



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FLORESTAIS

CLAYTON VIEIRA FRAGA FILHO

**MODELAGEM PARA PROGNÓSE DO CRESCIMENTO E PRODUÇÃO DE
EUCALIPTO EM DIFERENTES NÍVEIS DE ABORDAGEM**

JERÔNIMO MONTEIRO – ES
2016

CLAYTON VIEIRA FRAGA FILHO

**MODELAGEM PARA PROGNÓSE DO CRESCIMENTO E PRODUÇÃO DE
EUCALIPTO EM DIFERENTES NÍVEIS DE ABORDAGEM**

Tese apresentada ao Programa de Pós-Graduação em Ciências Florestais do Centro de Ciências Agrárias da Universidade Federal do Espírito Santo, como parte das exigências para obtenção do Título de Doutor em Ciências Florestais na Área de Concentração Ciências Florestais.

Orientador: Gilson Fernandes da Silva

Coorientador: Adriano Ribeiro de Mendonça

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial de Ciências Agrárias, Universidade Federal do Espírito Santo, ES, Brasil)

F811m Fraga Filho, Clayton Vieira, 1978-
Modelagem para prognose do crescimento e produção de eucalipto em diferentes níveis de abordagem / Clayton Vieira Fraga Filho. – 2016.
280 f. : il.

Orientador: Gilson Fernandes da Silva.
Coorientadores: Adriano Ribeiro de Mendonça.
Tese (Doutorado em Ciências Florestais) – Universidade Federal do Espírito Santo, Centro de Ciências Agrárias.

1. Eucalipto. 2. Produção florestal. 3. Modelos mistos. 4. Mensuração florestal. I. Silva, Gilson Fernandes da. II. Mendonça, Adriano Ribeiro de. III. Universidade Federal do Espírito Santo. Centro de Ciências Agrárias. IV. Título.

CDU: 630

MODELAGEM PARA PROGNOSE DO CRESCIMENTO E PRODUÇÃO DE
EUCALIPTO EM DIFERENTES NÍVEIS DE ABORDAGEM

Clayton Vieira Fraga Filho

Tese apresentada ao Programa de Pós-Graduação em Ciências Florestais do Centro de Ciências Agrárias da Universidade Federal do Espírito Santo, como parte das exigências para obtenção do Título de Doutor em Ciências Florestais na Área de Concentração Recursos Florestais.

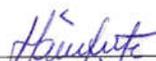
Aprovada em 04 de fevereiro de 2016.



Prof. Dr. Gilson Fernandes da Silva (Orientador)
Universidade Federal do Espírito Santo



Prof. Dr. Adriano Ribeiro de Mendonça (Coorientador)
Universidade Federal do Espírito Santo



Prof. Dr. Hélio Garcia Leite (Membro Externo)
Universidade Federal de Viçosa



Prof. Dr. Carlos Pedro Boechat Soares (Membro Externo)
Universidade Federal de Viçosa



Dr. Daniel Henrique Breda Binoti (Membro Externo)
DAP Engenharia Florestal

Aos meus pais Claiton e Magna por terem me dado amor à sua maneira e me mostrado como seguir meu próprio caminho. À minha mãe, especialmente por ter suportado meus momentos difíceis.

Aos meus avós, pelo carinho e amor.

À minha avó Nilda pelo exemplo de perseverança e pela força que tens. Pelas palavras sempre doces, mesmo nos momentos de correção e pelo amor que sempre me deu.

OFEREÇO

À minha esposa Fernanda e ao meu filho João Vitor, pelo amor que têm por mim, pelas palavras e gestos de carinho que me fortalecem. Sem vocês eu não teria conseguido, amo vocês.

DEDICO

"Por mais árdua que seja a luta, por mais distante que um ideal se apresente, por mais difícil que seja a caminhada, existe sempre uma maneira de vencer: a nossa fé!"

Agradecimentos

Agradeço à minha família pelo amor incondicional e ensinamentos que me deram, por serem minha referência – a todos vocês, especialmente aos meus tios e tias, por serem meus segundos pais e mães, pela sinceridade e amor. À tia Marilda, minha segunda mãe, pelo amor, carinho e por acreditar de forma incondicional em mim. Também à tia Leninha, minha terceira mãe (ou segunda? Sou privilegiado por ter três mães), pelo amor e por tudo que fez por mim desde a minha infância, pelo incentivo nos estudos, sem você eu não teria me apaixonado pelos livros.

Aos meus pais pela formação que me deram, por serem a fonte de tudo o que me tornei e pelo amor. Por TUDO.

À minha irmã Ewelyn, pelo amor e pelas palavras de apoio que me ajudaram nos momentos difíceis.

Aos meus sogros Vicente e Flávia (Espetáculo) pelo carinho, pela forma como me acolheram em suas vidas – como um filho. Pela confiança, e por terem concebido uma pessoa tão especial como minha esposa, Fernanda, a quem agradeço de modo especial. Você sabe meu amor que a nossa caminhada sempre nos apresentou desafios, são anos de muito esforço mútuo, e você sempre me apoiando e acreditando em mim. Mais que minha esposa, és minha companheira, minha fortaleza, minha amada. Sem você e nosso filho João Vitor – nosso maior presente, nosso tesouro – como seria possível? Você me fez e me faz ver a vida e o mundo de uma forma melhor e mais leve, seu amor me renova e me faz bem, desejo sempre fazer o mesmo para você.

Ao meu filho João Vitor, que na sua inocência de criança compreendeu a importância deste trabalho, às vezes me perguntava: “Papai, está muito difícil? ” E de modo muito especial sempre me dizia: “Papai, você vai conseguir”

Aos meus amigos pela verdadeira amizade, pelos momentos de felicidade e alegria que vivemos juntos. De forma especial agradeço aos amigos Murrice (Carlos Henrique) e Rudinho (Rudson) pelo incentivo constante, e por terem demonstrado sua verdadeira amizade, compreendendo os momentos em que não pude estar presente. Também agradeço aos meus amigos Wendel e Magda por sempre proferirem palavras de incentivo, por acreditarem no meu potencial e por tudo o que representam em nossas vidas.

Aos meus colegas do Departamento de Computação pelo apoio, especialmente aos amigos Bruno, Geraldo, Antonio e Helder, os dois últimos,

também meus colegas no doutorado, companheiros nas disciplinas, agradeço pela ajuda, pelo otimismo de sempre.

À Universidade Federal do Espírito Santo, ao Departamento de Ciências Florestais e da Madeira e ao Programa de Pós-graduação em Ciências Florestais e a todos os seus membros, pela oportunidade oferecida.

Ao Prof. Gilson pela imensa demonstração de amizade, companheirismo, sinceridade, despretensão, dedicação e serenidade na orientação e condução deste trabalho. Gilson é uma pessoa iluminada e extremamente dedicada – tem meu respeito e amizade verdadeira para toda a vida. Por mais que eu tente, faltarão palavras que expressem o que fez por mim.

Ao Prof. Adriano pela orientação, pela disponibilidade de sempre, pela sinceridade e ajuda.

Aos professores da Banca Examinadora, Hélio Garcia Leite, Carlos Pedro Boechat Soares e Daniel Binoti pelas contribuições, fundamentais para a melhoria do trabalho.

De modo especial, agradeço ao Prof. Hélio Garcia Leite, pelo exemplo de tranquilidade e pela pessoa acessível que é, inclusive nas madrugadas à dentro. Ainda há muito o que aprender, mas seus ensinamentos foram fundamentais para a minha formação. Também agradeço ao Daniel Binoti pela ajuda, pelos ensinamentos, pelas piadas e pelo exemplo de pessoa que é, sempre otimista, positivo e me dando força para seguir adiante. Ao Prof. Carlos Boechat, meu conterrâneo, pelos ensinamentos e parceria.

Conviver com vocês faz de mim uma pessoa melhor.

Um fato ocorrido dias antes da conclusão deste trabalho me chamou a atenção: após muitos dias de trabalho intenso, e muito desânimo e cansaço, olhei para o alto e vi uma Esperança (um inseto) bem acima da minha mesa de estudos. No momento não dei muita importância. No dia seguinte, viajamos para outra cidade e na casa da minha avó, continuei com o trabalho, ainda muito cansado. A simbologia do nome do inseto fez todo o sentido para mim naquele momento, pois, acima da minha mesa surgiu outra Esperança! Não bastasse, na noite anterior à defesa, já em minha casa, me deparei com mais uma – então, seria coincidência? Entendi naquele momento que não. Eu estava recebendo uma mensagem de força e fé, de um ser superior? De uma ou várias pessoas que rezaram e torceram por mim? Uma mensagem de Deus? De qualquer forma, agradeço por isso.

BIOGRAFIA

Clayton Vieira Fraga Filho, filho de Claiton Vieira Fraga e Magna Maria Schots Fraga, brasileiro, nascido em 23 de agosto de 1978 na cidade de Manhumirim-MG. Casado com Fernanda Vieira Mollica Fraga, e pai de João Vitor Mollica Fraga. Concluiu o ensino médio na Escola Estadual de Manhumirim em 1996, e ingressou no curso de Sistemas de Informação na Faculdade de Viçosa em 2001, concluindo o curso de graduação em 2005. Em maio de 2006 iniciou o mestrado em Ciência da Computação na Universidade Federal de Viçosa (UFV), defendendo sua dissertação em fevereiro de 2009. Durante sua trajetória profissional trabalhou como gerente de Centrais de Processamento de Dados, programador, analista de sistemas e gerente de projetos em empresas em diversos ramos e também desenvolvedoras de software, entre 1994 e 2006, quando ingressou no ensino superior. Em 2009 foi aprovado no concurso para professor efetivo do Departamento de Computação do Centro de Ciências Agrárias da Universidade Federal do Espírito Santo onde iniciou suas atividades. Em março de 2013 iniciou o doutorado no Programa de Pós-graduação em Ciências Florestais da mesma universidade, defendendo sua tese em fevereiro de 2016.

RESUMO

FRAGA FILHO, Clayton Vieira. **MODELAGEM PARA PROGNOSE DO CRESCIMENTO E PRODUÇÃO DE EUCALIPTO EM DIFERENTES NÍVEIS DE ABORDAGEM**. 2016. Tese (Doutorado em Ciências Florestais) – Universidade Federal do Espírito Santo, Jerônimo Monteiro, ES. Orientador: Prof. Dr. Gilson Fernandes da Silva. Coorientador: Prof. Dr. Adriano Ribeiro de Mendonça.

Modelos de crescimento e produção são eficientes e confiáveis para descrever a dinâmica de florestas e realizar a previsão da produção de longo prazo. Existem diferentes níveis de abordagens para estes modelos, tais como modelos em nível de povoamento, distribuição de diâmetros e árvores individuais. Este trabalho teve como objetivo geral comparar diferentes níveis para modelagem do crescimento e produção quanto à acurácia e à tendenciosidade das estimativas de produção em povoamentos de eucalipto, com e sem efeitos aleatórios. Foram utilizados dois critérios para a divisão de dados. No critério 1, dividiu-se os dados em dois grupos: o primeiro para ajuste, com parcelas com idade futura entre 24 e 60 meses, e o segundo para a projeção, com parcelas com idade futura entre 72 e 96 meses. No segundo critério, 70% das parcelas foram usadas no ajuste e 30% para a validação destes modelos. Foram propostas alternativas em nível de povoamento, sendo duas alternativas de modelagem fixa avaliadas para o critério 1, e duas para o critério 2. Além dessas, foram propostas três alternativas de modelagem mista para o critério 1 e sete alternativas para o critério 2. De acordo com análises realizadas para os modelos fixos, o mais acurado e menos viesado foi o Logístico. A modelagem mista para projeção foi superior à modelagem de efeitos fixos. Em seguida, foi realizada uma avaliação de alternativas de modelagem de distribuição de diâmetros com e sem efeito aleatório. Um modelo fixo foi avaliado para ambos os critérios: uma alternativa usando modelagem mista foi proposta para o critério 1 e três alternativas para o critério 2. Concluiu-se que as alternativas ajustadas e validadas com dados do critério 2 foram superiores, sendo a modelagem mista de distribuição de diâmetros a mais acurada e menos enviesada. A partir destes resultados, foram avaliadas alternativas fixas e mistas em nível de árvores individuais. Foram utilizados dois modelos fixos avaliados para o critério 1 e um modelo fixo avaliado para o critério 2. Cinco modelos mistos foram propostos com dados do critério 1 e

seis modelos para o critério 2. Dos modelos fixos, o modelo linear (critério 1) foi superior, contudo superado pela modelagem mista, tanto no ajuste quanto na validação (critérios 1 e 2). Por fim, os modelos mais acurados e menos enviesados obtidos previamente foram comparados. Para realizar a projeção (critério 1) utilizando a modelagem fixa, o modelo Logístico apresentou maior acurácia e menor tendenciosidade. O modelo linear misto de árvores individuais proposto neste trabalho foi superior ao resultado do modelo Logístico. Para o critério 2, o modelo misto adaptado de Clutter foi superior. Assim, conclui-se neste trabalho que a modelagem em nível de árvores individuais produziu resultados similares aos encontrados para os modelos em nível de povoamento, sendo ambos mais acurados que os resultados encontrados para os modelos de distribuição de diâmetros. Conclui-se também que a modelagem mista nestes três níveis proporcionou ganhos em acurácia e ausência de viés na prognose do crescimento e da produção de povoamentos florestais de eucalipto.

Palavras-chave: Mensuração florestal, Modelos de Crescimento e Produção, Modelos mistos.

ABSTRACT

FRAGA FILHO, Clayton Vieira. **GROWTH AND YIELD PROGNOSIS MODELING OF EUCALYPTUS AT DIFFERENT APPROACH LEVELS**. 2016. Thesis (Doctor in Forest Sciences) – Federal University of Espírito Santo, Jerônimo Monteiro, ES. Advisor: Prof. Dr. Gilson Fernandes da Silva. Co-advisor: Prof. Dr. Adriano Ribeiro de Mendonça.

Growth and yield models are efficient and reliable to describe forests dynamics and perform the prediction of long-term production. There are different approach levels for these models, such as whole-stand models, diameter distribution, and individual-tree. This work aimed to compare different levels for the modeling of the growth and production in terms of the accuracy and bias level of the production estimates in eucalyptus stands, with fixed and mixed models. Two criteria were used for dividing the data. In criterion 1, the data were splited into two groups: the first for fitting plots with future age between 24 and 60 months, and the second for projection, having plots with future age between 72 and 96 months. In the second criterion, 70% of the plots were used to models fitting and 30% for the validation of these models. Alternatives at the whole-stand level were proposed, being two of them fixed model alternatives to criterion 1, and two for criterion 2. In addition, three mixed modeling alternatives were proposed for criterion 1 and seven alternatives for criterion 2. According to analyzes for fixed models, the Logistics model was more accurate and less biased. Mixed modeling projection presented superior results compared to fixed effects modeling. Furthermore, an evaluation was performed with fixed and mixed modeling alternatives for diameter distribution level. A fixed model was evaluated for both criteria: an alternative using mixed modeling was proposed for criterion 1 and three alternatives for criterion 2. The conclusion was that alternatives adjusted and validated with criterion 2 data proved better, and the diameters distribution with mixed model was more accurate and less biased compared with approaches with fixed effects models. From these results, fixed and mixed alternatives were evaluated at the individual-tree level. Two fixed models were analyzed for criterion 1 and one fixed model for criterion 2. Five mixed models were proposed for criterion 1 and six models for criterion 2. Among the fixed models, the linear model (criterion 1) was better; however, it was outperformed by mixed modeling in both fit and validation (criteria 1 and 2). Finally, the most accurate and less biased models previously

obtained were compared. For performing the projection (criterion 1) using fixed modeling, the Logistic model is more accurate and less biased. The linear mixed model of individual-tree proposed in this work was better than Logistic model. For criterion 2, the Clutter's adapted mixed model was better. Thus, it was concluded in this work that modeling at individual-tree level produced similar results as the whole-stand level, being both more accurate than the results found for the diameter distribution models. It was also concluded that the mixed modeling at these three levels provided gains in accuracy and lack of bias in prognosis of growth and yield modeling for eucalyptus stands.

