazenda apresentada na figura 4B apresentou um comportamento muito diferente do resultado geral. A calibração fez com que as previsões fossem piores quando comparadas às previsões geradas pela parte fixa do modelo de efeitos mistos e com o modelo marginal. No entanto, os valores absolutos do RMSE de todos esses modelos são muito baixos, o que pode explicar este comportamento peculiar.

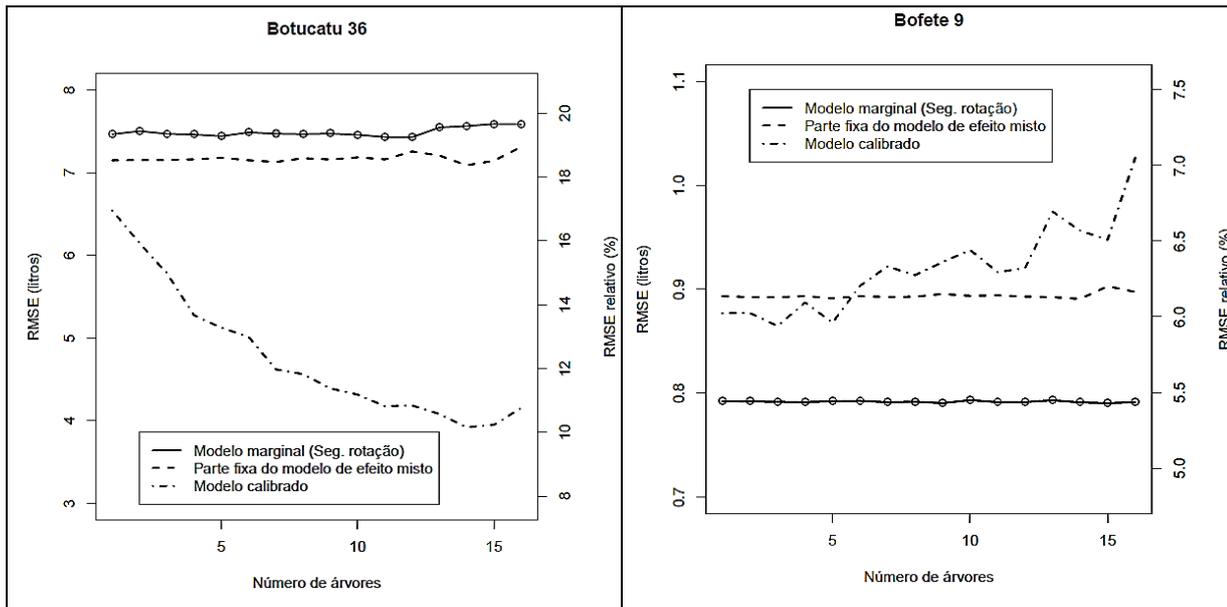


Figura 3.4 - Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE em duas fazendas usadas na segunda abordagem.

Terceira abordagem

Também a terceira abordagem apresentou uma tendência de redução do RMSE do modelo calibrado com o aumento do número de árvores- amostra (Figura 5). Esta tendência decrescente não mostrou qualquer sinal de estabilização até o número máximo de 20 árvores amostradas. A calibração cruzada reduziu o RMSE relativo de 11,6% para 10,8%, em comparação com as previsões marginais.

A Figura 3.5 B mostra o gráfico do viés dos modelos ajustados na terceira abordagem. O procedimento de calibração também foi eficaz para reduzir o viés, restando um viés relativo, de cerca de 4,8%. Este foi maior do que o viés observado na segunda abordagem. A parte fixa do modelo condicional teve um comportamento muito próximo ao modelo marginal em termos de RMSE e viés, mas ainda foi um pouco menos preciso do que o modelo marginal.

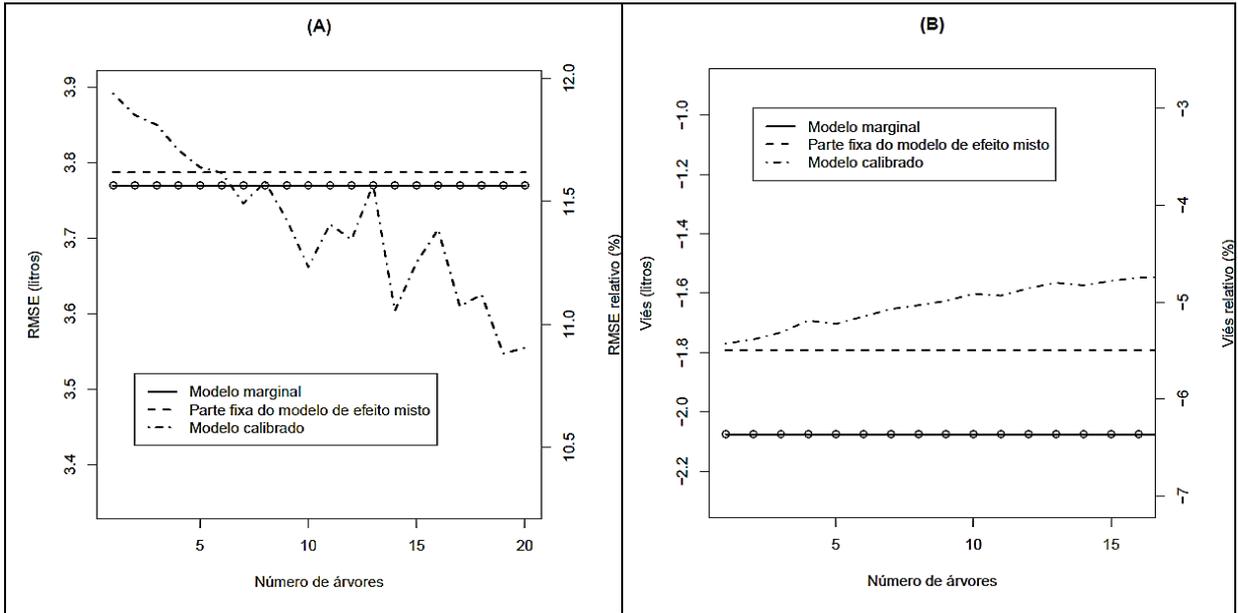


Figura 3.5 - Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE (A) e do Viés (B) na terceira abordagem.

Para completar, a Figura 3.6 mostra os gráficos do RMSE para duas fazendas usadas na terceira abordagem, sendo uma primeira fazenda com RMSE altos e uma segunda fazenda onde os modelos sem calibração já eram bastante precisos. Mais uma vez, a calibração foi altamente eficiente para a fazenda com RMSE inicial alto, e inútil na fazenda onde as previsões iniciais já eram bastante precisas.

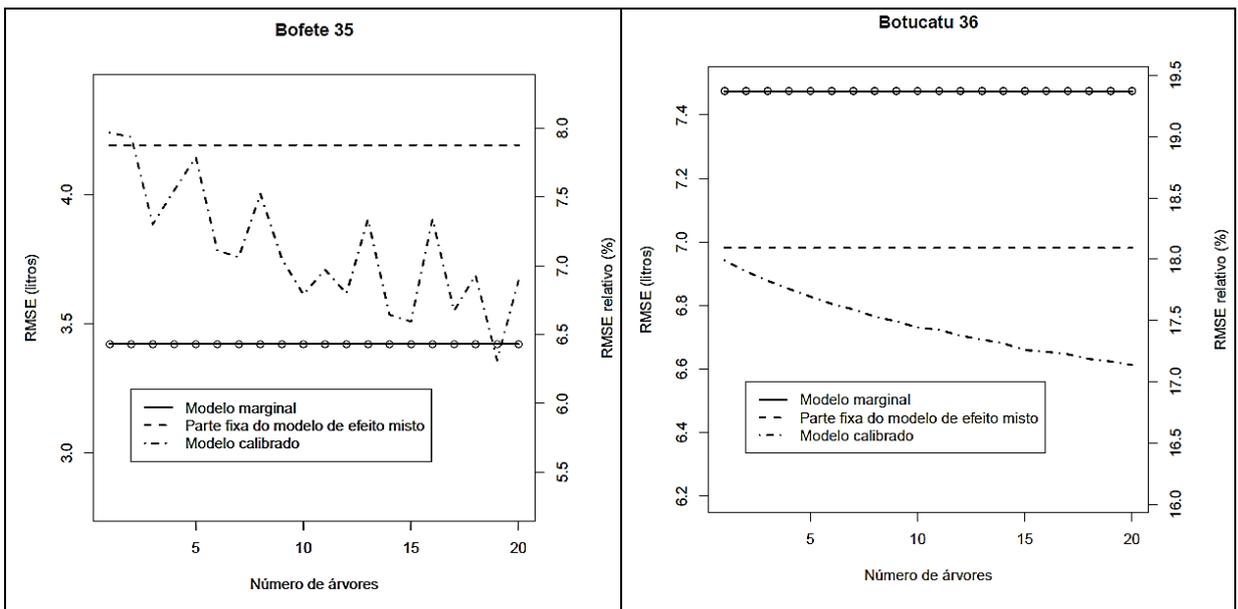


Figura 3.6 - Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE em duas fazendas usadas na terceira abordagem.

3.6 Discussão e Conclusões

Este trabalho apresentou três abordagens para calibração de modelos de efeitos mistos visando melhorar a qualidade das predições de volume em plantios de *Eucalyptus grandis*.