Keywords: Forest management, Growth and yield models, Mixed models.

SUMÁRIO

1. INTRODUÇÃO.....	19
1.1. OBJETIVOS.....	23
1.1.1. Objetivo geral	23
1.1.2. Objetivos específicos	23
1.2. Estrutura da Tese	23
2. REVISÃO DE LITERATURA.....	24
2.1. Modelos de Crescimento e produção.....	24
2.1.1. Modelos em nível de Povoamento	26
2.1.2. Modelos de Distribuição de diâmetros	28
2.1.3. Modelos de Árvore Individual	31
2.2. Modelos mistos	34
2.2.1. Formulação	39
2.2.2. Estimação	43
2.2.3. Diagnóstico do modelo	57
2.3. Qualidade dos Dados para processamento de Modelos de Crescimento e Produção.....	57
2.3.1. Localização e correção de erros	61
2.4. Descrição e processamento prévio dos dados	64
2.4.1. Descrição dos dados	64
2.4.2. Processamento dos dados	65
2.4.3. Consistência da base de dados	65
2.4.3.1. Avaliação da distribuição de frequência de cada estado ao longo do tempo	68
2.4.3.1. Identificação de árvores com DAP e HT = 0 e CODIGO diferentes de M, F ou A.....	69
2.4.3.1. Árvores com $DAP_2 < DAP_1$ entre medições consecutivas i e $i+1$, sendo i o ano da medição	70
2.4.3.1. Número de fustes mortos (CODIGO = M) ao longo das medições.....	72
2.4.3.1. Gráficos de interação com a média de uma variável ao longo das medições	72
2.4.4. Avaliação dos resultados	73
2.4.5. Capacidade produtiva, altura total e volume por árvore	74
2.4.5.1. Altura total	79

2.4.5.2 Volume	81
3. REFERÊNCIAS BIBLIOGRÁFICAS	82
CAPÍTULO 1 – MODELAGEM DO CRESCIMENTO E PRODUÇÃO DE EUCALIPTO, EM NÍVEL DE POVOAMENTO, COM E SEM EFEITO ALEATÓRIO	93
1. INTRODUÇÃO	93
2. MATERIAL E MÉTODOS.....	95
2.1. Dados utilizados	95
2.2. Divisão da base de dados para ajuste e validação dos modelos	95
2.3. Alternativas de modelagem em nível de povoamento	96
2.4. Avaliação dos resultados	100
3. RESULTADOS E DISCUSSÃO	102
3.1. Critério 1 – Alternativas usando a variável parcela como grupo de efeito aleatórios	102
3.2. Critério 2 – Alternativas usando outras variáveis como grupos de efeitos aleatórios	109
3.3. Comparação das alternativas ajustadas e validadas para os critérios definidos	124
4. CONCLUSÕES	127
5. REFERÊNCIAS BIBLIOGRÁFICAS.....	128
CAPÍTULO 2 – MODELAGEM DE CRESCIMENTO E PRODUÇÃO EM NÍVEL DE DISTRIBUIÇÃO DE DIÂMETROS COM E SEM EFEITO ALEATÓRIO	131
1. INTRODUÇÃO	131
2. MATERIAL E MÉTODOS.....	134
2.1. Descrição dos dados	134
2.2. Ajuste da Função de Densidade e Probabilidade	134
2.3. Divisão da base de dados para ajuste e validação dos modelos	134
2.4. Modelo de Distribuição de diâmetros	135
2.5. Alternativas de modelagem	136
2.6. Avaliação dos resultados	139
3. RESULTADOS E DISCUSSÃO	141
3.1. Critério 1 – Alternativas usando a variável parcela como grupo de efeito aleatório	142
3.1.1. Resultado das Alternativas de modelagem de distribuição de diâmetros	143

3.1.2. Resultados da aplicação das alternativas de modelagem de distribuição de diâmetros.....	149
3.2. Critério 2 – Alternativas usando outras variáveis como grupos de efeitos aleatórios	154
3.2.1. Alternativas de modelagem para se estimar a distribuição de diâmetros	154
3.2.2. Alternativas de modelagem de distribuição de diâmetros para se estimar o volume por classe de diâmetro.....	160
3.3. Comparação das alternativas ajustadas e validadas para os critérios definidos	166
4. CONCLUSÕES.....	167
5. REFERÊNCIAS BIBLIOGRÁFICAS.....	168
6. APÊNDICES	173
Apêndice A – Volumes estimados e observados obtidos no ajuste e na validação das alternativas 1(não mista) e 2 (mista) para o conjunto de dados do critério 2.	173
 CAPÍTULO 3 – MODELAGEM DO CRESCIMENTO E PRODUÇÃO DE POVOAMENTOS DE EUCALIPTO, EM NÍVEL DE ÁRVORES INDIVIDUAIS, COM E SEM EFEITO ALEATÓRIO	177
1. INTRODUÇÃO.....	177
2. MATERIAL E MÉTODOS.....	179
2.1. Descrição dos dados	179
2.2. Divisão da base de dados para ajuste e validação dos modelos	179
2.3. Alternativas de modelagem em nível de árvores individuais	179
2.4. Avaliação dos resultados	183
3. RESULTADOS E DISCUSSÃO	185
3.1. Avaliação dos modelos de crescimento em diâmetro e altura total	185
3.2. Critério 1 – Alternativas usando outras variáveis como grupos de efeitos aleatórios	189
3.2.1. Ajuste das alternativas de modelagem para crescimento em diâmetro	190
3.2.2. Ajuste das alternativas de modelagem para crescimento em altura total	196
3.2.3. Resultados de volume do modelo de árvore individual	201

3.3. Critério 2 – Alternativas usando o identificador da árvore como grupos de efeitos aleatórios	211
3.3.1. Ajuste das alternativas de modelagem para crescimento em diâmetro	212
3.3.2. Ajuste das alternativas de modelagem para crescimento em altura....	216
3.3.3. Resultados de volume do modelo de árvore individual	220
3.4. Comparação das alternativas ajustadas e validadas para os critérios definidos	229
4. CONCLUSÕES	230
5. REFERÊNCIAS BIBLIOGRÁFICAS	231
CAPÍTULO 4 – MODELAGEM DE CRESCIMENTO E PRODUÇÃO EM DIFERENTES NÍVEIS DE ABORDAGEM PARA POVOAMENTOS DE EUCALIPTO, COM E SEM EFEITOS ALEATÓRIOS.....	
1. INTRODUÇÃO	233
2. MATERIAL E MÉTODOS.....	236
2.1. Descrição dos dados	236
2.2. Divisão da base de dados para ajuste e validação dos modelos	236
2.3. Alternativas de modelagem sem efeito aleatório	236
2.4. Alternativas de modelagem com efeito aleatório	238
2.5. Avaliação dos resultados	241
3. RESULTADOS E DISCUSSÃO	243
3.1. Avaliação de modelos de crescimento e produção em diferentes níveis, utilizando modelagem fixa	243
3.1.1. Volumes totais obtidos por equações ajustados e validados com o conjunto de dados do critério 1, com modelagem fixa.....	243
3.1.2. Volumes totais obtidos por equações ajustados e validados com o conjunto de dados do critério 2, com modelagem fixa.....	248
3.2. Avaliação de modelos de crescimento e produção em diferentes níveis, utilizando modelagem mista	254
3.2.1. Volumes totais obtidos por equações calibradas com a parcela como grupo de efeitos aleatórios, com modelagem mista utilizando dados do critério 1	254
3.2.2. Volumes totais obtidos por equações calibradas para outras variáveis como grupo de efeitos aleatórios, utilizando modelagem mista para o conjunto de dados do critério 2.....	260

3.3. Comparação entre os diferentes níveis de modelagem para ambos os critérios com e sem efeito aleatório	265
4. CONCLUSÕES	267
5. REFERÊNCIAS BIBLIOGRÁFICAS	268
4. CONCLUSÕES GERAIS	271
APÊNDICES	272
Apêndice A – Valores σ^2 , σ_ε^2 , função de máxima verossimilhança restrita (REML) (21) obtido para o exemplo 1	272
Apêndice B – Matrizes $\hat{\mathbf{R}}$ e $\hat{\mathbf{V}}$ do exemplo 3 – Multinível com intercepto	274
Apêndice C – Matrizes $\hat{\mathbf{R}}$ e $\hat{\mathbf{V}}$ do exemplo 4 – Multinível com intercepto e inclinação	276

1. INTRODUÇÃO

Modelos de crescimento e produção florestal são de grande importância nos estudos realizados em Ciências Florestais, pois permitem simular a dinâmica natural do povoamento, fornecendo resultados fundamentais para uma série de tarefas dentro da empresa florestal, como: o planejamento estratégico da empresa, o planejamento da colheita, a aferição de resultados econômicos, dentre outros.

Diferente dos modelos de produção, que exprimem apenas condições correntes da floresta, os modelos de crescimento e produção permitem também inferir sobre condições futuras, sendo um dos principais componentes de um sistema de modelagem de povoamentos florestais (CAMPOS; LEITE, 2013). São definidos como um conjunto de relações matemáticas e tabelas, utilizados para descrever quantitativamente os povoamentos florestais (LEITE, 1990).

Em um breve histórico do desenvolvimento desse tipo de modelo, inicialmente as estimativas de produção eram feitas com base em tabelas de produção, na idade e na qualidade do local, sendo elaboradas por meio de processos gráficos (TREVIZOL JÚNIOR, 1985). Entretanto, esses processos são muito limitados e, principalmente pela sua subjetividade e tempo requerido, dificilmente permitem obter respostas mais fidedignas.

Em substituição aos processos gráficos, Mackinney et al. (1937) foram os primeiros a aplicar a técnica dos mínimos quadrados ao estudo do crescimento e produção. Schumacher (1939) corroborou a superioridade da técnica dos mínimos quadrados comparado ao processo gráfico apresentando um modelo de regressão cujas variáveis independentes eram a qualidade do local, a idade e a densidade do povoamento.

Por outro lado, até o início da década de sessenta, as prognoses eram feitas com equações independentes para explicação de cada componente do modelo, o que poderia conduzir a resultados inconsistentes e ilógicos (GUIMARÃES, 1994).

Com o intuito de superar esse inconveniente e desenvolver modelos consistentes e compatíveis em termos de crescimento e produção, reconheceu-se a relação derivada/integral entre os modelos de crescimento e produção (BUCKMAN, 1962; CLUTTER, 1963).

De acordo com esta relação, ao se integrar a equação de crescimento, obtém-se a produção, e ao derivar a equação de produção, obtém-se o crescimento, surgindo, assim, o conceito de modelos compatíveis.

Com base nesses desenvolvimentos e impulsionado por avanços nos campos da estatística e da computação, os conceitos anteriormente apresentados têm sido amplamente utilizados nas empresas florestais e são constantemente alvo de pesquisas.

Há diversas classificações para modelos de crescimento e produção, dentre as quais destacam-se os modelos empíricos: modelos de em nível de povoamento (MNP), modelos de distribuição de diâmetros (MDD) e modelos de árvores individuais (MAI) (VANCLAY, 1994). Tais modelos são representados por equações nas quais o resultado é válido apenas para as condições específicas em que foram estabelecidos, sendo altamente dependentes dos dados utilizados para seu ajuste (CAMPOS; LEITE, 2013).

Modelos de povoamento total são aqueles que consideram a produção por unidade de área. Esses modelos são tipicamente inflexíveis para avaliar multiprodutos por classe de diâmetro e grande número de tratamentos, sendo limitados a prever apenas a produção total (BURKHART, 1977). Estimam o crescimento e a produção a partir de atributos como idade, área basal e índice de local (CAMPOS; LEITE, 2013).

Como uma alternativa tem-se os modelos de árvore individual, que diferem do modelo de povoamento total pela necessidade de especificar o tamanho de cada árvore dentro do povoamento, em relação ao seu crescimento em diâmetro, altura, e, em alguns casos, do tamanho da copa e sua posição no dossel aumentando a complexidade para definição dos dados e de sistemas para seu ajuste e aplicação (VANCLAY, 1994).

Modelos de distribuição de diâmetros não são tão onerosos como os modelos de árvores individuais e nem apresentam flexibilidade limitada como os modelos de povoamento total, tornando-se, portanto, uma opção a ser considerada. Seu objetivo é fornecer informações sobre a estrutura do povoamento, permitindo assim a estimativa da produção por classes de diâmetro. Com tais informações é possível realizar o planejamento do corte, de desbastes, os tratamentos silviculturais, estimar multiprodutos madeireiros como por exemplo, madeira para serraria, celulose e energia (LEITE, 1990).

Modelos de crescimento e produção são constituídos por sistemas de equações, que por sua vez envolvem um grande número de funções, nas quais as variáveis dependentes de uma função, em alguns casos referem-se às variáveis preditas em outra. Tal característica induz a erros no processo de modelagem em

consequência da propagação de erros de uma função para outra (GERTNER, 1991). As premissas para utilização dos métodos dos mínimos quadrados ordinários consideram que a variável dependente em um modelo tenha distribuição normal, com observações independentes e identicamente distribuídas. Contudo, é comum a existência de correlação, ou autocorrelação, entre unidades de amostragem observadas ao longo do tempo, o que viola a suposição de independência (COSTA, 2003; CALEGARIO; CALEGARIO; et al., 2005).

Dados utilizados na modelagem e crescimento da produção florestal são caracterizados por serem desbalanceados e irregularmente espaçados (GRÉGOIRE et al., 1995), sendo assim, apresentam correlação e medidos ao longo de um horizonte de análise. Problemas de heterogeneidade de variância e, ou a autocorrelação estão presentes com frequência nas bases de dados florestais (CALEGARIO; MAESTRI; et al., 2005), decorrente do aumento das idades dos povoamentos, a autocorrelação residual, associadas à remedição das variáveis ao longo do tempo (OLIVEIRA, 2012).

Considerando tais características, podem ocorrer dificuldades de ajustes, sobretudo pela violação das pressuposições estatísticas necessárias às estimativas por mínimos quadrados ordinários, como a homocedasticidade e independência dos resíduos (CARVALHO et al., 2011; OLIVEIRA, 2012).

Carvalho et al. (2011) afirma que tal complexidade de ajuste dos modelos, associada ao elevado número de dados justificam o uso de técnicas estatísticas mais apuradas, como o uso de modelos mistos e a modelagem da heterocedasticidade. Modelos de efeitos aleatórios são caracterizados por utilizarem tanto parâmetros de efeitos fixos como aleatórios em dados com medidas repetidas, seja na modelagem linear ou não linear, de acordo com a demanda do estudo (MÔRA, 2015).

Modelos mistos são utilizados nas diversas áreas da ciência na modelagem de processos longitudinais, tais como na medicina, engenharia, ecologia e agricultura e ciências sociais (BROWN; PRESCOTT, 2006; BARBOSA, 2009; ZUUR et al., 2009; AMADOR, 2010; COSTA, 2010).