A primeira abordagem apresentou o procedimento de calibração do modelo em uma situação em que o ajuste, a predição e as amostras de calibração pertenciam à mesma população estatística. Para isso, foram utilizadas as florestas de *Eucalyptus grandis* em primeira rotação e o objetivo foi demonstrar a melhora das predições de volume por meio da aplicação convencional do procedimento de calibração (por exemplo, LAPPI; BAILEY, 1988). No entanto, essa abordagem traz um aspecto novo, já que foi a primeira vez em que esse método de calibração convencional foi aplicado à modelos de predição de volume.

Na segunda e terceira abordagens procurou-se lidar com a questão da melhora da precisão do modelo a partir de esforços limitados de coleta de dados na segunda rotação. Esta questão também foi tratada usando o procedimento de calibração, mas através de duas abordagens distintas. Na segunda abordagem, procurou-se determinar se um pequeno número de árvores-amostra em segunda rotação seria capaz de calibrar o modelo da primeira rotação, evitando-se assim esforços adicionais de coleta extensiva de dados de segunda rotação na construção de um novo modelo. A terceira abordagem, por outro lado, tentou reduzir os esforços através da utilização de um modelo bivariado constituído pelos modelos de primeira e segunda rotação. Nesse caso, assumiu-se que os dados de construção do modelo da segunda rotação estão disponíveis, mas a calibração é baseada unicamente nas árvores anteriormente derrubadas na primeira rotação. A calibração influenciou as predições da segunda rotação por meio da correlação dos efeitos aleatórios entre os dois modelos.

As três abordagens revelaram-se eficazes em melhorar a precisão da predição de volume. Os processos de calibração sempre reduziram os valores de RMSE e do viés, fazendo do modelo calibrado melhor que o correspondente modelo marginal. O modelo calibrado, mesmo com os menores tamanhos de amostra de calibração foi sempre mais preciso do que a parte fixa do modelo de efeitos mistos, Com amostras de menor tamanho, no entanto, o modelo calibrado foi menos preciso do que o modelo marginal.

Além de comparar os diferentes resultados, é necessário também considerar as necessidades de dados nas diferentes abordagens. Por exemplo, para calibrar o modelo da primeira rotação para as 8 fazendas da segunda rotação e utilizando 14 árvores por fazenda é necessário apenas 112 árvores, ao passo que os modelos marginais utilizados como

referência foram baseados em 443 árvores derrubadas em segunda rotação. Isso demonstra que o uso de modelos de efeitos mistos e da calibração podem reduzir significativamente os esforços de medição nas florestas de segunda rotação.

Considerando a terceira abordagem, a aplicação de um sistema de modelos de efeitos mistos, na fase de estimação mostrou resultados muito satisfatórios. Apesar da correlação entre os parâmetros aleatórios dos dois modelos ter sido moderada, o modelo calibrado foi muito eficaz na redução dos valores de RMSE em relação ao modelo marginal. Neste caso, apenas com uma árvore de primeira rotação amostrada, as previsões do modelo calibrado para a segunda rotação foram mais precisas do que as previsões do modelo marginal ajustado a partir do conjunto de dados completo de segunda rotação.

A calibração cruzada de um sistema de modelo de efeito misto aparentemente não relacionados (abordagem 3), a qual já tinha demonstrado bons resultados em Lappi (1991) e Maltamo et al. (2012), se mostrou consideravelmente promissora também neste estudo, ainda que a redução do viés tenha sido ligeiramente menor do que com as abordagens 1 e 2. Além disso, esta abordagem pode ser ainda mais útil em situações em que a correlação entre os parâmetros aleatórios seja mais elevada do que os valores observados neste estudo. Também é possível aplicar o método com árvores-amostra de ambas as rotações. Por exemplo: pode-se utilizar as árvores de primeira rotação disponíveis e possivelmente medir algumas árvores de segunda rotação se a precisão ainda não parecer satisfatória.

Os gráficos de RMSE das fazendas individuais nas três abordagens (Figuras 3.2, 3.4 e 3.6) mostram que, embora o procedimento de calibração tenha sido eficaz no âmbito geral, não foi eficaz para todas as fazendas. Especialmente nos casos em que o modelo marginal mostrou bons resultados, representado pelo baixo valor do RMSE, a abordagem de modelos de efeito misto e calibração não são eficazes. Além disso, a diferença de desempenho do modelo marginal e da parte fixa do modelo de efeitos mistos nestas fazendas também foi relativamente pequena. Este resultado lógico indica que a calibração pode não ser necessária em todas as fazendas, já que, fazendas onde os efeitos aleatórios têm um valor próximo de zero não se beneficiam de calibração. É claro que a possível utilidade da calibração deve ser conhecida antes de se coletar as árvores-amostra, a fim de evitar esforços desnecessários de inventário. Um procedimento prático poderia envolver a retirada de uma pequena amostra para a estimativa inicial provisória dos efeitos aleatórios, seguidos de uma campanha de medição mais intensa se a calibração se mostrar promissora.

Uma característica marcante dos resultados foi que as previsões marginais foram mais precisas do que as previsões baseadas na parte fixa do modelo de efeitos mistos. Um

comportamento semelhante foi observado também em outros estudos (LAPPI; MALINEN, 1994; DE MIGUEL et al., 2012). De acordo com Lappi e Malinen (1994) este fenômeno pode estar relacionado com o fato de que os parâmetros específicos de povoamento estão correlacionados com as médias específicas de grupo (fazendas) dos preditores do modelo. Neste caso, o modelo marginal não explica apenas a dependência dos preditores e da variável resposta dentro da mesma fazenda, mas também o efeito da variação entre as fazendas. Neste estudo, se observou correlações em torno de 0,4 entre os parâmetros aleatórios específicos de fazenda e as médias, específicas de fazenda, da altura total e do DAP. Isto pode, pelo menos parcialmente, explicar o comportamento observado das predições. No entanto, as predições do modelo marginal podem ser enviesadas em situações distintas e onde a distribuição dos preditores seja diferente da dos dados de ajuste dos modelos.

Referências

- CALAMA, R.; MONTERO, G. Interregional nonlinear height-diameter model with random coefficients for Scotch pine in Spain. **Canadian Journal of Forest Research**, Vancouver, v. 34, n. 1, p. 150-163, 2004.
- CALEGARIO, N.; DANIELS, R.F.; MAESTRI, R.; NEIVA, R. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal *Eucalyptus* plantation case study. **Forest Ecology and Management**, Winsconsin, n. 204, p. 11-20, 2005.
- DE MIGUEL, S.; MEHTÄTALO, L.; SHATER, Z.; KRAID, B.; PUKKALA, T. Evaluating marginal and conditional predictions of taper models in the absence of calibration data. **Canadian Journal of Forest Research**, Vancouver, v. 42, n. 7, p. 1383-1394, 2012.
- DIAZ-BALTEIRO, L.; RODRIGUES, L.C.E. Optimal rotations on Eucalyptus plantations including carbon sequestration: a comparison of results in Brazil and Spain. **Forest Ecology and Management**, Winsconsin, v. 229, n. 1/3, p. 247-258, July 2006.
- DORADO, F.C.; DIEGUEZ-ARANDA, U.; ANTA, M.B.; RODRIGUES, M.S.; VON GADOW, K. A generalized height-diameter model including random components for radiata pine in north-western Spain. **Forest Ecology and Management**, Winsconsin, v. 229, n. 1-3, p. 202-213, July 2006.
- EERIKÄINEN, K. Stem volume models with random coefficients for *Pinus kesiya* in Tanzania, Zambia, and Zimbabwe. **Canadian Journal of Forest Research**, Vancouver, v. 31, n. 5, p. 879-888, 2001.

FANG, Z.; BAILEY, R.L.; SHIVER, B.D. A multivariate simultaneous prediction system for stand growth and yield with mixed and random effects. **Forest Science**, Bethesda, v. 47, n. 4, p. 550-560, 2001.

FIEUWS, S., VERBEKE, G. Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. **Biometrics**, London, v. 62, n. 7, p. 424-431, 2006.

GOLDSTEIN, H. **Multilevel statistical models**. 4th ed. Chichester: Wiley, 2011. 358 p.

HUSCH, B.; BEERS, T.W.; KERSHAW, J.A. **Forest mensuration**. 4th ed. New York: Wiley, 2003. 443 p.

LAPPI, J. **Mixed linear models for analyzing and predicting stem form variation of scots pine**. Milano: Instituti Forestalis Fenniae, 1986. 69 p. (Communicationes, 134).

_____. Calibration of height and volume equations with random parameters. **Forest Science**, Bethesda, v. 37, p. 781-801, 1991.

_____. A longitudinal analysis of h/d curves. **Forest Science**, Bethesda, v. 43, p. 555-570, 1997.

_____. A multivariate, nonparametric stem-curve prediction method. **Canadian Journal of Forest Research**, Vancouver, v. 36, p. 1017-1027, Apr. 2006.

LAPPI, J.; BAILEY, R. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. **Forest Science**, Bethesda, v. 34, p. 907-927, 1988.

LAPPI, J.; MALINEN, J. Random-parameter height/age models when stand parameters and stand age are correlated. **Forest Science**, Bethesda, v. 40, n. 4, p. 715-731, Nov. 1994.