Nas ciências florestais, Môra (2015) pesquisou modelos não lineares de efeitos aleatórios aplicada à funções de afilamento. Grégoire et al. (1995) realizaram estudos do crescimento em área basal de *Pinus scrobis* L. e *Pseudotsuga menziensis* utilizando modelos mistos. Calegario et al (2005) citam trabalhos nos quais a modelagem de efeitos aleatórios foi usada no crescimento de *Pinus sp*

(FANG, 1999), e na modelagem do perfil longitudinal de três espécies florestais (GARBER; MAGUIRE, 2003). Destaca-se ainda o uso de modelos lineares de efeitos aleatórios em multinível na estimativa de crescimento de povoamentos homogêneos de *Eucalyptus* em dois níveis de modelagem aleatória: unidades amostrais e árvores individuais dentro das unidades amostrais (CALEGARIO; MAESTRI; et al., 2005). Fu et al.(2013) desenvolveram um modelo misto não linear para crescimento da copa. Ercanli et al. (2015) desenvolveram modelos lineares e não lineares mistos para prever o diâmetro a 1,3 m de altura a partir do diâmetro do tronco. Outros autores pesquisaram e desenvolveram modelos não lineares mistos para funções de afilamento, Yang et al. (2009) e Arias-Rodil et al. (2015). Outras pesquisas envolvendo modelagem mista na área florestal podem ser citadas: (CRECENTE-CAMPO et al., 2010; ZHAO et al., 2013; TIMILSINA; STAUDHAMMER, 2013; XU et al., 2014; BOHORA; CAO, 2014; MEHTÄTALO et al., 2014; SOARES et al., 2015; DONG et al., 2015; VISMARA et al., 2015).

Entretanto, na medida em que os avanços computacionais ocorrem, a modelagem em diferentes níveis de modelos de crescimento e produção utilizando modelos mistos frente ao uso de modelos fixos, apresenta estimativas mais acuradas de volume para diferentes níveis?

1.1. OBJETIVOS

1.1.1. Objetivo geral

Comparar diferentes níveis para modelagem do crescimento e produção, quanto à acurácia e tendenciosidade das estimativas de produção em povoamentos de eucalipto.

1.1.2. Objetivos específicos

- Avaliar e comparar alternativas de modelagem de crescimento e produção de eucalipto, em nível de povoamento, com e sem efeitos aleatórios.
- Avaliar e comparar alternativas de modelagem da distribuição de diâmetros de povoamentos de eucalipto com e sem efeitos aleatórios.
- Avaliar alternativas de modelagem de povoamentos de eucalipto, em nível de árvores individuais, com e sem efeitos aleatórios.
- Avaliar a modelagem de crescimento e produção em diferentes níveis de abordagem para povoamentos de eucalipto, com e sem efeitos aleatórios.

1.2. Estrutura da Tese

De modo a atingir os objetivos, a tese foi estruturada em quatro capítulos:

Capítulo 1 – Modelagem do crescimento e produção de eucalipto, em nível de povoamento, com e sem efeito aleatório;

Capítulo 2 – Modelagem de crescimento e produção em nível de distribuição de diâmetros com e sem efeito aleatório;

Capítulo 3 – Modelagem do crescimento e produção de povoamentos de eucalipto, em nível de árvores individuais, com e sem efeito aleatório;

Capítulo 4 – Modelagem de crescimento e produção em diferentes níveis de abordagem para povoamentos de eucalipto, com e sem efeitos aleatórios.

2. REVISÃO DE LITERATURA

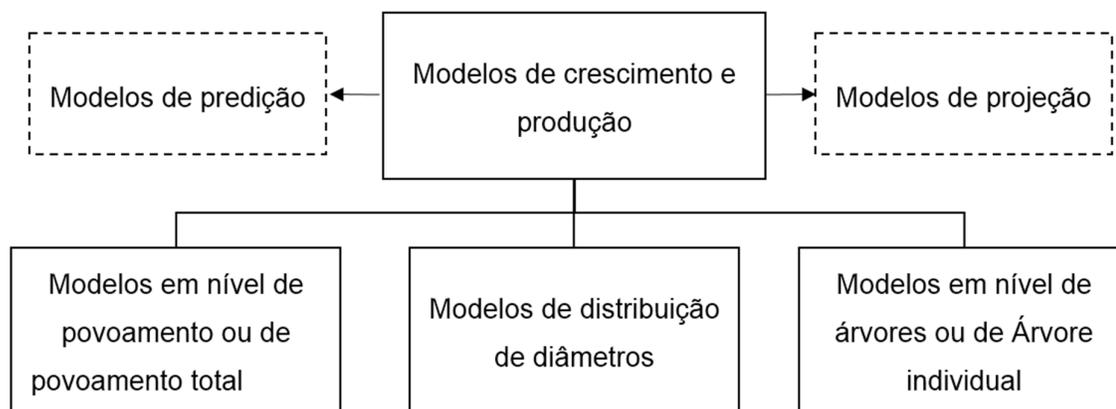
2.1. Modelos de Crescimento e produção

Os elementos principais do manejo de florestas equiâneas são: classificação das terras – quanto à sua localização, tamanho das árvores, estoques de crescimento e de colheita, espécies, solos, topografia e outros atributos; prescrições, como a rotação, regime de corte, preparo do solo, espaçamento, desbaste, método de colheita, entre outros, denominados tratamentos silviculturais. predição ou projeção do crescimento e da produção (CAMPOS; LEITE, 2013).

Sendo assim, para realizar o gerenciamento do planejamento florestal é preciso compreender os processos de crescimento e produção, para que o manejo possa ser decidido principalmente em função da prognose de produções futuras com base em informações atuais. Neste sentido, um modelo de crescimento e produção pode ser constituído por uma ou mais equações, formando um sistema, incluindo uma ou mais variáveis independentes, sendo empregados para estimarem o crescimento (diâmetro a 1,3 m de altura, altura total), e a produção (área basal e volume).

Uma classificação dos modelos crescimento e produção é apresentada na Figura 1, de acordo com sua estrutura e aplicação.

Figura 1 – Tipos de modelos de crescimento e produção quanto a sua estrutura e aplicação.



Fonte: Adaptado de Campos e Leite (2013).

Modelos de povoamento total são aqueles que consideram a produção por unidade de área. Esses modelos são tipicamente inflexíveis para avaliar multiprodutos e grande número de tratamentos, e são limitados a predizer apenas a

produção total (BURKHART, 1977). Como exemplo de aplicações desse tipo de modelo, pode-se citar os trabalhos de Trevizol Júnior (1985), Couto et al. (1989), Campos (1996), Silva et al. (2003), Soares et al. (2004), Cruz et al. (2008), Mendonça (2010), Salles et al. (2012).

Modelos de distribuição de diâmetros permitem estimar a produção por classe de diâmetro, possibilitando uma análise de produtividade em relação às dimensões das árvores do povoamento. Nos modelos de árvore individual árvore é considerada como unidade básica na modelagem do crescimento do povoamento (MARTINS, 2011).

Do ponto de vista de aplicação, tais modelos podem ser usados para predição e para projeção. No primeiro caso, parâmetros de uma equação ou de um sistema de equações são correlacionados com atributos do povoamento na idade atual, tais como idade, altura, diâmetro, área basal, mortalidade e ingresso, de acordo com o nível pretendido. Na projeção os parâmetros são estimados em uma idade futura em função dos mesmos atributos e dos parâmetros obtidos em uma idade atual (LEITE et al., 2013).

Modelos de distribuição de diâmetros não são tão onerosos como os modelos de árvores individuais e nem apresentam flexibilidade limitada como os modelos de povoamento total. Seu objetivo é fornecer informações sobre a estrutura do povoamento, permitindo assim a estimativa da produção por classes de diâmetro. Com tais informações é possível realizar o planejamento do corte, de desbastes, os tratamentos silviculturais, estimar multiprodutos como por exemplo, madeira para serraria, celulose e energia (LEITE, 1990).

Existem diversas maneiras de aplicar esse tipo de modelagem, na qual uma das mais comuns são as distribuições de probabilidade para caracterizar a distribuição dos diâmetros do povoamento florestal.

Vários estudos na literatura podem ser encontrados para avaliar metodologias científicas da aplicação deste tipo de modelo, testando diferentes distribuições de probabilidade, métodos de ajuste e de recuperação dos parâmetros estimados, como as obras de Binoti et al. (2011), Guimarães (1994), Oliveira Neto (2010), Santos (2012), Schneider et al. (2008), Soares et al. (2007).

Uma alternativa para estimar a produção por classe de diâmetros e multiprodutos são os modelos de árvore individual, que diferem do modelo de povoamento total pela necessidade de especificar o tamanho de cada árvore dentro do povoamento, em relação ao seu crescimento em diâmetro, altura, e, em alguns

casos, do tamanho da copa e sua posição no dossel (VANCLAY, 1994). Tais modelos podem ser muito detalhados e onerosos em termos de obtenção de dados, podendo ser de distância dependente e independente (MUNRO, 1974).

Ao contrário dos modelos de povoamento total, os modelos de árvore individual são menos utilizados no Brasil, por sua complexidade, uma vez que são constituídos por vários submodelos, e apresentam variáveis de difícil medição e avaliação, como aquelas que dizem respeito à mortalidade e às dimensões das copas das árvores (MARTINS, 2011). Portanto, são modelos onerosos e complexos na sua aplicação, mesmo que tenham a perspectiva de oferecer resultados mais fidedignos. Maiores detalhes de como aplicar este nível de modelagem podem ser encontrados nos trabalhos de Vanclay (1994), Hasenauer (2000), Martins (2011).

2.1.1. Modelos em nível de Povoamento

Incluem os modelos que exprimem a produção de madeira em termos de unidade de área, usando atributos em nível de povoamento (variáveis independentes), como idade, área basal e índice de local e outras, pelas quais diferem entre si (GUIMARÃES, 1994). Sendo Y a variável dependente, eles podem ser subdivididos em (CAMPOS; LEITE, 2013):

- a) $Y = f(I)$;
- b) $Y = f(I, S)$; e
- c) $Y = f(I, S \text{ ou } Hd, B, N \text{ ou } q)$,

em que: I é a idade do povoamento; S é o Índice de local; Hd é a altura dominante; B é a Área basal; N é o número de árvores por hectare; e q é o diâmetro médio.

Espera-se que, ao incluir mais variáveis independentes ao modelo, estimativas mais consistentes possam ser obtidas. Alguns destes modelos são apresentados na Tabela 1.

Tabela 1 – Principais modelos em nível de povoamento

Relação funcional	Modelo
$V = e^{\beta_0 + \beta_1 \left(\frac{1}{l}\right)} + \varepsilon$	Schumacher
$V = \frac{\beta_0}{1 + \beta_1 e^{-\beta_2 l}} + \varepsilon$	Logístico
$V = \beta_0 + e^{-e^{\beta_1 - \beta_2 l}} + \varepsilon$	Gompertz
$V = \frac{\beta_0}{(1 + \beta_1 e^{-\beta_2 l})^{1/\beta_3}} + \varepsilon$	Richards
$\text{Ln}B_2 = \text{Ln}B_1 \left(\frac{l_1}{l_2}\right) + \alpha_0 \left(1 - \frac{l_1}{l_2}\right) + \alpha_1 \left(1 - \frac{l_1}{l_2}\right) S + \varepsilon$ $\text{Ln}V_2 = \beta_0 + \frac{\beta_1}{l_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon$	Clutter (1963)

B_2 é a área basal do povoamento na idade futura ($\text{m}^2 \text{ha}^{-1}$); B_1 é a área basal do povoamento na idade atual; l_1 é a idade atual, em meses; l_2 , a idade futura; S é o índice de local; V_2 é o volume de madeira com casca do povoamento, na idade futura ($\text{m}^3 \text{ha}^{-1}$); com α_1 e α_2 , que são os coeficientes do modelo de área basal; $\beta_1 \dots \beta_4$ são os coeficientes do modelo volumétrico; e ε é o erro aleatório.

Oliveira (2012) destaca o modelo de Clutter (1963) dentre as principais relações funcionais utilizadas em estudos de crescimento e produção em nível de povoamento, embora tenha sido utilizado inicialmente em povoamentos de *Pinus taeda*, nos Estados Unidos, tem sido muito utilizado no Brasil para diferentes genótipos.

Considera-se que Clutter é um modelo em nível de povoamento por permitir que independente da classe de tamanho da árvore, a variável estimada é o volume por unidade de área. É considerado explícito já que a produção em volume é calculada diretamente e apresenta a propriedade de consistência, pois estimativas podem ser obtidas pela projeção da área basal ano a ano, ou diretamente de um para qualquer outro ano, sem que haja necessidade de intervalos regulares (MAGALHÃES, 2014).

Campos e Leite (2013) cita que as relações funcionais $Y = f(l)$ devem ser evitadas em áreas extensas e heterogêneas, pela possibilidade de ocorrerem diferentes tendências de crescimento. Já relações $Y = f(l, S)$ podem resultar em estimativas consistentes, permitindo considerar o efeito da capacidade produtiva, desde que existam curvas de índices de local consistentes. Por último, relações funcionais $Y = f(l, S \text{ ou } Hd, B, N \text{ ou } q)$, podem ser aplicados em áreas heterogêneas, sendo o crescimento, a produção e a idade técnica de colheita diferenciados por índice de local e densidade inicial.

É importante salientar que apenas as considerações dessas variáveis independentes não é garantia de consistência dos efeitos, sendo necessário considerar o ajuste dos dados aos modelos apropriados.

2.1.2. Modelos de Distribuição de diâmetros

Nos modelos de distribuição de diâmetros são realizadas estimativas do número de árvores em classes (frequência) de diâmetro por hectare, nas idades presentes e futura, proporcionando maior detalhamento da produção em relação ao modelo de povoamento total (LEITE, 1990).

A descrição da estrutura diamétrica de uma floresta ou de uma espécie pode ser realizada de maneira eficiente por meio do emprego de funções densidade de probabilidade (fdp.) (MACHADO et al., 2009; CAMPOS; LEITE, 2013).

Para cada fdp existe a função de distribuição acumulada $F(x)$ que representa sua integral, cuja área delimitada entre o intervalo $[a, b]$, encontrada sob a curva define a probabilidade das árvores estarem contidas nela, obtido pela função de distribuição acumulativa (CAMPOS; LEITE, 2013).

Sendo o diâmetro uma variável contínua, é importante apresentar propriedades básicas das funções de densidade de probabilidade de tal modo se x satisfaz as seguintes condições:

$f(x) \geq 0$ para todos os valores de x dentro do intervalo considerado, ou seja, $a < x < b$;

$$\int_a^b f(x)dx = 1;$$

Sendo assim, a fdp não considera valores pontuais, e sim com intervalos infinitesimais - ela informa a probabilidade de a variável X assumir um valor naquele intervalo. Assim a probabilidade de x ocorrer num intervalo depende da integral:

$$P = (a \leq x \leq b) = \int_a^b f(x) dx$$

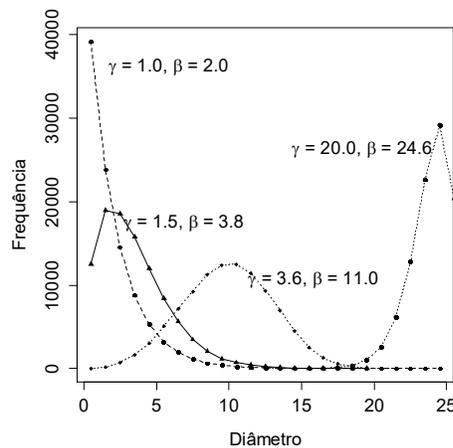
Já a função de distribuição acumulada $F(x)$ é o resultado da integral de $f(x)$ e estima a probabilidade que a variável x assumira um valor menor ou igual a x .

$$F(x) = \int_a^b f(x) dx, \text{ cujas propriedades são:}$$

- i. Não é decrescente;
- ii. Quando x tende a mais infinito $f(x) = 1$;
- iii. Quando x tende a menos infinito $f(x) = 0$;
- iv. $P(a \leq x \leq b) = F(b) - F(a)$ para $a < b$

Diversos estudos foram feitos para explicar a estrutura dos povoamentos empregando diferentes funções de probabilidade, aplicando as funções de densidade S_B de Johnson, Normal, Lognormal, Gama e Beta (PLÁCIDO et al., 2004; EISFELD et al., 2005; BINOTI et al., 2011; DALLA LANA et al., 2013). Dentre estas, a função Weibull se destaca como a mais utilizada com o propósito de construir modelos de distribuição de diâmetros (BAILEY; DELL, 1973; BATISTA, 1989; NEWTON et al., 2005; SCHNEIDER et al., 2008; BINOTI et al., 2010, 2011; FIGURA, 2010; SOARES et al., 2011; GORGOSO et al., 2012; GORGOSO-VARELA; ROJO-ALBORECA, 2014; DIAMANTOPOULOU et al., 2015). A ampla utilização da função Weibull pode ser explicada por sua flexibilidade em descrever várias formas conforme pode ser observado na Figura 2.

Figura 2 – Função de Densidade de Probabilidade Weibull, para diferentes valores de γ (forma) e β (escala).



Fonte: Adaptado de Leite (1990).

Na Tabela 2 são apresentadas distribuições Weibull com suas respectivas funções de densidade e probabilidade.

Tabela 2 – Diferentes funções da distribuição Weibull empregadas nos modelos de distribuição de diâmetros

Distribuição	Função de Densidade probabilística
Weibull 3 parâmetros (3P) completa	$f(x) = \left(\frac{\gamma}{\beta}\right) \left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{x-\alpha}{\beta}\right)^\gamma\right]$ $\alpha \leq x < \infty, \alpha \geq 0, \beta > 0 \text{ e } \gamma > 0$
Weibull 2 parâmetros (2P) completa	$f(x) = \left(\frac{\gamma}{\beta}\right) \left(\frac{x}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{x}{\beta}\right)^\gamma\right]$ $x \geq 0, \beta > 0 \text{ e } \gamma > 0$
Weibull 2 parâmetros (2P) truncada à esquerda	$f(x) = \left(\frac{\gamma}{\beta}\right) \left(\frac{x}{\beta}\right)^{\gamma-1} \exp\left[\left(\frac{t}{\beta}\right)^\gamma - \left(\frac{x}{\beta}\right)^\gamma\right]$
Weibull 2 parâmetros (2P) truncada à direita	$f(x) = \frac{\left(\frac{\gamma}{\beta}\right) \left(\frac{x}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{x}{\beta}\right)^\gamma\right]}{1 - \exp\left[-\left(\frac{t}{\beta}\right)^\gamma\right]}$
Weibull 2 parâmetros (2P) truncada à direita e à esquerda	$f(x) = \frac{\left(\frac{\gamma}{\beta}\right) \left(\frac{x}{\beta}\right)^{\gamma-1} \exp\left[\left(\frac{t}{\beta}\right)^\gamma - \left(\frac{x}{\beta}\right)^\gamma\right]}{1 - \exp\left[-\left(\frac{t}{\beta}\right)^\gamma\right]}$
Weibull 3 parâmetros (3P) truncada à esquerda	$f(x) = \left(\frac{\gamma}{\beta}\right) \left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} \exp\left[\left(\frac{t-\alpha}{\beta}\right)^\gamma - \left(\frac{x-\alpha}{\beta}\right)^\gamma\right]$

$$\begin{aligned}
 & \text{Weibull 3 parâmetros (3P)} \\
 & \text{truncada à direita} \\
 & f(x) = \frac{\left(\frac{\gamma}{\beta}\right)\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{x-\alpha}{\beta}\right)^\gamma\right]}{1 - \exp\left[-\left(\frac{t-\alpha}{\beta}\right)^\gamma\right]} \\
 \\
 & \text{Weibull 3 parâmetros (3P)} \\
 & \text{truncada à direita e à esquerda} \\
 & f(x) = \frac{\left(\frac{\gamma}{\beta}\right)\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} \exp\left[\left(\frac{t-\alpha}{\beta}\right)^\gamma - \left(\frac{x-\alpha}{\beta}\right)^\gamma\right]}{1 - \exp\left[-\left(\frac{t-\alpha}{\beta}\right)^\gamma\right]}
 \end{aligned}$$

em que: α é o parâmetro de locação; β é o parâmetro de escala; γ é o parâmetro de forma; x é a variável de interesse, no caso o diâmetro; t é o ponto de truncamento;

Na função Weibull, o parâmetro de locação (α) diz respeito a posição da curva sobre o eixo das abscissas, o parâmetro escala (β) controla as dimensões que a curva assume, e finalmente o parâmetro forma (γ) define as diferentes inclinações da curva.

O parâmetro de locação tem sido considerado como o diâmetro mínimo do povoamento, motivo pelo qual é eliminado da função, assumindo o valor inicial desde a origem dos dados (FIGURA, 2010). Tal abordagem é utilizada uma vez que o parâmetro de locação apresenta baixa correlação com atributos do povoamento e sua exclusão diminui a complexidade para realizar a projeção dos parâmetros de forma e de escala (BINOTI et al., 2010).

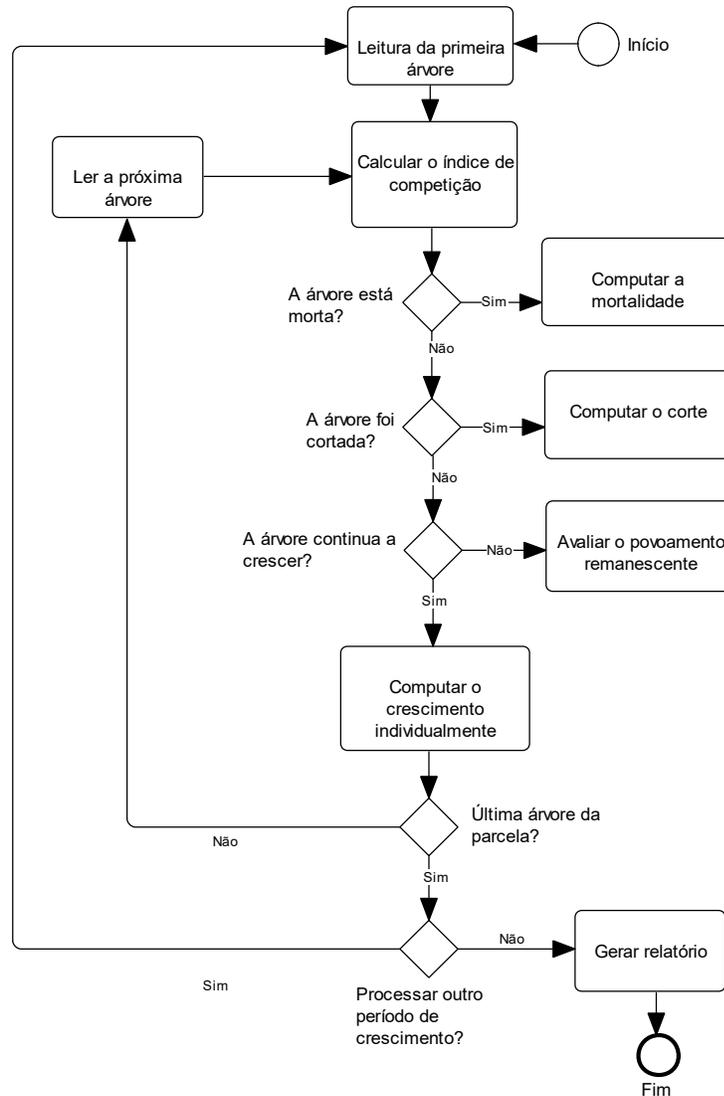
2.1.3. Modelos de Árvore Individual

Modelos em nível de árvore individual simulam o crescimento de cada árvore, em diâmetro, altura e copa, simulando a mortalidade, calculando o crescimento e o volume, de forma individual. Fornecem informações detalhadas sobre a dinâmica e a estrutura de um povoamento florestal, como a distribuição de volume por classe de tamanho, agrupadas por unidade de área, ou não (MARTINS, 2011; CAMPOS; LEITE, 2013).

São três os submodelos componentes deste nível de modelo: probabilidade de mortalidade, regeneração ou ingresso e crescimento dimensional (altura e diâmetro).

O fluxograma apresentado na Figura 3 representa o esquema geral do modelo de árvore individual utilizado por Martins (2011).

Figura 3 – Fluxograma do modelo de árvore individual



Fonte: Adaptado de Martins (2011) e Castro et al. (2013)

Dentre todas as árvores do povoamento, o submodelo de mortalidade simula as árvores mortas em cada idade, e as remove do conjunto e para as árvores vivas simula o crescimento dimensional em diâmetro e altura, realizando o cálculo do seu crescimento e do volume (MARTINS, 2011). A mortalidade pode ser classificada em dois tipos: regular e irregular, sendo a primeira ocasionada por fatores não catastróficos como competição, fatores genéticos e senescência, a qual pode ser mensurada. Já a mortalidade irregular ocorre por causas não intrínsecas à árvore,

como incêndios, ataques de pragas, danos mecânicos, portanto difícil de ser prevista e mensurada. Sendo assim a mortalidade é um aspecto essencial, no entanto, de difícil avaliação e modelagem (CASTRO et al., 2013).

Dentre os diferentes tipos de métodos para estimar a mortalidade do tipo regular, destacam-se os que utilizam índices de competição; predição da árvore morrer em um determinado período de tempo por condições fisiológicas do indivíduo; estimação empírica da mortalidade a partir da ocorrência de limites críticos que possibilitem sua previsão; e cálculo da probabilidade da mortalidade da árvore de um grupo de indivíduos com similaridade de características (MARTINS, 2011).

A regeneração ou ingresso representa outro submodelo, e diz respeito ao número de indivíduos que medidos em uma idade i não foram medidos em idades anteriores, por não apresentarem tamanho mínimo de acordo com o determinado.

O crescimento dimensional das árvores é o terceiro submodelo, que pode ser abordado de duas maneiras: 1) utilizando funções de crescimento; 2) utilizando funções modificadoras de crescimento potencial.

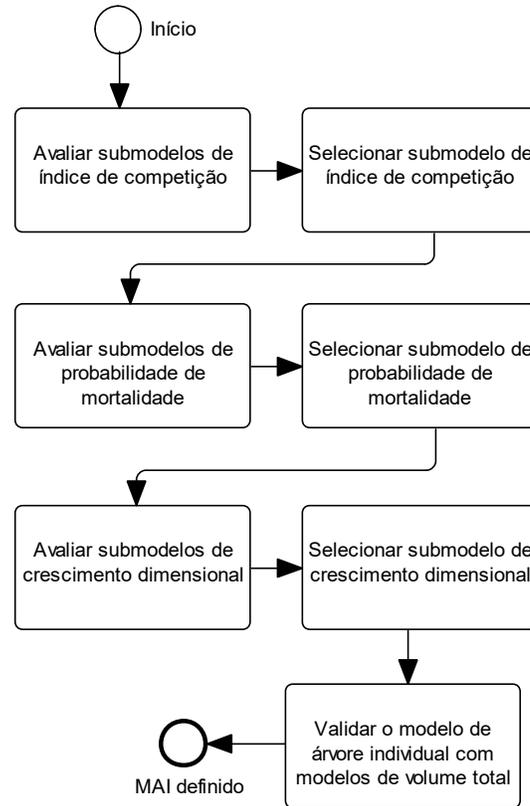
A estimativa do crescimento dimensional pode ser realizado para uma idade futura a partir da idade atual e de outras como o índice de competição, idade, índice de local (CASTRO, 2011). É possível ainda realizar a modelagem do crescimento por meio do crescimento potencial, em geral avaliado pelas árvores em crescimento sem o efeito da competição ou árvores dominantes. O crescimento das árvores em concorrência é determinado por uma função modificadora do crescimento potencial, que tem o propósito de penalizar os efeitos da competição sobre o crescimento, utilizando índices de competição e variáveis de povoamento para reduzir o crescimento das árvores submetidas à competição (CASTRO, 2011).

Munro (1974) citado por Martins (2011) propôs uma categorização para os modelos de arvores individuais, segundo a maneira na qual a competição entre as árvores é considerada: modelos dependentes de distância e independentes de distância.

Nos modelos dependentes de distância para cada árvore i , utiliza-se a localização das árvores vizinhas ou competidoras, variáveis do povoamento, suas dimensões iniciais para descrever a competição, determinando como consequência, seu crescimento ou mortalidade. A segunda categoria, a localização não é considerada (CASTRO, 2011).

O fluxograma na Figura 4 representa o procedimento necessário para a obtenção de um modelo de árvore individual.

Figura 4 – Determinação do modelo de árvores individuais



Fonte: O autor

2.2. Modelos mistos

Nos modelos de regressão linear múltipla procura-se prever uma variável dependente Y , assumindo sua relação linear com mais de uma variável independente, ou seja, inúmeras variáveis independentes k cujo modelo da i -ésima observação pode ser expressa como

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad (1)$$

Onde $i = 1, 2, \dots, n$.

As pressuposições estabelecidas para os modelos lineares múltiplos são análogas àquelas da regressão linear simples.

Escrevendo (1) para cada uma das n observações, tem-se

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \dots + \beta_k X_{1k} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \dots + \beta_k X_{2k} + \varepsilon_2 \\
&\vdots \\
Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \dots + \beta_k X_{nk} + \varepsilon_n
\end{aligned} \tag{2}$$

Tais equações podem ser escritas matricialmente como

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{3}$$

Ou ainda na forma matricial

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{4}$$

em que \mathbf{Y} é um vetor $n_i \times 1$ referente à variável resposta (dependente), \mathbf{X} é uma matriz $n_i \times p$ de covariáveis observadas (variáveis independentes), $\boldsymbol{\beta}$ é um vetor $p \times 1$ de parâmetros fixos desconhecidos, $\boldsymbol{\varepsilon}$ é um vetor $n_i \times 1$ de erros aleatórios e $p = k + 1$

Modelos lineares como o descrito na expressão (4) apresentam apenas fatores fixos, além do erro experimental (ε), que é sempre aleatório, são denominados modelos fixos (CAMARINHA FILHO, 2003). Tais modelos consideram medidas em apenas um (1) nível, de acordo com a unidade experimental considerada (*individual-level analysis*). Modelos apenas com fatores aleatórios são denominados modelos aleatórios, contudo, não faz parte deste estudo apresentá-los.

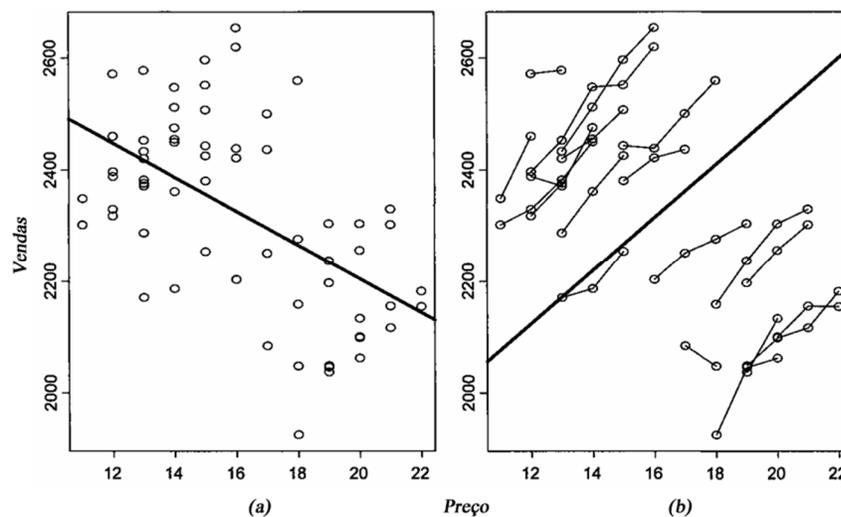
Modelos que apresentam em sua composição tanto efeitos fixos, como também incorporam em sua estrutura efeitos aleatórios (além do erro experimental) são denominados modelos mistos (NUNES, 2010).