LAPPI, J.; MEHTÄTALO, L.; KORHONEN, K.T. Generalizing sample tree information. In: KANGAS, A.; MALTAMO, M. (Ed.). **Forest inventory: methodology and application**. Helsinki: Springer, 2006. p. 85-106.

MALTAMO, M.; MEHTÄTALO, L.; JARI, V.; PACKALEN, P. Predicting and calibrating tree attributes by means of airborne laser scanning and field measurements. **Canadian Journal of Forest Research**, Vancouver, 2012. In press.

McCULLOCH, C.E.; SEARLE, S.R. **Generalized, linear, and mixed models**. New York: Wiley-Interscience, 2001. 325 p.

MEHTÄTALO, L. Height-diameter models for scots pine and birch in Finland. **Silva Fennica**, Helsinki, v. 39, n. 1, p.55-66, 2005a.

_____. Localizing a predicted diameter distribution using sample information. **Forest Science**, Bethesda, v. 51, n. 4, p. 292-302, 2005b.

MEHTÄTALO, L.; COMAS, C.; PUKKALA, T.; PALAHÍ, M. Improving predicted diameter distribution by using a small sample of diameters. In press.

PESONEN, A.; EERIKÄINEN, K.; MALTAMO, M.; TAHVANAINEN, T. Models for predicting tree and stand development on larch plantations in Hallormsstaour, Iceland. **New Forests**, Montreal, v. 37, p. 63-83, 2009.

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

SCHUMACHER, F.X.; HALL, F.S. Logarithmic expression of timber-tree volume. **Journal of Agricultural Research**, Boston, v. 47, p. 719-734, 1933.

WANG, M.; BHATTI, J.; WANG, Y.; VAREM-SANDERS, T. Examining the gain in model prediction accuracy using serial autocorrelation for dominant height prediction. **Forest Science**, Bethesda, v. 57, n. 3, p. 241-251, 2011.

4 APLICAÇÃO DE MODELOS LINEARES DE EFEITO MISTO E CALIBRAÇÃO NA PREDIÇÃO DA BIOMASSA EM FLORESTAS TROPICAIS MULTIESPECÍFICAS

Resumo

Este trabalho apresenta aplicações da calibração do modelo linear de efeito misto na predição da biomassa de árvores de espécies nativas amostradas numa floresta Ombrófila densa localizada no município de Registro, SP. Partiu-se do modelo de potência, em sua forma linearizada, para o desenvolvimento modelo de efeitos mistos. Para isso, alguns dos parâmetros deste modelo foram considerados como sendo aleatórios ao longo das diferentes parcelas e ao longo das diferentes espécies. O ajuste do modelo foi feito considerando esses dois níveis. Já a calibração foi realizada em nível de espécie, ignorando o efeito de parcela e em nível de parcela, ignorando o efeito de espécie. Para tanto, partiu-se de um pequeno número de árvores-amostra que não foram usadas no ajuste do modelo. Os resultados mostram que o procedimento de calibração fornece predições mais confiáveis em nível de espécie do que os modelos tradicionais. Em nível de parcela, a calibração não foi efetiva.

Palavras-chave: Modelos de biomassa; Modelos de efeitos mistos; Calibração

Abstract

This paper presents applications of linear mixed-effects model and calibration in predicting tree biomass of native species collected in a rain forest located in Registro, SP. We started from the power function, in its linearized form, for developing the mixed-effects model. Thus, some of the parameters of this model were assumed to be random among different plots and species. The model fitting was done considering these two levels while the calibration was performed at the species level, ignoring the effect of plot and plot-level, ignoring the species effects. To this end, were used a small number of sample trees that were not used in model fitting stage. The results show that the calibration procedure of mixed effects models provides more reliable predictions at species level than the traditional models. Regarding to the plot-level, the calibration was not effective.

Keywords: Biomass models; Mixed-effects models; Calibration

4.1 Introdução

No processo de aplicação de modelos de predição de biomassa uma seleção criteriosa e cuidadosa pode gerar predições de biomassa adequadas para situação em estudo e para tomadas de decisão em relação à biomassa. Entretanto, a aplicação de modelos inadequados pode resultar em predições totalmente equivocadas, gerando uma representação falsa da

situação estudada. Os equívocos podem ser tanto a superestimativa quanto a subestimativa, dependendo do modelo aplicado e da situação de estudos (CLARK; CLARK, 2000).

Os procedimentos de seleção de modelos deveriam, sempre que possível, incluir a comparação das predições com dados empíricos de biomassa específicos da situação em estudos, ainda que para uma amostra pequena, uma vez que as informações obtidas na literatura são raramente suficientes para que se possa antever qual dos modelos publicados resultará em predições adequadas ou inadequadas na situação de estudo.

As equações recomendadas pelo “Intergovernmental Panel on Climate Change” (IPCC) para utilização na floresta tropical (BROWN, 1997; CHAVE, 2005) sugerem certa generalidade quanto à sua aplicação, já que, foram ajustados a partir de dados coletados de diversos biomas localizados nos trópicos. No entanto, como esses modelos são baseados em dados empíricos, sua utilização é sítio específico e pode levar a erros de predição importantes (ARAÚJO et al., 1999). Uma possível alternativa para contornar esse problema seria construir modelos específicos para o local onde se quer aplicá-los. No entanto, existem dificuldades legais no sentido da liberação ambiental para obtenção de grandes amostras destrutivas, necessárias para construção de modelos de boa precisão.

Um agravante da aplicação desses modelos empíricos em florestas tropicais reside no fato de que os modelos são construídos a partir de uma amostra de determinadas espécies e no processo de inventário florestal este é aplicado em parcelas onde muitas vezes a composição de espécies é muito distinta da amostra de ajuste do modelo

Esses problemas poderiam ser minimizados a partir da construção de um modelo que pudesse ser calibrado às diferentes situações, lançando mão de um pequeno número de amostras destrutivas. Para tanto, é preciso construir um modelo de predição a partir da perspectiva de modelos de efeitos mistos (LAPPI, 1991, 1997; CALAMA; MONTEIRO, 2004; CASTEDO-DORADO et al., 2006).

Essa abordagem se caracteriza por tomar alguns dos parâmetros do modelo de predição como sendo fixos e outros como sendo aleatórios ao longo de determinado grupo pré-definido. O modelo pode, então, ser calibrado para um determinado grupo a fim melhorar a qualidade das predições dentro do grupo em questão. No contexto da Floresta Tropical, esses grupos podem ser, por exemplo, diferentes formações florestais, diferentes parcelas, as diferentes espécies, etc,

No meio florestal, Biging (1985), usou a abordagem de parâmetros aleatórios na perspectiva de curvas de índice de sítio e Lappi e Bailey (1988) e Temesgen e von Gadow (2004), utilizaram essa abordagem na predição de altura em níveis de árvore e de

povoamento. Já Lappi (1986) e Eerikainen (2001), por sua vez, usaram os parâmetros aleatórios numa perspectiva de modelos de predição de volume e funções de afilamento. Entre outras aplicações destacam-se: os modelos de altura em função do diâmetro (LAPPI, 1997; CALAMA; MONTERO 2004; MEHTÄTALO, 2005a, DORADO et al., 2006), modelos de altura dominante (LAPPI; BAILEY, 1988; CALEGARIO et al., 2005; PESONEN et al., 2009) e modelos de distribuição de diâmetro (MEHTÄTALO, 2005, 2010).

Modelos de efeitos mistos, onde se consideram diversos grupos simultaneamente, foram utilizados numa perspectiva de modelo não lineares para predizer crescimento (HALL; BAILEY, 2001; HALL; CLUTTER, 2004; ZHAO et al., 2005; NOTHDURFT; KUBLI; LAPPI, 2006; YANG et al., 2009), biomassa e volume de diferentes compartimentos (ZANG; BORDERS, 2004; SUNANDA; JAYARAMAN, 2006; EERIKAINEN, 2009). Em florestas tropicais multiespecíficas, no entanto, a abordagem é ainda pouco explorada.

A hipótese deste trabalho é de que em florestas inequiâneas multiespecíficas, onde a composição de espécies muda de acordo com a unidade amostral, torna-se interessante a aplicação da abordagem de efeitos aleatórios de múltiplos níveis. Neste caso, os níveis parcela e espécie poderiam permitir a obtenção de um modelo calibrável, a uma determinada espécie ou a uma determinada parcela que não foram incluídas na amostra de ajuste.

Diante do que foi exposto acima, o presente trabalho tem como objetivo construir uma equação de predição de biomassa arbórea a partir de dados coletados numa floresta tropical localizada no município de Registro, SP e calibra-las à uma situação específica de aplicação usando uma pequena amostra destrutiva representativa dessa situação. Para isso propõe-se o uso da metodologia de modelos de parâmetros aleatórios de múltiplos níveis e calibração proposta por Lappi (1991) e usada em Lappi (2001) e Erikäinen (2009).

Pretende-se, assim, dar uma contribuição à metodologia recomendada pelo IPCC no que se refere a projetos de sequestro de carbono, através da redução do viés de predição das equações gerais aplicadas a situações locais.

4.2 Material

4.2.1 Área de estudo e coleta de dados

Os dados foram coletados numa área de supressão florestal localizada às margens do aeroporto de Registro, SP. A área do aeroporto de Registro encontra-se inserida no bioma mata Atlântica e com frequência a vegetação que ocorre sobre as planícies litorâneas tem sido chamada de vegetação de restinga, ou incluída dentro do complexo da restinga.