Os efeitos aleatórios são úteis para modelar a variabilidade que não é explicada pelas covariáveis fixas do modelo de regressão, muitas vezes causada pela falta de outra covariável não incluída, como também pela variabilidade existente entre os diferentes níveis de agrupamento dos dados (NUNES, 2010).

Considere um exemplo apresentado por Demidenko (2013) para transição conceitual da modelagem de efeitos fixos para a modelagem de efeitos aleatórios, com uma regressão linear simples.

Na abordagem de efeitos aleatórios, considera-se que os dados sejam agrupados, e desta forma, uma das variáveis presentes na base de dados identifica subgrupos nos dados apresentados. Para ilustrar, considere uma relação linear simples entre o preço (x) e vendas de mercadorias (y), com $\{(x_k, y_k), k = 1, \dots, K\}$ como sendo a amostra de observações coletadas do preço e vendas de diversas mercadorias. No gráfico apresentado na Figura 5 (a) de y versus x observa-se um relacionamento linear com uma inclinação negativa.

Figura 5 – Tendência de regressão clássica (a) e de modelo misto (b).



Fonte: Adaptado de (DEMIDENKO, 2013).

Na estatística clássica, assume-se que os pares (x_k, y_k) são independentes e identicamente distribuídos (IID), com a linha de regressão $E(y | x) = \alpha + \beta x$. No entanto, considerando-se que os dados sejam agrupados, em que cada grupo é uma mercadoria. Na Figura 5 (b), observamos que os pontos de cada mercadoria estão conectados, e assim temos uma análise diferente daquela obtida no modelo clássico de regressão: o aumento dos preços de cada mercadoria poderia aumentar suas vendas.

O modelo assumido na abordagem clássica foi:

$$y_k = \alpha + \beta x_k + \varepsilon_k \quad (5)$$

em que $k = 1, \dots, K$, e ε_k são variáveis aleatórias independentes e identicamente distribuídas com média zero e variância constante σ^2 .

Neste sentido, presume-se que os dados foram obtidos de mercadorias semelhantes, homogêneas. Como apresentando em (b), as mercadorias não são homogêneas e variam consideravelmente no preço e vendas. Então, um modelo adequado seria assumir que cada mercadoria tem suas vendas específicas, ou seja, um intercepto específico:

$$y_{ij} = \alpha_i + \beta x_{ij} + \varepsilon_{ij} \quad (6)$$

em que $i = 1, \dots, N$, $j = 1, \dots, n_i$.

A indexação ij foi adotada, pois os dados estão agrupados, onde i corresponde à *iésima* mercadoria, j corresponde à *jésima* observação da *iésima* mercadoria, com n_i é o número de observações para a *iésima* mercadoria, e α_i são os interceptos específicos de cada mercadoria. O número total de observações é

$$K = \sum_{i=1}^N n_i.$$

O pressuposto central do modelo de efeitos aleatórios é que o intercepto $\{\alpha_i, i = 1, \dots, N\}$ são aleatórios e pertence a uma população em geral que pode ser expressa como:

$$\alpha_i = \alpha + b_i \quad (7)$$

em que α é o intercepto e b_i é o efeito aleatório da venda específica de cada mercadoria, dentre as vendas da população.

Com isso, vendas específicas por mercadorias são consideradas, mas, assumimos que as mercadorias representam a economia da região e, portanto, um aumento no preço afeta as vendas em todas as mercadorias.

Juntos, os modelos (6) e (7) definem um modelo linear misto, no qual os parâmetros α e β são efeitos fixos (parâmetros médios da população), e b_i é o efeito aleatório com média zero e variância σ_b^2 , independente de ε_{ij} , que combinados, formam o modelo:

$$y_{ij} = \alpha_i + \beta x_{ij} + \eta_{ij} \quad (8)$$

em que $\eta_{ij} = \varepsilon_{ij} + b_i$ é o erro aleatório composto, $i = 1, \dots, N$, $j = 1, \dots, n_i$,

Os modelos lineares mistos são flexíveis para permitir que o coeficiente angular (inclinação) e o intercepto variem entre os indivíduos, possuindo dois componentes: um intra-individual e outro entre indivíduos (variação no intercepto e inclinação individual). Neste tipo de modelo, as medidas dos indivíduos não precisam ser igualmente espaçadas e balanceadas e as análises podem ser conduzidas com os dados de indivíduos que tiveram medidas perdidas (parcelas que não foram medidas, árvores que não foram medidas dentro de parcelas, por exemplo) ou que apresentam ausência de informação em algum momento do estudo. Portanto, na ausência de dados, os modelos mistos oferecem mais uma vantagem de utilização frente aos outros tipos de modelos para dados longitudinais e de medidas repetidas (QUEIROZ, 2012).

O uso de modelos de efeito aleatório envolve três etapas: (i) a estimativa dos parâmetros do modelo; (ii) predição dos efeitos aleatórios, e (iii) realização de predições específicas para cada indivíduo. A primeira etapa corresponde ao ajuste do modelo, enquanto a segunda é definida também como calibração, e a terceira, predição específica para o indivíduo, ou *subject-specific prediction* (SS) (ARIAS-RODIL et al., 2015). Vismara et al. (2015) explica que calibração de um modelo misto pode ser realizada para um determinado grupo de dados em uma estrutura hierarquizada, sendo assim, os parâmetros do modelo podem variar entre estes grupos.

Modelos mistos permitem identificar as diferentes fontes estocásticas de variabilidade que não são explicadas pela parte fixa do modelo, dividindo-se a variância residual em diferentes componentes (CALAMA; MONTERO, 2005). Nesta abordagem, os parâmetros de um modelo ajustado podem variar entre os grupos, sendo assim, o modelo pode ser calibrado para um determinado grupo afim de melhorar a qualidade das predições dentro do referido grupo (VISMARA et al., 2015).

Para Pinheiro e Bates (2006), os modelos lineares mistos têm sido um tema crescente na estatística nos últimos 50 anos, já que permite a modelagem de correlação intra-indivíduo, muitas vezes presente em dados agrupados, construindo assim uma ferramenta conveniente para modelagem dessa dependência.

2.2.1. Formulação

O modelo linear misto é expresso na forma matricial por:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon} \quad (9)$$

em que: \mathbf{Y} é um vetor $n \times 1$ referente à variável resposta (dependente); \mathbf{X} uma matriz $n \times p$ de especificação ou covariáveis observadas (variáveis independentes) de efeitos fixos; $\boldsymbol{\beta}$ um vetor $p \times 1$ de parâmetros fixos desconhecidos; \mathbf{Z} uma matriz $n \times k$ de especificação (conhecida e de posto completo) dos efeitos aleatórios; \mathbf{b} um vetor $q \times 1$ de efeitos aleatórios com vetor de média zero; $\boldsymbol{\varepsilon}$ um vetor $n \times 1$ de erros aleatórios com vetor de média 0; \mathbf{G} uma matriz $k \times k$ de covariância e variância dos efeitos aleatórios; \mathbf{R} uma matriz $n \times n$ de covariância residual $\text{cov}(\boldsymbol{\varepsilon}) = \mathbf{R}$, n número de observações; q o número de variáveis aleatórias definidas no modelo; p o número de variáveis fixas observadas (covariáveis fixas do modelo); h o número de elementos únicos (grupos) identificados na variável definida como nível um do modelo misto; $k = q \times h$ o número de colunas da matriz \mathbf{Z} .

A matriz \mathbf{Z} pode conter quaisquer covariáveis aleatórias que se relacionam com o modelo especificado.

De (9) tem-se ainda que:

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} \quad (10)$$

$$\text{var}(\mathbf{Y}) = \mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R} \quad (11)$$

em que: $\mathbf{V}_{n \times n}$ é uma matriz é a matriz de covariâncias de \mathbf{Y} , que permite a modelagem de diversas estruturas de covariância para \mathbf{R} . Salienta-se que quando $\mathbf{R} = \sigma^2\mathbf{I}$ e $\mathbf{Z} = 0$, o modelo misto passa a ser um modelo linear clássico (GOUVEIA, 2013).

Bates et al (2014) definem a matriz $\mathbf{Z}_{n \times k}$ a partir das matrizes $\mathbf{W}_{n \times h}$ e $\mathbf{A}_{n \times q}$, sendo $\mathbf{W}_{n \times h}$ a matriz formada para identificar o posicionamento dos elementos únicos (grupo) no conjunto de dados do nível 1 e $\mathbf{A}_{n \times q}$ a matriz com as covariáveis do modelo aleatório. A construção da $\mathbf{Z}_{n \times k}$ é realizada produto Khatri-Rao (BATES et al., 2014) definido em seguida.

$$\mathbf{Z} = (\mathbf{W}' \odot \mathbf{A}')' = \begin{bmatrix} \mathbf{W}'_1 \otimes \mathbf{A}'_1 \\ \mathbf{W}'_2 \otimes \mathbf{A}'_2 \\ \vdots \\ \mathbf{W}'_n \otimes \mathbf{A}'_n \end{bmatrix} \quad (12)$$

em que: \otimes é o produto direto ou *Kronecher*.

Em seguida é apresentado um exemplo demonstrando a formulação das matrizes a partir de dados de inventário florestal para dois projetos apresentados na Tabela 3. Destaca-se que o objetivo do exemplo é de ilustrar a formulação matricial de um modelo linear misto, e não realizar uma análise dos resultados, uma vez que a amostra não apresenta variabilidade suficiente.

Tabela 3 – Dados de inventário florestal.

Projeto	Parcela	Árvore	Ano de Medição	Idade	HT	DAP	Classe de local
1	1	1	3	36	12,30	12,6	23,5
1	1	1	4	48	20,50	15,9	23,5
1	1	1	5	60	21,30	17,2	23,5
1	2	1	3	36	14,77	11,9	30,5
1	2	1	4	48	20,61	14,8	30,5
1	2	1	5	60	24,16	16,0	30,5
2	1	1	1	24	19,20	14,3	37,5
2	1	1	2	36	25,90	17,7	37,5
2	1	1	3	48	30,10	20,3	37,5

Fonte: O autor

Destaca-se que dados de inventário florestal são caracterizados por serem desbalanceados (GRÉGOIRE et al., 1995), cujas medidas são obtidas em momentos diferentes para as parcelas e/ou número diferente de medidas repetidas por parcela. Observa-se na Tabela 3 que o projeto 1 consiste de 2 parcelas, e o projeto 2 de apenas uma parcela, com medições realizadas em períodos distintos. Para obter o modelo linear misto, considere como exemplo um modelo hipsométrico a partir do modelo linear simples apresentado em (13).

$$HT = \beta_0 + \beta_1 DAP + \varepsilon_i \quad (13)$$

em que: HT é a altura total, DAP é o diâmetro a 1,3 m de altura e β_0 e β_1 coeficientes de regressão (fixos).

O exemplo 1 considera apenas o intercepto aleatório para a variável parcela (é calibrado para o grupo parcela), ou seja, parcela como grupo de efeitos aleatórios, conforme apresentando em (14).

$$HT = \beta_0 + \beta_1 DAP + b_0 + \varepsilon_i \quad (14)$$

em que: b_0 é o coeficiente de regressão do modelo aleatório.

Sendo assim, espera-se que o modelo considere a variação presente no nível do indivíduo ε_i como também aquela proveniente do coeficiente aleatório b_{0i} no nível da parcela.

O vetor \mathbf{Y} da variável dependente HT e a matriz \mathbf{X} da parte fixa do modelo apresentado em (14) com os dados da Tabela 3 são dadas por

$$\mathbf{Y} = \begin{bmatrix} 12,30 \\ 20,50 \\ 21,30 \\ 14,77 \\ 20,61 \\ 24,16 \\ 19,20 \\ 25,90 \\ 30,10 \end{bmatrix}, \text{ e } \mathbf{X} = \begin{bmatrix} 1 & 12,6 \\ 1 & 15,9 \\ 1 & 17,2 \\ 1 & 11,9 \\ 1 & 14,8 \\ 1 & 16,0 \\ 1 & 14,3 \\ 1 & 17,7 \\ 1 & 20,3 \end{bmatrix}$$

Ao formular a matriz \mathbf{Z} faz-se necessário definir a matriz $\mathbf{W}'_{3 \times 9}$, onde $n = 9$, correspondente ao número de observações e 3 é o número de parcelas apresentadas na Tabela 3.

$$\mathbf{W}' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Na especificação do modelo apresentado em (14) é definido apenas o intercepto aleatório, sem que sejam especificadas as covariáveis em sua formulação aleatória, sendo assim define-se o vetor $\mathbf{A}_{9 \times 1} = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$. Realizando o produto Khatri-Rao (BATES et al., 2014) definido em (12), obtém-se a matriz

$$\mathbf{Z}_{9 \times 3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \text{ sendo a primeiras e segundas colunas correspondentes às}$$

parcelas 1 e 2 do projeto 1, e a coluna 3 correspondente a parcela 1 do projeto 2.

Destaca-se que o número de colunas é definido por $k=1 \times 3$, onde 1 é o número de covariáveis do modelo aleatório, neste caso 1 apenas o intercepto e 3 a quantidade de parcelas dispostas nos dados utilizados no exemplo. Os demais componentes do modelo mistos serão apresentados adiante na seção que trata a estimação.

Para ilustrar a formulação de um modelo misto com uma covariável aleatória representando a inclinação, considere o modelo (15) no qual o índice de local é considerado.

$$HT = \beta_0 + \beta_1 DAP + b_0 + b_1 S + \varepsilon_i \quad (15)$$

em que: S é a covariável aleatória: índice de local.

Define-se o exemplo 2 no qual o vetor \mathbf{Y} da variável dependente HT, as

matrizes \mathbf{X} e $\mathbf{W}'_{3 \times 9}$ são os mesmos do exemplo 1, temos a matriz $\mathbf{A}_{9 \times 2} = \begin{bmatrix} 1 & 23,5 \\ 1 & 23,5 \\ 1 & 23,5 \\ 1 & 30,5 \\ 1 & 30,5 \\ 1 & 30,5 \\ 1 & 37,5 \\ 1 & 37,5 \\ 1 & 37,5 \end{bmatrix}$ e a

$$\text{matriz } \mathbf{Z}_{9 \times 6} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 23,5 \\ 0 & 0 & 0 & 0 & 1 & 23,5 \\ 0 & 0 & 0 & 0 & 1 & 23,5 \\ 0 & 0 & 1 & 30,5 & 0 & 0 \\ 0 & 0 & 1 & 30,5 & 0 & 0 \\ 0 & 0 & 1 & 30,5 & 0 & 0 \\ 1 & 37,5 & 0 & 0 & 0 & 0 \\ 1 & 37,5 & 0 & 0 & 0 & 0 \\ 1 & 37,5 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ em } k = 2 \times 3, \text{ sendo 2 o número de}$$

covariáveis do modelo aleatório (15).