Cabe, no entanto, ressaltar que, segundo os critérios estabelecidos pelo sistema de classificação fisionômico ecológico do IBGE (VELOSO et al., 1991), a vegetação contida na área de estudo recebe a denominação de Floresta Ombrófila Densa das Terras Baixas.

As árvores foram amostradas na área de forma a se obter duas amostras com características distintas. Uma amostra composta de unidades amostrais superficiais (parcelas) e outra amostra composta por árvores individuais isoladas. As características das duas amostras são discutidas a seguir:

1) Parcelas

Foram locadas, aleatoriamente, 10 unidades amostrais de 400 m² em três estratos diferentes. Estes estratos foram estabelecidos pelo inventário florestal pré-realizado e estão relacionados aos estados de regeneração da comunidade florestal, classificados como: inicial, médio e avançado.

Cada estrato apresenta uma composição de espécies bastante diferente e as parcelas foram distribuídas nestes estratos, propositalmente, a fim captar essa variação. O número de árvores coletadas por parcela é apresentada na Tabela 4.1.

No entanto, na tentativa de aumentar o tamanho da amostra de ajuste, as árvores isoladas (Tabela 4.2) foram consideradas como sendo uma parcela de uma única árvore. Fazendo essa consideração tem-se ao todo 90 parcelas com todas as 219 árvores coletadas,

Tabela 4.1 – Distribuição das árvores nas parcelas que foram usadas nos procedimentos de calibração e predição.

Parcela	Número de árvores
1	21
2	11
3	23
4	15
5	17
6	13
7	12
8	11
9	7
10	9
Total	139

2) Árvores isoladas

A amostra foi composta das espécies mais frequentes na área e segundo dois critérios:

i) Dentre as espécies mais frequentes foram escolhidas 10 espécies que apresentavam maior distinção quanto à sua densidade básica a fim, a fim de conferir maior representatividade ao modelo de predição.

ii) Dentro de cada espécie, as árvores foram escolhidas seguindo um processo de amostragem estratificada, sendo cada classe de diâmetro definida como um extrato. Foram estabelecidas quatro classes de diâmetro cuja amplitude em cada classe variou de acordo a distribuição diamétrica observada da espécie. A amostragem foi uniforme nos extratos, ou seja, cada classe de diâmetro foi contemplada com dois indivíduos amostrados. O número de árvores coletadas por espécie é apresentado na Tabela 4.2.

Tabela 4.2 – Distribuição das árvores amostradas por espécie na parcela e isoladamente e usadas para o ajuste do modelo. O símbolo * indica as espécies que foram utilizadas na calibração e predição do modelo.

Espécie	Amostra			Espécie	Amostra		
	Isolada	Parcela	Total		Isolada	Parcela	Total
<i>Endlicheria paniculata</i> *	8	16	24	<i>Guarea macrophylla</i>	-	3	3
<i>Calophyllum brasiliense</i> *	8	13	21	<i>Leandra australis</i>	-	3	3
<i>Pera glabrata</i> *	8	12	20	<i>Miconia cinerascens</i>	-	3	3
<i>Myrcia racemosa</i> *	8	11	19	<i>Myrcia pubipetala</i>	-	3	3
<i>Alchornea triplinervia</i> *	8	9	17	<i>Campomanesia phaea</i>	-	2	2
<i>Tibouchina pulchra</i> *	8	7	15	<i>Cyathea sp.</i>	-	2	2
<i>Andira fraxinifolia</i> *	8	4	12	<i>Rapanea ferruginea</i>	-	2	2
<i>Gomidesia affinis</i> *	8	4	12	<i>Ficus sp.</i>	-	1	1
<i>Jacaranda puberula</i> *	8	2	10	<i>Gomidesia fenzliana</i>	-	1	1
<i>Psidium cattleianum</i> *	8	2	10	<i>Miconia sellowiana</i>	-	1	1
<i>Coussapoua microcarpa</i> *	-	6	6	<i>Nectandra oppositifolia,</i>	-	1	1
<i>Ilex theezans</i> *	-	6	6	<i>Ocotea brachybotra</i>	-	1	1
<i>Eugenia ceraciflora</i> *	-	5	5	<i>Posoqueria acutifolia</i>	-	1	1
<i>Andira fraxinifolia</i> *	-	4	4	<i>Rollinia sericea</i>	-	1	1
<i>Ocotea sp.</i> *	-	4	4	<i>Ternstroemia brasiliensis</i>	-	1	1
<i>Tapirira guianensis</i> *	-	4	4	<i>Tibouchina trichopoda</i>	-	1	1
Total de árvores isoladas							80
Total de árvores do grupo parcela							139
Total de árvores do grupo espécie*							189
Total Geral							219

4.2.2 Medição das variáveis dendrométricas

Com as árvores ainda em pé foram medidos o diâmetro com casca a 1,30cm do solo (dap), ou do conjunto de fustes caso a árvore fosse perfilhada, além da altura total (ht). Para o caso da árvore perfilhada, foi calculado o diâmetro equivalente da árvore através da soma das áreas seccionais dos diferentes fustes.

4.2.3 Medição da biomassa

A biomassa lenhosa foi determinada para os diferentes compartimentos das árvores e esses compartimentos foram definidos de acordo com a literatura (ARAUJO et al., 1996; HIGUCHI et al., 1998; SANQUETTA, 2002; SOARES et al., 2002; SEGURA, 2005).

Desta forma, cada árvore foi desfolhada, sendo suas folhas pesadas no campo com célula de carga e ensacadas. Do conjunto de folhas destacadas foi retirada uma amostra de aproximadamente 100 g, para determinação do peso seco, após secagem em estufa de circulação forçada de ar a $75 \pm 2^\circ\text{C}$. Em seguida, cada árvore foi desganhada e todos os seus galhos foram pesados no campo com dinamômetro, inclusive a ponta da árvore com diâmetro menor que 5 cm. No conjunto de 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 100 g, para determinação do seu peso seco, após secagem em estufa de circulação forçada de ar a $103 \pm 2^\circ\text{C}$.

O tronco da árvore foi seccionado e pesado com auxílio do dinamômetro. Discos de madeira com casca de aproximadamente 5 cm de espessura, nos seguintes níveis de medição: 0, 50, e 100% da altura comercial (até 5 cm de diâmetro) foram extraídos. Os discos foram pesados para determinação do peso seco, após secagem em estufa de circulação forçada de ar a $75 \pm 2^\circ\text{C}$.

A partir da razão entre massa verde (Kg) e massa seca (Kg) das amostras multiplicada por 100, foi obtido o teor de umidade dos galhos, folhas e tronco das árvores em porcentagem (%). A biomassa de galhos e folhas foi obtida de acordo com a seguinte expressão:

$$bc_i = m_i(1 - u_i/100), \quad (4.1)$$

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

A biomassa do tronco das árvores, com casca foi obtida de acordo com a mesma expressão:

$$bt_i = m_i(1 - u_i/100), \quad (4.2)$$

onde, i é o índice que se refere às árvores; bt_i 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 (%).

A biomassa total de cada árvore foi obtida pela soma da biomassa do fuste e biomassa da copa. As Tabelas 4.3 e 4.4 trazem a amplitude de variação dos dados coletados de biomassa, bem como, das respectivas variáveis dendrométricas.

Tabela 4.3 – Informações acerca da amplitude de variação dos dados coletados nas parcelas.

Parcela	Biomassa total (Kg)			DAP (cm)			Altura total (m)		
	Min.	Média	Max.	Min.	Média	Max.	Min.	Média	Max.
1	3,6	40,8	347,6	4,5	11,5	26,7	9,7	18,1	32,9
2	6,3	23,6	464,4	5,4	13,9	25,5	10,9	20,6	31,2
3	3,3	19,2	96,1	3,8	13,4	38,2	7,1	19,5	44,9
4	5,4	23,2	79,9	3,8	9,7	17,5	7,9	14,6	21,4
5	5,5	17,7	39,0	4,1	13,0	23,2	8,8	21,0	30,9
6	6,3	23,6	60,4	7,0	13,8	22,9	14,1	21,9	29,0
7	18,3	328,0	1429,0	3,8	9,3	13,0	8,0	12,6	15,9
8	4,4	168,1	1769,0	4,1	11,6	24,2	8,3	18,9	27,2
9	11,7	541,4	4200,0	4,4	8,8	13,0	5,4	11,4	15,6
10	8,2	22,8	60,5	4,1	12,2	22,9	8,7	18,7	33,5
Todas*	3,3	96,5	4200,0	4,3	13,2	81,0	4,8	9,7	23,6

Tabela 4.4 - Informações acerca da amplitude de variação das árvores coletadas por espécie.

Espécie	Biomassa total (Kg)			DAP (cm)			Altura total (m)		
	Min.	Média	Max.	Min.	Média	Max.	Min.	Média	Max.
<i>Endlicheria paniculata</i> *	3,3	20,5	96,2	5,0	9,8	27,4	8,0	11,6	17,4
<i>Calophyllum brasiliense</i> *	5,4	475,3	4200,1	5,0	25,7	81,5	5,6	8,1	12,1
<i>Pera glabrata</i> *	4,4	29,9	156,9	5,2	10,1	28,6	7,2	14,0	23,6
<i>Myrcia racemosa</i> *	4,5	16	40,5	5,0	7,3	15,7	5,9	9,7	15,2
<i>Alchornea triplinervia</i> *	8,3	84,9	302,1	7,2	16,6	43,7	5,6	6,9	9,2
<i>Tibouchina pulchra</i> *	7,7	44,6	121,4	5,3	12,4	19,5	5,9	9,2	13,8
<i>Andira fraxinifolia</i> *	9,3	149,8	584,1	6,2	23,8	70,2	6,3	10,0	15,1
<i>Gomidesia affinis</i> *	3,7	19,7	58	5,3	7,4	11,2	5,3	8,2	9,6
<i>Jacaranda puberula</i> *	7,1	26,6	55,5	7,2	12,0	17,3	6,7	9,8	11,9
<i>Psidium cattleianum</i> *	5,1	16,7	47,3	5,4	7,7	10,3	6,1	8,4	10,7
<i>Coussapoua microcarpa</i> *	225,3	538,9	1429,0	12,2	27,9	52,0	13,2	17,1	23,0
<i>Ilex theezans</i> *	5,7	66,9	212,7	7,0	13,2	22,5	7,9	11,7	14,9
<i>Eugenia ceraciflora</i> *	9,5	60,9	177,7	5,9	11,0	18,3	6,3	9,9	15,8
<i>Andira fraxinifolia</i> *	10,3	39,6	102,5	9,2	16,9	36,4	6,7	8,6	11,1
<i>Ocotea sp.</i> *	7,6	31,1	63,4	5,8	10,1	14,5	6,6	9,7	12,4
<i>Tapirira guianensis</i> *	3,6	39,5	136,9	5,8	9,6	18,0	7,9	10,2	16,5
	3,3	96,5	4200,0	4,3	13,2	81,0	4,8	9,7	23,6

4.2.3 Métodos

4.2.3.1 O modelo de biomassa

Para o desenvolvimento do modelo, partiu-se da função não linear que descreve a biomassa de árvores individuais em função do seu diâmetro à altura do peito (dap) e da sua altura (ht) (eq. 4.3). Esse modelo é chamado de modelo alométrico quando se tem a biomassa da árvore relacionada, apenas com o seu dap :

$$b = \beta_0 dap^{\beta_1} ht^{\beta_2} \quad (4.3)$$

No entanto, como a predição de efeitos aleatórios é mais conveniente utilizando os modelos lineares, o modelo proposto foi linearizado por transformação logarítmica. Assim, o modelo para predição da biomassa da i -ésima árvore é

$$\ln b_i = \beta_0 + \beta_1 \ln dap_i + \beta_2 \ln ht_i + \epsilon_i, \quad (4.4)$$

onde b_i é a biomassa total da árvore (Kg); dap_i é o diâmetro à altura do peito (cm); ht_i é a altura total da árvore (m), β_0 , β_1 , β_2 são os parâmetros do modelo e, ϵ_i é o erro aleatório residual, $\epsilon_i \sim NID(0, \sigma^2)$. Este é o modelo linear de referência que foi utilizado neste trabalho para obtenção das predições da média marginal.

A partir do modelo marginal (eq. 4.4), assume-se que a sua forma funcional é fixa, mas que alguns dos seus parâmetros podem variar de acordo com dois grupos distintos: as diferentes parcelas e as diferentes espécies que compõe a amostra. Desta forma obtém-se um modelo linear de efeitos mistos para a predição da biomassa, considerando uma estrutura de efeitos aleatórios cruzados.

Na modelagem de efeitos aleatórios cruzados não se tem grupos dentro de grupos. Ao invés disso, pode se agrupar os dados de modo que o pertencimento a um grupo não restringe o pertencimento a outro grupo (EERIKÄINEN, 2009).

Para o ajuste do modelo de efeitos mistos com efeitos aleatórios cruzados, testaram-se várias combinações de parâmetros para compor a parte aleatória. Com finalidade comparativa, procedimento semelhante também foi aplicado ajustando-se modelos de um único efeito para os dois grupos: espécie e parcela. A seleção do modelo de efeitos mistos

final, entre os candidatos foi feita através do teste de razão de verossimilhança e dos valores do critério de informação de Akaike (AIC). Assim, o modelo linear de efeitos mistos para a árvore i da espécie k na parcela p , selecionado, tem a seguinte forma:

$$\ln b_{kpi} = \beta_0 + \beta_1 \ln dap_{kpi} + \beta_2 \ln ht_{kpi} + \alpha_{0k} + \alpha_{2k} \ln ht_{kpi} + \gamma_{0p} + \gamma_{2p} \ln ht_{kpi} + \epsilon_{kpi} \quad (4.5)$$

onde α_{0k} e α_{2k} são os normalmente distribuídos efeitos aleatórios em nível de espécie e γ_{0k} e γ_{2k} são os normalmente distribuídos efeitos aleatórios em nível de parcela .

Para uma apresentação formal da predição de efeito aleatório, precisou-se reformular o modelo, apresentando-o na forma matricial para uma única espécie k na parcela p , como segue:

$$\ln b_{kp} = \mu_{kp} + Z_k \alpha_k + Z_p \gamma_p + \epsilon_{kp}, \quad (4.6)$$

onde o vetor $\ln b_{kp}$ inclui as biomassas das árvores observadas da espécie k na parcela, μ_{kp} inclui as predições com base na parte fixa do modelo de efeitos aleatórios, fornecendo o valor esperado do vetor $\ln b_{kp}$, Z_k é a matriz que define a parte aleatória em nível de espécie, $\alpha_k = (\alpha_{0k}, \alpha_{2k})'$ inclui os efeitos aleatórios de espécie, grupo-específicos, que têm distribuição normal com média 0 e matriz de variância e covariância matriz $Dk_{2 \times 2}$, definida por

$$Dk = var(\alpha_k) = var \begin{bmatrix} \alpha_{0k} \\ \alpha_{2k} \end{bmatrix} = \begin{bmatrix} var(\alpha_{0k}) & cov(\alpha_{0k}, \alpha_{2k}) \\ cov(\alpha_{2k}, \alpha_{0k}) & var(\alpha_{2k}) \end{bmatrix},$$

onde Z_p é a matriz que define a parte aleatória em nível de parcela, $\alpha_k = (\alpha_{0k}, \alpha_{2k})'$ inclui os efeitos aleatórios de espécie, grupo-específicos, que têm distribuição normal com média 0 e matriz de variância e covariância matriz $Dp_{2 \times 2}$, definida por

$$Dp = var(\alpha_p) = var \begin{bmatrix} \alpha_{0p} \\ \alpha_{2p} \end{bmatrix} = \begin{bmatrix} var(\alpha_{0p}) & cov(\alpha_{0p}, \alpha_{2p}) \\ cov(\alpha_{2p}, \alpha_{0p}) & var(\alpha_{2p}) \end{bmatrix},$$

e ϵ_{kp} é o erro aleatório residual também normalmente distribuído com média 0 e variância

$$\text{var}(\epsilon_{kp}) = R = \sigma^2 I.$$

A matriz Dk , bem como a estrutura de R é comum à todas as espécies enquanto que a matriz Dp , bem como a estrutura de R é comum à todas as parcelas. Essas matrizes são estimadas na fase de ajuste do modelo. No entanto, os valores dos vetores de efeitos aleatórios são preditos posteriormente usando o melhor preditor linear não enviesado (BLUP). Como a predição é baseada nas árvores-amostra retiradas do conjunto de dados, as matrizes Z_k e Z_p são idênticas.

4.2.3.2 Calibração com dados empíricos

Para predizer os efeitos aleatórios e calibrar os modelos, foi usado o melhor preditor linear não enviesado (BLUP). No caso geral, considerando um vetor não observado b_1 e um vetor observado b_2 , o BLUP para b_1 é:

$$\widehat{b}_1 = \mu_1 + V_{12} V_2^{-1} (b_2 - \mu_2), \quad (4.7)$$

com variância

$$\text{var}(\widehat{b}_1 - b_1) = V_1 + V_{12} V_2^{-1} V_{12}', \quad (4.8)$$

onde \widehat{b}_1 é a predição do vetor não observado, μ_1 e μ_2 são as médias de b_1 e b_2 , V_1 e V_2 são as variâncias de b_1 e b_2 e V_{12} é a covariância entre V_1 e V_2 .

Neste trabalho, a forma geral do BLUP foi adaptada para a fim de realizar as predições em nível de parcela e em nível de espécie, predizendo os efeitos aleatórios destes dois grupos.

Aplicar o BLUP em nível de parcela significa assumir que dada uma árvore i localizada numa determinada parcela p , a predição será realizada em nível de parcela e os efeitos aleatórios preditos serão válidos para todas as árvores da parcela p , independente de que espécie ela pertença. Em contrapartida, aplicar o BLUP em nível de uma espécie k significa predizer os efeitos aleatórios para este grupo, tornando-os válidos para todas as árvores desta espécie k , independente em que parcela a árvore esteja localizada. Por consequência, para árvores da espécie k e da parcela p em questão, a calibração consiste na

adição de ambos os efeitos, de parcela e de espécie.

A forma do BLUP é a mesma para as duas situações, mudando apenas a matriz de efeitos aleatórios considerada. A forma geral do BLUP para essas duas situações, que são posteriormente referenciadas como situações 1 e 2, são apresentadas a seguir.