2.2.2. Estimação

Dentre os problemas no ajuste de modelos mistos, dado um conjunto de observações, destaca-se a estimação dos componentes de variâncias e covariâncias dos efeitos aleatórios, e a definição da metodologia adequada para a estimação dos efeitos fixos. Camarinha Filho (2002) argumenta que a seleção de um modelo está ligada à possibilidade de realizar estimação dos parâmetros e a predição do comportamento das variáveis aleatórias para um dado conjunto de observações, uma vez que variáveis aleatórias só podem ser preditas, não estimadas, uma vez que possuem valor fixo. Sendo assim, no elemento fixo do modelo devem-se estimar os próprios efeitos fixos, enquanto que nos aleatórios realiza-se a estimativa dos componentes de variância, em seguida a predição das variáveis aleatórias.

Em seu trabalho, Camarinha Filho (2002) apresenta diversos métodos para a estimação dos componentes de variância, destacando o método da Máxima Verossimilhança (*Maximum likelihood*). Tal método consiste na obtenção de estimadores que maximizem a função de densidade de probabilidade das observações a respeito dos efeitos fixos e os componentes de variância.

Seja o modelo misto, apresentando em (9):

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon}$$

Dado que os efeitos aleatórios \mathbf{b}_i e $\boldsymbol{\varepsilon}_i$ têm distribuição Normal, com média = 0 e matrizes de variâncias e covariâncias \mathbf{G} e \mathbf{R} , respectivamente, o vetor \mathbf{Y} terá distribuição Normal multivariada, com média $\mathbf{X}\boldsymbol{\beta}$ e matriz de variâncias e

covariâncias $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$, portanto $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{ZGZ}' + \mathbf{R})$. Sendo a função de máxima verossimilhança (CAMARINHA FILHO, 2002):

$$L = f(\mathbf{Y}, \mathbf{V}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\mathbf{V}|^{\frac{1}{2}}} e^{-\frac{1}{2}[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})]} \quad (16)$$

Fazendo $l = \log(L)$, obtém-se:

$$-\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{V}| - \frac{1}{2} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (17)$$

Sabendo que $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$ e $\mathbf{R} = \sigma^2 \mathbf{I}$, maximiza-se a l (17), obtendo-se os estimadores dos parâmetros $\boldsymbol{\beta}$ e σ^2 (COSTA, 2010):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} \mathbf{Y} \quad (18)$$

$$\hat{\sigma}^2 = \frac{(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{n} \quad (19)$$

Ao estimar a variância σ^2 com o método da máxima verossimilhança seu resultado é viesado, o que leva a parâmetros subestimados (FAUSTO et al., 2008). Isso ocorre em decorrência de dois aspectos (RESENDE et al., 1996): imposição de restrição de não negatividade e perda de graus de liberdade devido à estimação dos efeitos fixos.

Objetivando eliminar o problema com a perda de graus de liberdade, Patterson e Thompson (1971) apresentaram o método da Máxima Verossimilhança Restrita (*Restricted Maximum likelihood - REML*), tornando-se o principal método para estimativa de componentes de variância, mantendo as demais propriedades do método de máxima verossimilhança (RESENDE et al., 1996). Assim, a função da máxima verossimilhança restrita (REML) é dada por (BROWN; PRESCOTT, 2006):

$$L_{REML} = f(\mathbf{Y}, \mathbf{V}) = \frac{1}{(2\pi)^{\frac{n-p}{2}} |\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}|^{\frac{1}{2}} |\mathbf{V}|^{\frac{1}{2}}} e^{-\frac{1}{2}[(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})]} \quad (20)$$

em que: p o número de variáveis fixas observadas (covariáveis fixas do modelo);

Para $l_{REML} = \log(L_{REML})$ em (20), obtém-se:

$$-\frac{n-p}{2} \log(2\pi) - \frac{1}{2} \log|\mathbf{V}| - \frac{1}{2} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) - \frac{1}{2} \log|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}| \quad (21)$$

A função l_{REML} não permite a estimação do coeficiente de efeitos fixos em $\boldsymbol{\beta}$ e por esta razão o estimador $\hat{\boldsymbol{\beta}}$ definido em (18) é utilizado (GAŁECKI; BURZYKOWSKI, 2013). Conhecida a matriz \mathbf{V} em (18), tem-se o melhor estimador não viesado (*Best linear unbiased estimator – BLUE*) (CAMARINHA FILHO, 2002).

O estimador $\hat{\sigma}^2$ é dado por:

$$\hat{\sigma}_{REML}^2 = \frac{(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{n-p} \quad (22)$$

Os efeitos aleatórios refletem desvios ao valor esperado da população $\mathbf{X}\boldsymbol{\beta}$, sendo necessário obter os melhores preditores (PINHEIRO; BATES, 2006). Embora tais preditores possam ser obtidos por métodos desenvolvidos por Henderson (1975), os métodos Bayesianos são preferíveis (NUNES, 2010). Nestes o valor esperado da distribuição *a posteriori* é o valor predito de \mathbf{b} , dado por (23):

$$\hat{\mathbf{b}} = \mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \quad (23)$$

Neste modelo $\hat{\mathbf{b}}$ recebe o nome de estimador *Bayes empírico* (EB) ou, melhor preditor linear não viesado (*Best linear unbiased predictor – BLUP*). Empíricos, devido ao fato de os valores dos efeitos fixos e componentes de variância são substituídos pelas estimativas de máxima verossimilhança restrita (NUNES, 2010).

Nota-se que o estimador do componente $\hat{\sigma}^2$ apresentado em (19) e (22) está em função de $\hat{\boldsymbol{\beta}}$ (18), nesse sentido ambos só podem ser encontrados por métodos numéricos iterativos (CAMARINHA FILHO, 2002; COSTA, 2010). Dentre os métodos adotados, destaca-se o método de *Newton-Raphson*, o método de *Fisher* e o algoritmo *EM* (*Expectation-Maximization*) (BARBOSA, 2009).

Para a estimativa da variável resposta, temos:

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad (24)$$

$$\hat{Y}_1 = \mathbf{X}\hat{\beta} + \mathbf{Z}\hat{b} \quad (25)$$

Métodos quasi-Newton são uma classe de algoritmos eficientes que eliminam a necessidade de cálculo das segundas derivadas para verificação de ponto de máximo. O método BFGS é o método quasi-Newton amplamente utilizado, e é descrito por (BASTIANI, 2012):

Passo 0: Faça $i = 0$ e atribua valores iniciais para o vetor $\theta_0 \in \mathbb{R}^{p+q}$ e para a matriz $\mathbf{M}_0 \in \mathbb{R}^{(p+q) \times (p+q)}$;

Passo 1: Calcule $\mathbf{g}_i = \nabla f(\theta_i)$, o gradiente de f . Se $\mathbf{g}_i = 0$, pare. O vetor θ_i é o mínimo. Caso contrário, faça $\mathbf{d}_i = (-\mathbf{M}_i)^{-1}\mathbf{g}_i$;

Passo 2: Calcule:

$$\lambda_i = \operatorname{argmin}_{\lambda > 0} f(\theta_i + \lambda \mathbf{d}_i)$$

$$\theta_{i+1} = \theta_i + \lambda_i \mathbf{d}_i$$

Passo 3: Atualize a matriz \mathbf{M} :

$$\mathbf{M}_{i+1} = \mathbf{M}_i - \frac{\mathbf{M}_i \mathbf{s}_i \mathbf{s}_i' \mathbf{M}_i}{\mathbf{s}_i' \mathbf{M}_i \mathbf{s}_i} + \frac{\mathbf{y}_i \mathbf{y}_i'}{\mathbf{s}_i' \mathbf{y}_i},$$

em que

$$\mathbf{s}_i = \lambda_i \mathbf{d}_i \text{ e } \mathbf{y}_i = \mathbf{g}_{i+1} - \mathbf{g}_i$$

Passo 4: Faça $i = i + 1$ e volte ao passo 1.

Bastiani (2012) afirma que a matriz \mathbf{M} é sempre positiva-definida, desde que a matriz \mathbf{M}_0 fornecida no passo 0 apresente tal propriedade, superando uma das limitações do método *Newton-Raphson*, em que a matriz $\mathbf{M}_i = -\mathbf{H}_i^{-1}$, \mathbf{H} a matriz hessiana, pode não ser positiva-definida caso se esteja longe do ponto ótimo.

O algoritmo numérico iterativo quasi-Newton denominado L-BFGS-B (*Limited-memory BFGS - Broyden-Fletcher-Goldfarb-Shanno-B*) é uma alternativa aos métodos numéricos citados para maximização de *ML* e *REML*. Ele utiliza uma quantidade limitada de memória computacional (ZHU et al., 1997) e apresenta ainda como vantagem a possibilidade do espaço paramétrico ficar restrito a um intervalo inferior e superior, especificados. Seu objetivo é minimizar uma função não linear de n variáveis, $\min f(x)$, sujeito a $x_i \leq x \leq x_u$, sendo i e u correspondentes aos limites inferior e superior, respectivamente.

De acordo com Bastiani (2012), o algoritmo apresenta o seguinte funcionamento:

- A cada iteração é atualizada uma aproximação da memória limitada BFGS para a matriz hessiana. A matriz de memória limitada presta-se a definir um modelo quadrático da função objetivo f .
- A direção de busca é em seguida calculada usando uma abordagem em 2 fases:
 - A projeção do método gradiente é usada para identificar o conjunto de variáveis ativas, ou seja, variáveis que serão analisadas em seus limites, minimizando-se o modelo quadrático de forma aproximada com relação a essas variáveis.
 - A direção de busca é definida para ser o vetor da presente iteração para a aproximação obtida e finalmente uma pesquisa em linha é realizada ao longo desta direção.

A estrutura do algoritmo L-BFGS-B é apresentado na Figura 6.

Figura 6 – Algoritmo L-BFGS-B – Estrutura.

Algoritmo 1: L-BFGS

Entradas: ponto inicial x_0 , tamanho inteiro $m > 0$, $k = 1$;

Saídas: x com o valor mínimo da função objetivo.

```

1  while não convergir do
2      calcule o gradiente  $\nabla f(x_k)$  na posição  $x_k$ 
3      computar a direção  $p_k$  usando o Algoritmo 2;
4      computar  $x_{k+1} = x_k + \alpha_k p_k$  onde  $\alpha_k$  é escolhido para satisfazer as condições;
5      if  $k > m$  then
6          descartar o par do vetor  $s_{k-m}, y_{k-m}$ 
7      end
8      atualize  $s_k = x_{k+1} - x_k, y_k = \nabla f(x_{k+1}) - \nabla f(x_k), k = k + 1$ ;
9  end

```

Fonte: Adaptado de (CHEN et al., 2014)

Na linha 3 do algoritmo 1 calcula-se a nova direção p_k usando s e y com o gradiente atual $\nabla f(x_k)$.

A abordagem mais comum para o cálculo da nova direção é apresentada na Figura 7.

Figura 7 – Algoritmo para determinação da direção.

Algoritmo 1: L-BFGS

Entradas: $\nabla f(x_k), s_i, y_i$ onde $i = k - m, \dots, k - 1$

Saídas: nova direção p

```

1   $p = -\nabla f(x_k);$ 
2  for  $i \leftarrow k - 1$  to  $k - m$  do
3       $\alpha_i \leftarrow \frac{s_i \cdot p}{s_i \cdot y_i};$ 
4       $p = p - \alpha_i \cdot y_i;$ 
5  end
6   $p = \left( \frac{s_{k-1} \cdot y_{k-1}}{y_{k-1} \cdot y_{k-1}} \right) p;$ 
7  for  $i \leftarrow k - m$  to  $k - 1$  do
8       $\beta = \frac{y_i \cdot p}{s_i \cdot y_i};$ 
9       $p = p + (\alpha_i - \beta) \cdot s_i;$ 
10 end
```

Fonte: Adaptado de (CHEN et al., 2014)

Na Figura 8 é apresentado o algoritmo para maximização da função de máxima verossimilhança (17).

Figura 8 – Algoritmo para maximização de máxima verossimilhança (17).

Algoritmo estimaVarianciasML: Estima as variâncias pelo método da máxima verossimilhança

Entradas: vetor θ_i de valores iniciais para variâncias, vetor \mathbf{Y} , matrizes \mathbf{X} e \mathbf{Z} , q , k

Saídas: valor obtido de ML, matrizes $\hat{\mathbf{G}}$, $\hat{\mathbf{R}}$ e $\hat{\mathbf{V}}$, vetor $\hat{\boldsymbol{\beta}}$, e variâncias estimadas $\hat{\sigma}^2$ e $\hat{\sigma}_e^2$.

```

1  leia  $\theta, \mathbf{Y}, \mathbf{X}, \mathbf{Z}, q, k$ 
2  while L-BFGS-B( $l$ ) <> max( $l$ ) do
3      defina  $\sigma^2 = \theta_1$ , e  $\sigma_e^2 = \theta_2$ 
4      defina  $\hat{\mathbf{R}} = \sigma_e^2 \mathbf{I}_n$ 
5      defina  $\hat{\mathbf{V}} = \mathbf{Z}\mathbf{Z}' + \hat{\mathbf{R}}$ 
6      obtenha  $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$ 
7      defina  $\hat{\mathbf{G}} = \sigma^2 \mathbf{I}_k$ 
8      calcule  $l = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|\mathbf{V}| - \frac{1}{2} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ 
9      atualize  $\theta$ 
10 end
```

Fonte: O autor

De maneira análoga ao algoritmo apresentado para a função de máxima verossimilhança (17), foi definido um algoritmo apresentado na Figura 9 para a maximização da função de máxima verossimilhança restrita (21).

Figura 9 – Algoritmo para maximização de máxima verossimilhança restrita (21).

Algoritmo estimaVariânciasREML: Estima as variâncias pelo método da máxima verossimilhança restrita

Entradas: vetor θ_i de valores iniciais para variâncias, vetor \mathbf{Y} , matrizes \mathbf{X} e \mathbf{Z} , q , k

Saídas: REML, matrizes $\hat{\mathbf{G}}$, $\hat{\mathbf{R}}$ e $\hat{\mathbf{V}}$, vetor $\hat{\boldsymbol{\beta}}$, e variâncias estimadas $\hat{\sigma}^2$ e $\hat{\sigma}_e^2$.

```

1  leia  $\theta$ ,  $\mathbf{Y}$ ,  $\mathbf{X}$ ,  $\mathbf{Z}$ ,  $q$ ,  $k$ 
2  obtenha  $p$  e  $n$ 
3  while L-BFGS-B( $I_{REML}$ ) <> max( $I_{REML}$ ) do
4      defina  $\sigma^2 = \theta_1$ , e  $\sigma_e^2 = \theta_2$ 
5      defina  $\hat{\mathbf{R}} = \sigma_e^2 \mathbf{I}_n$ 
6      defina  $\hat{\mathbf{V}} = \mathbf{Z}\mathbf{Z}' + \hat{\mathbf{R}}$ 
7      obtenha  $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$ 
8      defina  $\hat{\mathbf{G}} = \sigma^2 \mathbf{I}_k$ 
9      calcule  $I_{REML} = -\frac{n-p}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) - \frac{1}{2}\log|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|$ 
10     atualize  $\theta$ 
11  end

```

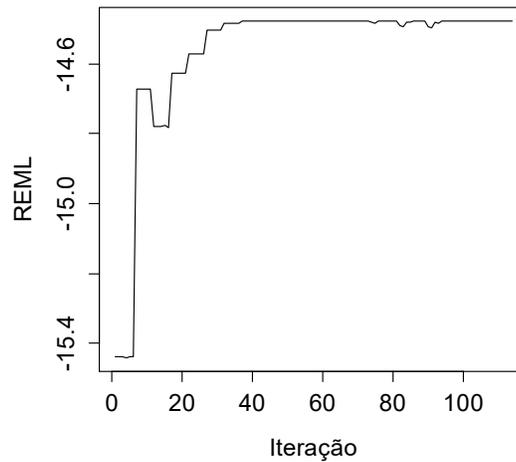
Fonte: O autor

Considerando o exemplo 1 e os dados da Tabela 3 e suas respectivas matrizes $\mathbf{X}_{9 \times 2}$ e $\mathbf{Z}_{9 \times 3}$ definidas, bem como o vetor $\mathbf{Y}_{9 \times 1}$, realiza-se a estimação dos componentes de variância por meio do algoritmo apresentado na Figura 9 para a maximização da função de máxima verossimilhança restrita (21), sendo tais matrizes entradas para o algoritmo. Além das matrizes, deve-se informar o número de covariáveis de \mathbf{Z} , $q = 1$, e $k =$ o número de colunas da matriz \mathbf{Z} . Outra entrada fundamental é um vetor θ_i , sendo $i=1\dots 2$, já que o interesse é a estimação das variâncias σ^2 no intercepto aleatório no nível da parcela, e σ_e^2 para os erros aleatórios.