4.2.3.2.1 Predição em nível de espécie (Situação 1)

Para a calibração em nível de espécie o BLUP se torna:

$$\widehat{\alpha}_k = 0 + D_k Z_k' (Z_k D_k Z_k' + R)^{-1} (\ln b_k - \widehat{\mu}_k), \quad (4.9)$$

e a variância da predição de $\widehat{\alpha}_k$ é

$$\text{var}(\widehat{\alpha}_k - \alpha_k) = D_k - D_k Z_k' (Z_k D_k Z_k' + R)^{-1} Z_k D_k \quad (4.10)$$

onde D_k , Z_k e R são matrizes relacionadas com o modelo descrito na seção 4.3.1.

Nesta situação, o vetor observado b_1 inclui a biomassa das árvores-amostra e o vetor não observado b_2 inclui os efeitos aleatórios para a espécie em questão. As biomassas observadas das árvores amostra da espécie k são escritos no vetor $\ln b_k$ e as predições correspondentes que utilizam os diâmetros e alturas medidas, no vetor $\widehat{\mu}_k$.

4.2.3.2 Predição em nível de parcela (Situação 2)

Para a predição em nível de parcela o BLUP se torna:

$$\widehat{\gamma}_p = 0 + D_p Z_p' (Z_p D_p Z_p' + R)^{-1} (\ln b_p - \widehat{\mu}_p), \quad (4.11)$$

e a variância da predição de $\widehat{\gamma}_p$ é

$$\text{var}(\widehat{\gamma}_p - \gamma_p) = D_p - D_p Z_p' (Z_p D_p Z_p' + R)^{-1} Z_p D_p \quad (4.12)$$

onde D_p , Z_p e R são matrizes relacionadas com o modelo descrito na seção 4.3.1.

Nesta situação, o vetor observado b_1 inclui a biomassa das árvores-amostra e o vetor não observado b_2 inclui os efeitos aleatórios para a parcela em questão. As biomassas observadas das árvores amostra da parcela p , são escritas no vetor $\ln b_p$ e as previsões correspondentes que utilizam os diâmetros e alturas medidas, no vetor $\widehat{\mu}_p$.

4.2.3.3 Aplicação a dados empíricos

A fim de avaliar a eficiência da calibração foi aplicado um procedimento de validação cruzada ao conjunto de dados de biomassa descrito na seção 4.2. Este procedimento consiste em retirar um elemento e de um determinado grupo, por vez, dentre o conjunto de elementos disponíveis, no grupo. Sendo assim, quando o grupo é o conjunto de espécies os elementos retirados são as k espécies e quando o grupo é representado pelo conjunto de parcelas os elementos retirados são as p parcelas.

A implementação do procedimento de validação cruzada consiste em três etapas: (i) Ajuste dos modelos das equações 4.4 e 4.5 ao conjunto de dados, (ii) calibração do modelo usando três árvores-amostra, selecionadas aleatoriamente do elemento e do grupo selecionado e que foi retirado do conjunto de dados de ajuste e (iii) a previsão da biomassa das árvores do grupo selecionado que não foram utilizadas no passo de calibração (iv) e avaliação das previsões contra as biomassas reais.

Como o grupo espécie apresentou alguns elementos com apenas poucas árvores, somente a fase de ajuste (i) foi aplicada ao conjunto de dados completo. Nas demais fases foram usadas somente às espécies que apresentaram pelo menos quatro árvores (Tabela 4.2).

Com relação ao grupo de parcelas o ajuste foi feito com os dados completos, considerando que árvores isoladas constituem parcelas de um único indivíduo. Na amostra de calibração e de previsão estes dados foram reduzidos às parcelas apresentadas na Tabela 4.1.

4.2.3.3.1 Estimação

Com todas as árvores de determinado grupo, exceto as do elemento e retirado, um modelo linear de efeito fixo (modelo marginal) é ajustado por meio de mínimos quadrados ordinários. Através deste procedimento, as estimativas de $\beta_0, \beta_1, \beta_2$ para a eq. 4.4 são obtidas.

Com as mesmas árvores, o modelo linear de efeitos mistos foi ajustado utilizando máxima verossimilhança restrita. Isso resulta nas estimativas dos parâmetros fixos (β_0, β_1 e

β_2), as variâncias e covariâncias dos efeitos aleatórios ($\text{var}(\alpha_{0k}; \alpha_{2k})$, $\text{var}(\gamma_{0k}; \gamma_{2k})$) e na variância do resíduo (ϵ_{kpi}) da eq. 4.5.

4.2.3.3.2 Calibração

Na calibração, os efeitos aleatórios dos modelos de efeitos mistos foram preditos usando um pequeno número de árvores-amostra do elemento e que não foi utilizada na estimação do modelo. As árvores do elemento em questão foram classificadas em classes de tamanho de acordo com seu diâmetro, sendo amostradas uniformemente em todas as classes de diâmetro. Este procedimento foi repetido mil vezes, para evitar a influência dos erros de amostragem nos resultados finais. Os modelos se tornam calibrados, adicionando-se os valores preditos dos dois efeitos aleatórios às estimativas de parâmetros fixos da equação (4.5). A adição desses efeitos aleatórios é feita segundo a descrição de aplicação do BLUP, no modelo de efeitos Aleatórios cruzados, apresentada na seção 4.2.3.2.

4.2.3.3.3 Predição

Nesta fase, utilizou-se os modelos marginais, a parte fixa dos modelos de efeitos mistos, e os modelos calibrados, para prever as biomassas das árvores do elemento e , que não foram utilizadas na etapa de calibração. Como as predições estão na escala logarítmica, as biomassas preditas foram transformados de volta para a escala original, usando a abordagem de correção de viés proposta por (LAPPI, 2006):

$$b_{ei} = \frac{e^{[\ln b_{ei} + \sigma]} + e^{[\ln b_{ei} - \sigma]}}{2} \quad (4.13)$$

Onde σ é o desvio padrão do erro de predição de volume em escala logarítmica e o índice e , denota o elemento escolhido, podendo ser k para o grupo espécie e p para o grupo parcela.

4.2.3.3.4 Critérios de avaliação

A avaliação de cada abordagem foi feita através do através do erro quadrático médio

(RMSE) e do viés, como segue:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (b_{ei} - \bar{b}_{ei})^2}{n}} \quad (4.14)$$

$$VIÉS = \frac{\sum_{i=1}^n (b_{ei} - \bar{b}_{ei})}{n} \quad (4.15)$$

O RMSE e o viés foram calculados usando todas as predições a fim de comparar a resposta em desempenho de calibração nos dois grupos: espécie e parcela. Estes dois critérios foram também calculados na escala relativa, dividindo-os pelo valor da biomassa média observada e multiplicando-se por 100.

4.2.4 Resultados

Todas as hipóteses consideradas para estes modelos foram atendidas tornando-os adequados para as aplicações propostas neste trabalho. As estimativas dos parâmetros para os modelos marginais (eq. 4.2) e condicionais (eq. 4.3), ajustados ao conjunto de dados completo, descrito na seção 4.2, são apresentadas na Tabela 4.5. Estes estão sendo reportados visando futuras aplicações, já que na proposta de validação cruzada deste trabalho sempre um elemento é retirado de um dos grupos na fase de estimação.

A partir da observação da Tabela 4.5, é possível notar uma contribuição relativa importante do efeito parcela na variância total, principalmente em comparação ao efeito de espécie. Isto porque os valores do desvio padrão dos efeitos aleatórios de parcela são maiores que os valores do desvio padrão dos efeitos de espécie.

Cabe ressaltar, que esse modelo foi ajustado aos dados completos enquanto que os modelos aplicados no procedimento de calibração e predição consideram apenas um subconjunto desses dados, conforme descrito na seção 4.2.

Tabela 4.5 – Estimativas dos parâmetros do modelo marginal e do modelo condicional ajustados aos dados completos.

Parâmetros	Estimativas	
	Modelo marginal	Modelo condicional
$\widehat{\beta}_0$	-2.71008 (0.24502)	-2.87141 (0.38733)
$\widehat{\beta}_1$	1.64740 (0.09153)	1.59305 (0.08773)
$\widehat{\beta}_2$	0.94106 (0.16701)	1.06588 (0.21455)
$d.p.(\alpha_{0k})$	-	0.42231
$d.p.(\alpha_{2k})$	-	0.18005
$d.p.(\gamma_{0p})$	-	1.31495
$d.p.(\gamma_{2p})$	-	0.60214
$cor(\alpha_{0k}, \alpha_{2k})$	-	-0,998
$cor(\gamma_{0p}, \gamma_{2p})$	-	-0.864
$d.p.(\epsilon_{kpi})$	0,4926	0.39543

Situação 1: Calibração e predição em nível de espécie

A Figura 4.1 A mostra o gráfico do RMSE dos modelos ajustados na situação de calibração em nível de espécie. Nesta abordagem, o RMSE da parte fixa do modelo de efeito fixo foi menor que o RMSE do modelo marginal, O modelo calibrado apresentou uma tendência de redução do valor do RMSE, com o aumento do número de árvores utilizadas no procedimento de calibração. Com apenas uma única árvore o modelo calibrado é 5% melhor que o modelo marginal em termos de precisão, chegando a quase 8% com três árvores usadas na calibração.

A Figura 4.1 B mostra o gráfico do viés dos modelos ajustados nesta mesma situação de calibração. O procedimento de calibração também foi eficaz para reduzir o viés. Com apenas uma única árvore, o modelo calibrado reduz o viés pela metade em relação ao modelo marginal. No entanto, a inclusão de novas árvores não foi efetiva na redução do viés e o viés remanescente, após a utilização de 3 árvores na calibração, foi de cerca de 12 %.

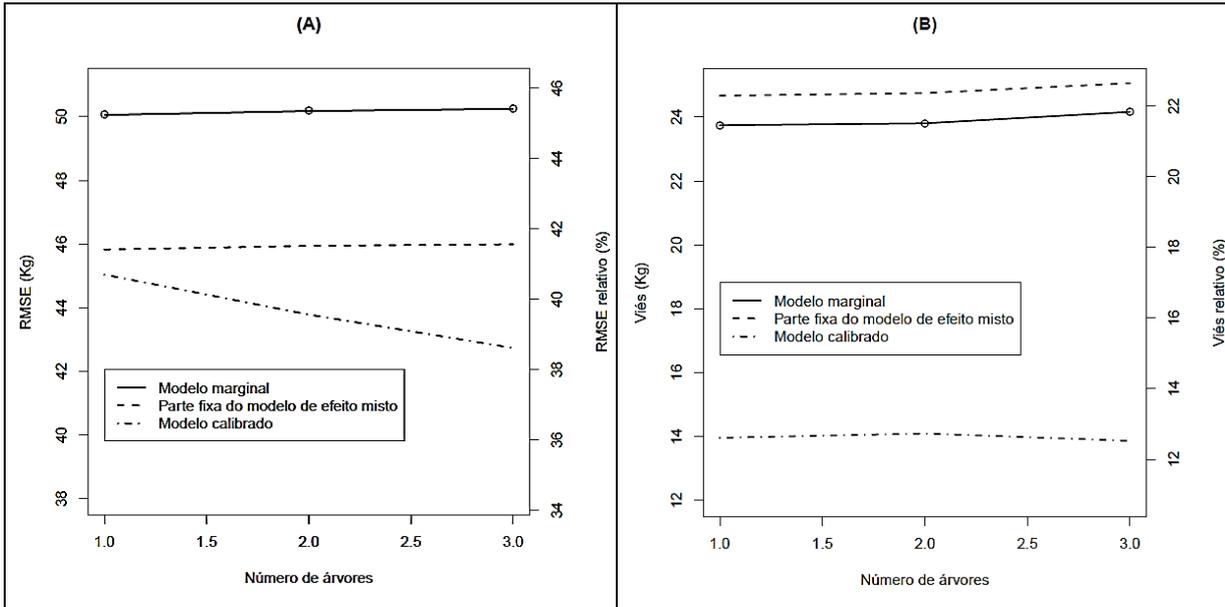


Figura 4.1 – Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE (A) e do Viés (B) na situação 1 (calibração em nível de espécie).

Situação 2: Calibração e predição em nível de parcela

A Figura 4.2 A mostra o gráfico do RMSE dos modelos ajustados na situação de calibração em nível de parcela. Nesta abordagem, o RMSE da parte fixa do modelo de efeito fixo foi sempre muito menor que o RMSE do modelo marginal. O modelo calibrado, no entanto, não apresentou uma tendência de redução do RMSE com o aumento do número de árvores utilizadas no procedimento de calibração. Pelo contrário, o modelo calibrado é ligeiramente pior que o modelo de efeito fixo, demonstrando que a calibração, neste caso, não foi efetiva.

O gráfico do viés dos modelos ajustados para predição em nível de parcela (Figura 4.2 B) mostra o mesmo comportamento do gráfico de RMSE. Ou seja, o procedimento de calibração também não foi eficaz para reduzir o viés, mas o viés da parte fixa do modelo de efeitos mistos e do modelo calibrado é bem menor que o viés do modelo marginal.

Cabe ressaltar, no entanto, que os valores, tanto do RMSE quanto do viés nesta situação, foram inferiores aos observados quando a calibração e a predição são feitas em nível de parcela. O RMSE relativo permanece alto (31%) enquanto que o viés relativo, de cerca de 4,5%, se torna bastante aceitável.

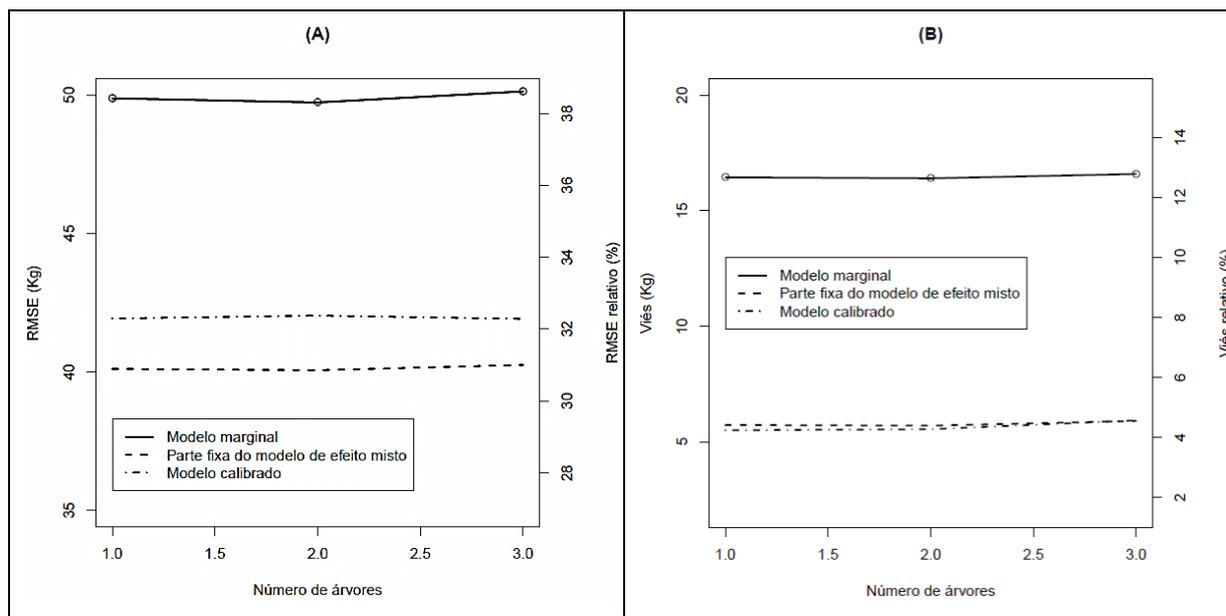


Figura 4.2 - Efeito do aumento do número de árvores amostradas nos valores absolutos e relativos do RMSE (A) e do Viés (B) na situação 2 (calibração em nível de parcela).

4.3 Discussão e Conclusões

O presente estudo apresentou um modelo de efeitos mistos em dois níveis (espécie e parcela) para a predição de biomassa numa floresta tropical, localizada em Registro, SP. A calibração do modelo foi feita nestes dois níveis e objetivou melhorar a qualidade das predições de biomassa na floresta em questão.

Primeiramente, foi realizado um procedimento de calibração em nível de espécie, ignorando o efeito da parcela. Num segundo momento, o procedimento de calibração foi realizado em nível de parcela ignorando o efeito de espécie.

A abordagem de múltiplos efeitos aleatórios já foi utilizada no meio florestal para modelar efeitos aninhados, tais como compartimentos da árvore inseridos no nível de árvore, árvores inseridas no nível de parcelas, etc. Esse tipo de análise é frequentemente utilizada em levantamentos com medidas repetidas (). No entanto, a utilização de um modelo de múltiplos efeitos aleatórios cruzados e calibração traz uma contribuição metodológica, principalmente do ponto de vista de aplicação às florestas tropicais multiespecíficas.

A calibração em nível de espécie se mostrou eficaz em melhorar a qualidade das predições de biomassa das diferentes espécies. O processo de calibração sempre reduziu o valores de RMSE e do viés, fazendo do modelo calibrado melhor que o correspondente modelo marginal.

A calibração em nível de parcela, no entanto, não foi efetiva em melhorar a qualidade

das predições. Nesta situação a parte fixa do modelo de efeito misto gerou predições muito mais precisas que o equivalente modelo marginal enquanto que o modelo calibrado piora ligeiramente a qualidade dessas predições.

Algo que chama atenção nos resultados é o fato de que no ajuste do modelo global apresentado na Tabela 4.5 a contribuição relativa do efeito de parcela é muito grande em relação ao resíduo do modelo. No procedimento de ajuste, no entanto, considerou-se todas as 90 parcelas, pois as árvores isoladas foram consideradas como sendo parcelas que continham uma única árvore. Mas de fato, somente 10 parcelas, totalizando 139 árvores, foram usadas no procedimento de calibração e predição.

Em contrapartida, em nível de espécie só foram excluídas dos dados de calibração e predição espécies com menos de 4 representantes. Desta forma, na amostra de calibração e predição tinha-se uma amostra com 16 espécies e 189 árvores, que são valores bem mais próximos dos dados completos.

De qualquer forma, a inclusão do nível de parcela no modelo contribuiu para melhorar as predições, pois quando comparado com os modelos de um único nível o modelo de dois níveis foi mais bem avaliado em termos de AIC e do teste da razão de verossimilhança. Isto porque, muito do efeito atribuído à parcela no modelo global é de fato um efeito de árvore individual, que dada as características extremamente variáveis da amostra em questão, se tornam muito relevantes.