O gráfico com os valores da função de máxima verossimilhança restrita (REML) obtidos a cada iteração é apresentado na

Figura 10. Os valores correspondentes das variâncias e REML a cada iteração são apresentados no Apêndice A.

Figura 10 – Valores da função de máxima verossimilhança restrita (21) para o exemplo 1.



Fonte: O autor

A convergência do algoritmo pelo método L-BFGS-B ocorreu na 114ª iteração, sendo o valor máximo da função de máxima verossimilhança restrita (REML) (21), e a estimativa das variâncias obtidas é apresentada na Tabela 4.

Tabela 4 – Estimativa dos componentes de variância do exemplo 1 com REML.

σ^2	σ_ε^2	REML
3,612702	0,715864	-14,47735

Fonte: O autor

A estimativa da variância com valor de 3,612702 indica que existe variabilidade “dentro da parcela”, representando heterogeneidade entre indivíduos. A variância que não pôde ser explicada dentro do grupo representado pela parcela σ_ε^2 foi de 0,715864.

A matriz $\hat{\mathbf{G}}$ de covariância dos efeitos aleatórios estimada para o exemplo 1

é denotada por $\mathbf{G} = \begin{bmatrix} 3,612702 & 0 & 0 \\ 0 & 3,612702 & 0 \\ 0 & 0 & 3,612702 \end{bmatrix}_{3 \times 3}$, sendo a matriz de

covariância residual $\mathbf{R} = \begin{bmatrix} 0,7159 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0,7159 \end{bmatrix}_{9 \times 9}$ e a matriz de variância e

covariância de \mathbf{Y} dada pela matriz de bloco diagonal $\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & & \\ & \ddots & \\ & & \mathbf{V}_k \end{bmatrix}_{9 \times 9}$, sendo

$$\mathbf{V}_k = \begin{bmatrix} 4,32856 & 3,61270 & 3,61270 \\ 3,61270 & 4,32856 & 3,61270 \\ 3,61270 & 3,61270 & 4,32856 \end{bmatrix}.$$

Os coeficientes dos efeitos fixos obtidos foram $\boldsymbol{\beta} = \begin{bmatrix} -10,091 \\ 1,987602 \end{bmatrix}$, o que torna

possível a obtenção de $\hat{\mathbf{b}} = \mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ dado por (23):

$$\hat{\mathbf{b}} = \begin{bmatrix} 3,612702 & 0 & 0 \\ 0 & 3,612702 & 0 \\ 0 & 0 & 3,612702 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{V}_1^{-1} & & \\ & \ddots & \\ & & \mathbf{V}_k^{-1} \end{bmatrix}_{9 \times 9} \cdot \begin{bmatrix} -2,65277 \\ -1,01185 \\ -2,79574 \\ 1,205357 \\ 1,285033 \\ 2,448499 \\ 0,86831 \\ 0,810462 \\ -0,1573 \end{bmatrix} = \hat{\mathbf{b}} = \begin{bmatrix} 0,475734 \\ 1,544296 \\ -2,02003 \end{bmatrix},$$

$$\text{sendo } \mathbf{V}_k^{-1} = \begin{bmatrix} 0,960137 & -0,43679 & -0,43679 \\ -0,43679 & 0,960137 & -0,43679 \\ -0,43679 & -0,43679 & 0,960137 \end{bmatrix}$$

Os interceptos do modelo misto apresentado em (14) para o nível de cada parcela apresentado no modelo podem ser obtidos a partir de $\hat{\mathbf{b}}$ preditos e $\hat{\boldsymbol{\beta}}$ pela expressão (26):

$$\beta_0 + b_0 \quad (26)$$

Para os $k=3$ grupos formados pelas parcelas dispostas na Tabela 3, obtém-se os novos interceptos, apresentados na Tabela 5.

Tabela 5 – Parâmetros obtidos para o modelo misto apresentado em (14).

Projeto	Parcela	β_0	b_0	$\beta_0 + b_0$	β_1
1	1		-2,0200	-12,11105	
1	2	-10,091022	1,5443	-8,54672	1,987602
2	1		0,4757	-9,6153	

Fonte: O autor

De posse dos parâmetros estimados, a equação (14) apresentado na forma matricial, pode ser representado na forma funcional por (27).

$$\hat{Y} = \beta_0 + b_0 + \beta_1 DAP \quad (27)$$

Dada a expressão (27), pode-se obter o valor estimado de HT da árvore 1, parcela 1 do projeto 1, na idade 36 meses apresentada na Tabela 3 por:

$$\hat{HT} = -12,11105 + 1,987602 \cdot 12,6 = 12,93274$$

Na forma matricial, a estimativa de HT no nível de parcela pode ser obtida por (28).

$$\hat{Y} = X\beta + Zb \quad (28)$$

Portanto para o exemplo 1, tem-se:

$$\hat{Y}_1 = \begin{bmatrix} -2,65277 \\ -1,01185 \\ -2,79574 \\ 1,205357 \\ 1,285033 \\ 2,448499 \\ 0,86831 \\ 0,810462 \\ -0,1573 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0,475734 \\ 1,544296 \\ -2,02003 \end{bmatrix} = \hat{Y}_1 = \begin{bmatrix} 12,93274 \\ 19,49182 \\ 22,07571 \\ 15,10574 \\ 20,86979 \\ 23,25491 \\ 18,80742 \\ 25,56527 \\ 30,73304 \end{bmatrix}$$

Os erros aleatórios no nível da parcela podem ser obtidos por

$$\varepsilon_1 = \mathbf{Y} - \hat{\mathbf{Y}}_1 = \begin{bmatrix} -0,63274 \\ 1,008176 \\ -0,77571 \\ -0,33894 \\ -0,25926 \\ 0,904203 \\ 0,392576 \\ 0,334728 \\ -0,63304 \end{bmatrix}$$

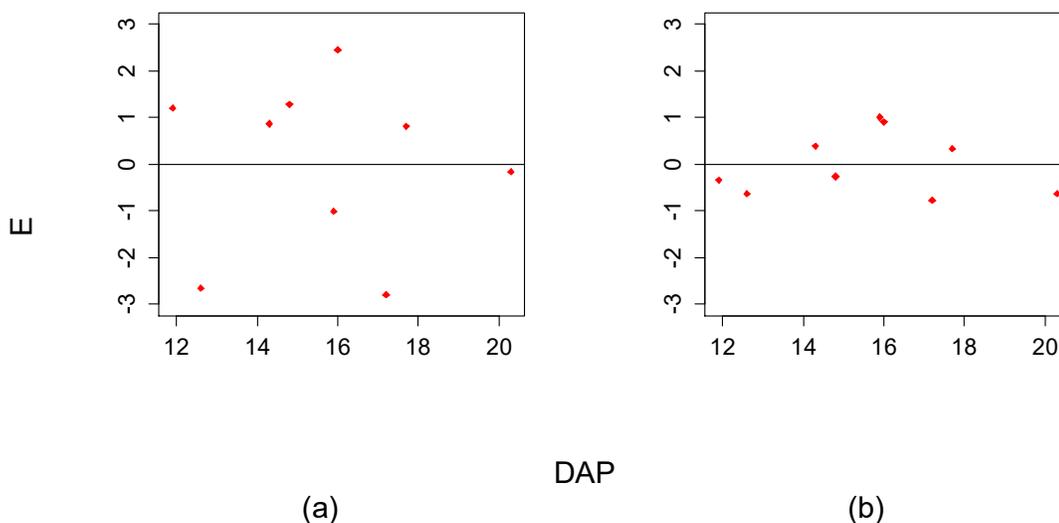
Considerando apenas os coeficientes referentes ao efeito fixo do modelo

(14), temos uma regressão linear clássica simples, cujo $\hat{\mathbf{Y}} = \begin{bmatrix} 14,95277 \\ 21,51185 \\ 24,09574 \\ 13,56144 \\ 19,32549 \\ 21,71061 \\ 18,33169 \\ 25,08954 \\ 30,2573 \end{bmatrix}$ para o nível

da população, e erros $\varepsilon = \mathbf{Y} - \hat{\mathbf{Y}} = \begin{bmatrix} -2.65277 \\ -1.01185 \\ -2.79574 \\ 1.205357 \\ 1.285033 \\ 2.448499 \\ 0.86831 \\ 0.810462 \\ -0.1573 \end{bmatrix}$

Graficamente, tem-se os resíduos apresentados na Figura 11.

Figura 11 – Resíduos modelo linear simples clássico (a) e linear misto (b).



Fonte: O autor

Observa-se pela análise gráfica residual que o modelo misto foi mais adequado para os dados apresentados no exemplo 1, por apresentar uma menor dispersão.

Com o propósito de identificar possíveis tendências nas estimativas de HT, a raiz quadrada do erro médio foi calculada para o modelo linear simples clássico (13) e linear misto (14) (ARAÚJO JÚNIOR et al., 2010):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n}} \quad (29)$$

em que: RMSE é a raiz quadrada do erro médio;

Os resultados são apresentados na Tabela 6 refletem numericamente o que pôde ser observado na Figura 11.

Tabela 6 – RMSE para o modelo linear simples clássico (13) e linear misto (14).

Modelo	RMSE
Fixo (equação 10)	0,214
Misto (equação 11)	0,080

Fonte: O autor

A convergência do algoritmo pelo método L-BFGS-B ocorreu na iteração 468, sendo o valor máximo da função de máxima verossimilhança restrita (REML) (21), e a estimativa das variâncias obtidas são apresentadas na Tabela 7.

Tabela 7 – Estimativa dos componentes de variância do exemplo 2, com o modelo (15) com REML.

σ^2	σ_s^2	σ_ε^2	REML
3,612702	$6,055 \cdot 10^{-6}$	0,715867	- 14,47772

em que σ_s^2 é a variância da inclinação da covariável S do modelo (15).

Os coeficientes do modelo misto apresentado em (15) para exemplo 2 foram obtidos com a execução do algoritmo definido na Figura 9 para o nível de cada parcela com intercepto e inclinação, e os $k=3$ grupos formados pelas parcelas dispostas na Tabela 3 são apresentados na Tabela 8.

Tabela 8 – Parâmetros obtidos modelo misto apresentado em (15).

Projeto	Parcela	β_0	b_0	$\beta_0 + b_0$	β_1	b_1
1	1		-2,0200	-12,109246		$-7,954 \cdot 10^{-5}$
1	2	-10,091590	1,5443	-8,549110	1,987602	$7,892 \cdot 10^{-5}$
2	1		0,4757	-9,616415		$2,9892 \cdot 10^{-5}$

Fonte: O autor

Os coeficientes dos efeitos fixos apresentados na Tabela 8, $\beta = \begin{bmatrix} -10,091 \\ 1,987602 \end{bmatrix}$,

demonstrando que $\hat{\mathbf{b}} = \mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\beta)$ dado por (23):

$$\hat{\mathbf{b}} = \begin{bmatrix} 3,609784 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6,06 \cdot 10^{-6} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3,609784 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6,06 \cdot 10^{-6} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3,609784 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6,06 \cdot 10^{-6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 37,5 & 37,5 & 37,5 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 30,5 & 30,5 & 30,5 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 23,5 & 23,5 & 23,5 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$$\begin{bmatrix} 4,328987 & 3,613128 & 3,613128 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3,613128 & 4,328987 & 3,613128 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3,613128 & 3,613128 & 4,328987 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4,331276 & 3,615417 & 3,615417 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3,615417 & 4,331276 & 3,615417 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3,615417 & 3,615417 & 4,331276 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4,334158 & 3,6183 & 3,6183 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3,6183 & 4,334158 & 3,6183 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3,6183 & 3,6183 & 4,334158 \end{bmatrix} \cdot \begin{bmatrix} -2,65277 \\ -1,01185 \\ -2,79574 \\ 1,205357 \\ 1,285033 \\ 2,448499 \\ 0,86831 \\ 0,810462 \\ -0,1573 \end{bmatrix} = \hat{\mathbf{b}} = \begin{bmatrix} 0,475175659 \\ 2,99 \cdot 10^{-5} \\ 1,542479881 \\ 7,89 \cdot 10^{-5} \\ -2,01765554 \\ -7,95 \cdot 10^{-5} \end{bmatrix}, \text{ sendo } \hat{\mathbf{b}}_{1 \times 6}, \text{ as linhas ímpares referem-se ao}$$

intercepto e pares aos coeficientes de inclinação aleatórios. Representando em

$$\text{colunas, tem-se } \hat{\mathbf{b}}_{3 \times 2} = \begin{bmatrix} 0,475176 & 2,99 \cdot 10^{-5} \\ 1,54248 & 7,89 \cdot 10^{-5} \\ -2,01766 & -7,95 \cdot 10^{-5} \end{bmatrix}.$$

Sabendo que $q > 1$, uma vez que o exemplo 2 considera a variável S como covariável aleatória, faz-se necessário reformular a matriz \mathbf{G} de dimensão $k \times k$, já que $h = 3$ e $q = 2$, então $k=6$. A matriz \mathbf{G} construída considera que a covariância entre o resíduo do intercepto e da inclinação é 0, e as variâncias são iguais para os grupos formados em h . Portanto \mathbf{G} é bloco-diagonal para h grupos, conforme demonstrado em (30).

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & 0 & 0 \\ 0 & \mathbf{G}_2 & 0 \\ 0 & 0 & \mathbf{G}_h \end{bmatrix}_{k \times k} \quad (30)$$

em que $\mathbf{G}_h = \begin{bmatrix} \sigma_{b_0}^2 & 0 \\ 0 & \sigma_{b_1}^2 \end{bmatrix}_{q \times q}$, sendo $\sigma_{b_0}^2 =$ variância do intercepto, e $\sigma_{b_1}^2$ a variância

da inclinação (HERSHBERGER; MOSKOWITZ, 2002).

De posse dos parâmetros estimados, o modelo (15) apresentado na forma matricial, pode ser representado na forma funcional pela equação (31).

$$\hat{Y} = \hat{\beta}_0 + \hat{b}_0 + \hat{\beta}_1 DAP + \hat{b}_1 S \quad (31)$$

Dado o modelo (15) pode-se obter a estimativa de HT da árvore 1, parcela 1 do projeto 1, na idade 36 meses por:

$$\hat{HT} = -12.109246 + 1.987602 \cdot 12.3 + -7.954 \cdot 10^{-5} \cdot 23.5 = 12.33639$$

2.2.3. Diagnóstico do modelo

Concluída a modelagem é essencial efetuar o diagnóstico do modelo ajustado para avaliar se os pressupostos de distribuição para modelos mistos são válidos e se o modelo proposto é sensível a observações incomuns. Diferente dos modelos clássicos de regressão cuja literatura sobre diagnósticos é abundante, para os modelos mistos estas são de difícil interpretação (QUEIROZ, 2012). Contudo, segundo Pinheiro e Bates (2006), dois são os tipos básicos de pressupostos da distribuição de modelos mistos:

1. Os erros dentro de um grupo são independentes e identicamente distribuídos seguindo a distribuição Normal, com média 0 e variância σ^2 , e independentes dos efeitos aleatórios;
2. Os efeitos aleatórios são normalmente distribuídos, com média 0 e matriz de covariância \mathbf{G} e são independentes para diferentes grupos.