O fato da calibração em nível de parcela ter sido inútil, indicar que não há muita variação entre árvores além da variação entre as espécies nos utilizados na predição (dados com várias árvores por parcela).

O sucesso da calibração em nível de espécie indica que é possível calibrar o modelo para uma nova espécie com apenas uma única árvore-amostra. Em termos práticos, essa metodologia pode ser aplicada numa situação onde em determinada parcela se encontrasse uma nova espécie que não fora incluída no processo de construção do modelo.

Referências

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

_____. Comparison of formulae for biomass content determination in a tropical rain forest site in the state of Pará, Brazil. **Forest Ecology and Management**, Winsconsin, v. 117, n. 1/3, p. 43-52, May 1999.

BIGING, G.S. Improved estimates of site index curves using a varying-parameter model. **Forest Science**. Bethesda, v. 31, p. 248-259, 1985.

BROWN, S. **Estimating biomass and biomass change of tropical forests: a primer**. Rome: FAO, 1997. 55 p. (FAO. Forestry Paper, 134).

CALEGARIO, N.; DANIELS, R.F.; MAESTRI, R.; NEIVA, R. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal *Eucalyptus* plantation case study. **Forest Ecology and Management**, Winsconsin, n. 204, p. 11-20, 2005.

CALAMA, R.; MONTERO, G. Interregional nonlinear height-diameter model with random coefficients for Scotch pine in Spain. **Canadian Journal of Forest Research**, Vancouver, v. 34, n. 1, p. 150-163, 2004.

CASTEDO-DORADO, F.; DIEGUEZ ARANDA, U.; BARRIO ANTA, M.; SANCHEZ RODRIGUES, M.; GADOW, K.V. A generalized height-diameter model including random components for radiata pine plantations in northwestern Spain. **Forest Ecology and Management**, Winsconsin, v. 229, n. 1/3, p. 202-213, July 2006.

CHAVE, J.; ANDALO, C.; BROWN, S.; CAIRNS, M.; CHAMBERS, J.C.; EAMUS, D., FÖLSTER, H.; FROMARD, F.; HIGUCHI, N.; KIRA, T.; LESCURE, J.; NELSON, B.W.; OGAWA, H.; PUIG, H.; RIÉRA, B.; YAMAKURA, T. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. **Oecologia**, Heidelberg, v. 145, n. 1, p. 87-99, June 2005.

CLARK, D.B.; CLARK, D.A. Landscape-scale variation in forest structure and biomass in a tropical rain forest. **Forest Ecology and Management**, Winsconsin, v. 137, n. 1/3, p. 185-198. Oct. 2000.

DORADO, F.C.; DIEGUEZ-ARANDA, U.; ANTA, M.B.; RODRIGUES, M.S.; VON GADOW, K. A generalized height-diameter model including random components for radiata pine in north-western Spain. **Forest Ecology and Management**, Winsconsin, v. 229, n. 1/3, p. 202-213, July 2006.

EERIKAINEN, K. A multivariate linear mixed-effects model for the generalization of sample tree heights and crown ratios in the finnish national forest inventory. **Forest Science**, Bethesda, v. 55, n. 6, p. 480-493, June 2009.

HALL, D.B.; BAILEY, R. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. **Forest Science**, Bethesda, v. 47, n. 3, p. 311-321, Mar. 2001.

HALL, D.B.; CLUTTER, M. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. **Biometrics**, Georgia, v. 60, p.16 -24, Mar. 2004.

- LAPPI, J. Calibration of height and volume equations with random parameters. **Forest Science**, Bethesda, v. 37, n. 3, p. 781-801, 1991.
- _____. A longitudinal analysis of height/diameter curves. **Forest Science**, Bethesda, v. 43, n. 4, p. 555-570, 1997.
- LAPPI, J.; BAILEY, R.L. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. **Forest Science**, Bethesda, v. 34, p. 907-927, 1988.
- MEHTÄTALO, L. Height-diameter models for scots pine and birch in Finland. **Silva Fennica**, Helsinki, v. 39, n. 1, p. 55-66, 2005a.
- _____. Localizing a predicted diameter distribution using sample information. **Forest Science**, Bethesda, v. 51, n. 4, p. 292-302, 2005b.
- MEHTÄTALO, L.; COMAS, C.; PUKKALA, T.; PALAHÍ, M. Improving predicted diameter distribution by using a small sample of diameters. In press.
- NOTHDURFT, A.; KUBLIN, E.; LAPPI, J. A non-linear hierarchical mixed model to describe tree height growth. **European Journal of Forest Research**, Estocolmo, v. 125, p. 281-289, Mar. 2006.
- OVERMAN, J.P.M.; WITTE, H.J.L.; SALDARRIAGA, J.G. Evaluation of regression models for above-ground biomass determination in Amazon rainforest. **Journal of Tropical Ecology**, Cambridge, v. 10, n. 2, p. 207-218, May 1994.
- PESONEN, A.; EERIKAINEN, K.; MALTAMO, M.; TAHVANAINEN, T. Models for predicting tree and stand development on larch plantations in Hallormsstaour, Iceland. **New Forests**, Oxford, v. 37, p. 63-83, 2009.
- RAO, C.R. Simultaneous estimation of parameters in different linear models and applications to biometric problems. **Biometrics**, London, v. 31, p. 545-554, 1975.
- SUNANDA, C.; JAYRAMAN, K. Prediction of stand attributes of even-aged teak stands using multilevel models. **Forest Ecology and Management**, Winsconsin, v. 236, p. 1-11, May. 2006
- TEMESGEN, H.; GADOW, K.V. Generalized height-diameter models—an application for major tree species in complex stands of interior British Columbia. **European Journal of Forest Research**, Freising, v. 123, p. 45-51, 2004.
- VONESH, E.F.E; CARTER, R.L. Efficient inference for random-coefficient growth curve models with unbalanced data. **Biometrics**, London, v. 43, p. 617-628, 1987.
- YANG, Y.; HUANG, S.; MENG, S.; TRINCADO, G.; VANDERSCHAAF, C.L. A multilevel individual tree basal area increment model for aspen in boreal mixed wood stands. **Canadian Journal of Forest Research**, Vancouver, v. 39, p. 2203-2214, Feb. 2009.

ZHANG, Y.; BORDERS, B.E. Using a system mixed-effects modeling method to estimate tree compartment biomass for intensively managed loblolly pines - an allometric approach. **Forest Ecology and Management**, Winsconsin, v. 194, p. 145–157, Feb. 2004.

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

ANEXO

Tabela - Parâmetros dos modelos de biomassa testados no segundo capítulo.

<i>Estimativas dos parâmetros dos modelos testados com seus respectivos desvios padrão</i>				
EXPRESSÃO DA MÉDIA	COEFICIENTES (DEVIO PADRÃO)			
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
$\beta_0 d_{ij}^{\beta_1}$	-2,8900 (0,1367)	2,6497 (0,0435)	-	-
$\beta_0 d_{ij}^{\beta_1} p_{ij}^{\beta_2}$	-2,1729 (0,1042)	2,6649 (0,0305)	-0,3457 (0,2846)	-
$\beta_0 d_{ij}^{\beta_1} \overline{db} l_j^{\beta_2}$	-2,6636 (0,0988)	2,6356 (0,0315)	0,4645 (0,0389)	-
$\beta_0 d_{ij}^{\beta_1} \overline{db} j^{\beta_2}$	-2,2209 (0,0899)	2,6214 (0,0268)	0,8172 (0,0551)	-
$\beta_0 d_{ij}^{\beta_1} db_{ij}^{\beta_2}$	-2,0235 (0,0941)	2,5366 (0,0271)	0,7396 (0,0457)	-
$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2}$	-3,0710 (0,1627)	2,4679 (0,0918)	0,2715 (0,1216)	-
$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} p_{ij}^{\beta_3}$	-2,4047 (0,1176)	2,3327 (0,0655)	0,4872 (0,0883)	-0,3786 (0,0279)
$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db} l_j^{\beta_3}$	-2,8981 (0,1234)	2,4080 (0,0673)	0,3448 (0,0927)	0,4752 (0,0381)
$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} \overline{db} j^{\beta_3}$	-2,4551 (0,1066)	2,3525 (0,0586)	0,3975 (0,0803)	0,8474 (0,0523)
$\beta_0 d_{ij}^{\beta_1} h_{ij}^{\beta_2} db_{ij}^{\beta_3}$	-2,266 (0,1126)	2,3090 (0,0549)	0,3490 (0,0762)	0,7516 (0,0434)
$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1}$	-3,5638 (0,1801)	0,9911 (0,0198)	-	-
$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} p_{ij}^{\beta_2}$	-2,6775 (0,1363)	0,9950 (0,0140)	-0,4119 (0,0298)	-
$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db} l_j^{\beta_2}$	-3,3950 (0,1426)	0,9952 (0,0156)	0,5092 (0,0427)	-
$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} \overline{db} j^{\beta_2}$	-2,8443 (0,1253)	0,9841 (0,0132)	0,9127 (0,0586)	-
$\beta_0 (d_{ij}^2 h_{ij})^{\beta_1} db_{ij}^{\beta_2}$	-2,7202 (0,1335)	0,9599 (0,0138)	0,8034 (0,0499)	-