Os métodos mais utilizados para avaliação são aqueles que utilizam representação gráfica como o de resíduos e valores ajustados (NUNES, 2010). Como alternativa a avaliação de qualidade do ajuste pode ser feita pela comparação dos dados observados com o modelo ajustado, no sentido de avaliar sua adequação, também por métodos gráficos.

2.3. Qualidade dos Dados para processamento de Modelos de Crescimento e Produção

A qualidade das estimativas e resultados obtidos por modelos de crescimento e produção nos diferentes níveis depende de muitos fatores, sendo a qualidade dos dados utilizados no ajuste o principal deles (MAGALHÃES, 2014).

Dados eletrônicos desempenham um papel fundamental na sociedade da tecnologia da informação e comunicação (TIC): elas são gerenciadas por aplicativos

empresariais e governamentais, por todos os tipos de aplicativos na Web, e são fundamentais em todas as relações entre governos, empresas e cidadãos, sendo a "qualidade" de tais dados e seus efeitos relacionados, cada tipo de atividade da sociedade TIC são cada vez mais críticas.

Batini e Scannapieca (2006) afirmam que a qualidade dos dados nos processos decisórios e operacionais é reconhecida por várias instituições e organizações, como por exemplo, na declaração da missão do *European Statistical System*: "Nós fornecemos à União Europeia e ao mundo com informações de alta qualidade sobre economia e a sociedade no nível Europeu, nacional e regional e tornamos a informação disponível a todos para fins de tomada de decisão, pesquisa e debate.". Afirmam ainda que a qualidade dos dados também é uma questão essencial para os processos operacionais das empresas e organizações.

O *Data Warehousing Institute* em um relatório de 2002 sobre a qualidade dos dados mostra que há uma diferença significativa entre a percepção e a realidade sobre a qualidade dos dados em muitas organizações, e que o problema de qualidade de dados custou as empresas da América do Norte 600 milhões de dólares em um ano.

A qualidade dos dados é definida por dois fatores relacionados: o quão bem ele atende às expectativas dos consumidores de dados (o quão bem ele é capaz de servir aos propósitos de sua utilização ou utilizações) e quão bem ele representa os objetos, eventos e conceitos que é criado para representar. A fim de medir se os dados atende às expectativas ou está "apto para a utilização", expectativas e usos precisam ser definidos (SEBASTIAN-COLEMAN, 2013).

Lima (2010) destaca a importância dada à qualidade de dados, quando relata que o Fundo Monetário Internacional (FMI) disponibiliza um modelo para aprimoramento e avaliação das informações: *Data Quality Assessment Framework* (DQAF). A autora ainda declara que o escritório de estatística da comunidade europeia (*Eurostat*) adaptou esse modelo para a comunidade europeia, criando o *Quality Declaration of the European Statistics System*.

Na tecnologia da informação o desenvolvimento de metodologias para avaliação da qualidade dos bancos de dados se baseia nos princípios de gerenciamento da qualidade total.

O DQAF destaca em um conjunto de características de dados associados cinco dimensões da qualidade dos dados: integridade, oportunidade, validade,

consistência e integridade (MALAVERRI; MEDEIROS, 2012; SEBASTIAN-COLEMAN, 2013). Sebastian-Coleman (2013) apresenta cada definição:

- *Completeness*: implica em ter todas as partes necessárias ou adequadas dos dados. A primeira condição é a existência de completude, afinal devem existir dados antes que possa ser classificado completo. Em muitas empresas, as expectativas sobre os dados podem não ser atendidas porque são baseadas no pressuposto de que certos dados já existem quando não existem, ou que os dados existem numa forma utilizável, mas isso não é a realidade. Um conjunto de dados é completo na medida em que ele contém variáveis (ou atributos) necessárias e um número suficiente de registros, e na medida em que são preenchidas de acordo com as expectativas dos consumidores de dados. Para um conjunto de dados para ser completo, pelo menos, três condições devem ser atendidas: o conjunto de dados deve ser definida de modo a incluir todos as variáveis desejados (*width*); o conjunto de dados deve conter a quantidade desejada de dados (*depth*); e as variáveis devem ser preenchidas até ao ponto desejado (*density*). Também pode ser entendido em relação ao processamento de dados, por exemplo, ter todos os dados necessários ou convenientes para iniciar o processamento e garantir que parte dos dados não foi retirado de um conjunto de dados quando ele é processado
- *Pontualidade*: tem sido definida como o grau em que os dados representam a realidade a partir do ponto desejado no tempo. No que diz respeito ao processamento, a pontualidade está associada com a disponibilidade de dados, o grau em que os clientes têm os dados necessários no momento certo.
- *Validade*: A validade é o grau em que os dados estão em conformidade com um conjunto de regras de negócio, às vezes expressa como um padrão ou representado dentro de um domínio de dados definido. Validade é diferenciada de precisão e exatidão, cada um dos quais requer comparação com objetos reais. Validade se limita a medir contra substitutos ou alguns valores alternativos de objetos reais. Estes substitutos podem ser instanciados como dados, permitindo validade para ser medida diretamente a partir de um conjunto de dados de uma forma que a precisão e exatidão não pode ser.

- *Consistência*: A consistência é o grau em que os dados estão em conformidade com um conjunto equivalente de dados, normalmente um conjunto produzido sob condições semelhantes, ou um conjunto produzido pelo mesmo processo ao longo do tempo. Medidas de consistência podem ser definidas em conformidade com as regras de negócio (dependências que podem ser estabelecidas tecnicamente se a lógica pode ser claramente expressa), ou podem revelar padrões lógicos criados dentro dos dados que refletem as conexões entre as situações do mundo real que os dados representam.
- *Integridade*: Representa o grau em que os dados estão em conformidade com regras de relacionamento de dados (como definido pelo modelo de dados) que se destinam a assegurar a apresentação completa, consistente e válida de dados que representam os mesmos conceitos, normalmente considerando outras dimensões.

Dados de inventário florestal e aqueles necessários ao ajuste de modelos de crescimento e produção nos diferentes níveis devem apresentar tais características, ou a qualidade da informação produzida pode não representar qualquer aderência com a realidade das florestas plantadas.

A validade de dados deve ser considerada não apenas pontualmente em cada medição, mas de forma a manter a integridade entre medições ao longo do período de existência de uma parcela, e os indivíduos que a compõem. Sendo assim, do ponto de vista da completude, para modelos de crescimento e produção, completo é o dado caracterizado por medida repetidas do mesmo indivíduo.

A integridade deve ser mantida na medida em que os dados produzidos, por exemplo para as prognoses apresentem a relação de origem-destino e sua rastreabilidade garantida. Os resultados de uma prognose devem ser rastreados para os modelos pelos quais foram produzidas, pelas estatísticas de ajuste, análises realizadas, até o nível da árvore individual para modelos de árvores individuais.

Uma característica fundamental é a pontualidade, sobretudo para o processamento, espera-se que o dado esteja disponível no momento em que seja necessário realizar um estudo de prognose para avaliação de disponibilidade de múltiplos produtos em uma idade futura. Caso ocorra a indisponibilidade, os clientes que demandam a informação podem ficar frustrados causando estresse no ambiente organizacional e até mesmo levando membros da equipe a criarem dados irreais, ou usar métodos não recomendados para sua geração.

Tanto a completude quanto a consistência devem ser avaliadas em cada medição e entre medições. A falta de informações sobre o DAP de uma ou mais árvores em algumas parcelas pode prejudicar o ajuste de um modelo que seria promissor para determinado propósito. Contudo, mesmo na presença do valor número do DAP é necessário avaliar, ou verificar sua validade frente ao comportamento esperado para esta variável. Faz-se ainda necessária a consistência destas variáveis do ponto de vista longitudinal, ou seja, entre o conjunto de medições realizadas.

Por um erro cometido em diferentes medições, por exemplo, a inversão da ordem de entrada em uma parcela em diferentes períodos de inventário pode acarretar em comportamento decrescente do valor de HT das mesmas árvores ao longo do tempo. Tal característica é fundamental para garantir a qualidade dos dados para ajuste de modelos de crescimento e produção.

2.3.1. Localização e correção de erros

Localização e correção de erros são úteis toda vez que os dados foram coletados a partir de fontes passíveis de erros (por exemplo, aqueles em que a entrada manual foi realizada) ou adquiridos a partir de fontes cuja confiabilidade não é totalmente garantida e sofre pressão, por exemplo, para cumprimento de prazos, metas e outros (BATINI; SCANNAPIECA, 2006).

Após a localização de registros errados, a fim de corrigir erros, que pode executar sobre eles a aquisição de novos dados. Porém sabe-se que este tipo de atividade é geralmente muito caro, e, em todos os contextos em que são recolhidos dados para fins estatísticos, o uso de edições é geralmente preferido para corrigir dados errados.

O problema de localização de erros por meio de edição de campos errados é normalmente referido como o problema de edição de imputação. Batini e Scannapieca (2006) apresentam um modelo teórico para estes problemas, sendo seus principais objetivos: Os dados em cada registro deve satisfazer todas as edições mudando o menor número de campos possível. Este é o chamado princípio da mudança mínima.

Quando a entrada de dados é necessária, deseja-se manter a distribuição de frequência conjunta de valores nos diferentes campos entre diferentes registros. Sendo assim, ao modificar o registro, por exemplo o valor de DAP, é importante observar se tal modificação pode alterar o comportamento das variáveis

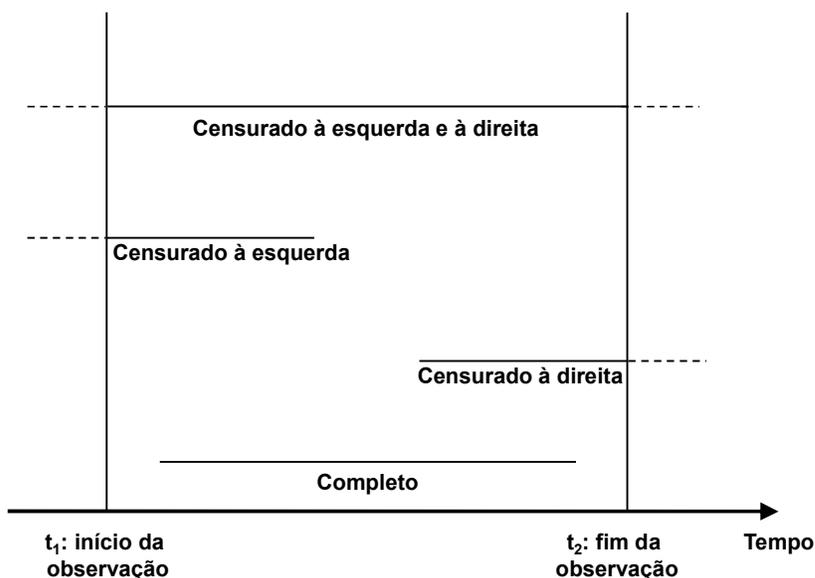
relacionadas a ele no registro e também no conjunto representado por outros registros. Caso um DAP numa idade inicial seja modificado para um valor comum apenas em idades perto do corte, o comportamento global de árvores naquela idade pode ser modificado.

No que diz respeito a dados longitudinais, medidos ao longo do tempo, dois tipos de incompletudes podem ser identificados: dados truncados e dados censurados (BATINI; SCANNAPIECA, 2006).

Dados truncados correspondem a observações que são descartados a partir do conjunto de dados analisados. Por exemplo, parcelas abaixo de 24 meses ou árvores com DAP abaixo de 5 cm. Já os dados censurados correspondem aos dados que sabemos com certeza não foram medidos antes de um período $p1$ (dados censurados à esquerda) ou depois de um período $p2$ (dados censurados à direita). Como um exemplo de dados censurados à esquerda, assumir que estamos interessados em realizar a classificação de sítio de um povoamento; poderíamos ter apenas dados disponíveis depois de uma medição, dados de medições anteriores poderiam não estar disponíveis.

A Figura 12 ilustra os casos mencionados.

Figura 12 – Tipos de dados incompletos.



Fonte: Adaptado de Batini e Scannapieca (2006).

Para a detecção deste tipo de problema, histogramas e distribuições de frequência podem ser utilizados por um especialista no domínio.

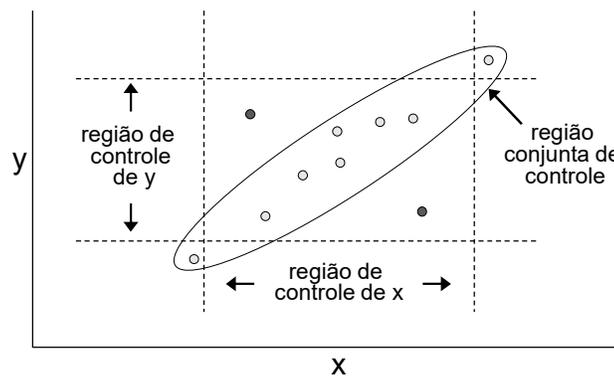
A descoberta de *outliers* consiste em identificar valores altos ou baixos demais em relação à variável analisada. Tipicamente, um *outlier* é atribuível a uma das seguintes causas na medição de dados:

1. o objeto não foi corretamente observado, o dado gravado, ou durante o registro no banco de dados;
2. Trata-se de uma população diferente, em relação a outros valores, mesmo considerando o diferentes períodos de medições.
3. Que é correto, mas representa um evento raro.

Em todos os casos é necessário analisar a fim de identificar a origem do desvio. No caso 2 pode ocorrer de uma mesma parcela permanente ser medida em 2 inventários por equipes diferentes e em um dos casos a equipe inverteu a ordem de caminhada na parcela.

Gráficos de controle são úteis na detecção de *outliers*, porém representam apenas paredes de variáveis, conforme pode ser observado na Figura 13.

Figura 13 – Gráficos de Controle



Fonte: Adaptado de Batini e Scannapieca (2006).

No que diz respeito a dados longitudinais, medidos ao longo do tempo, dois tipos de incompletudes podem ser identificados: dados truncados e dados censurados (BATINI; SCANNAPIECA, 2006).

Métodos numéricos para a detecção de *outliers* como o Z-Score Modificado (ANDRIOTTI, 2005) que utiliza a mediana amostral e o desvio absoluto da mediana amostral (MAD):

$$MAD = mediana(|z_t - \tilde{z}|) \quad (32)$$

em que \tilde{z} é a mediana amostral da série.

O Z-Score Modificado é, então, calculado pela expressão:

$$ZM = \frac{0.6745(z_t - \bar{z})}{MAD} \quad (33)$$

Consideram-se *outliers* as observações cujo ZM (em módulo) sejam superiores a 3,5.

2.4. Descrição e processamento prévio dos dados

2.4.1. Descrição dos dados

Na condução deste estudo foram utilizados dados de povoamentos não desbastados com 20 genótipos (clones de eucalipto) localizados na região Centro-oeste do estado de MG plantados no espaçamento 3 x 3 m. Os dados foram provenientes de 376 parcelas permanentes, variando entre 274 e 356 m², medidas entre 2006 e 2012, sendo o número de parcelas distribuídas de acordo com o número de medições, por idade apresentadas na Quadro 1.

Em todas as parcelas foram medidos anualmente os diâmetros a 1,30 m de altura (*DAP*) de todas as árvores, alturas totais (*HT*) de dez árvores e alturas das árvores dominantes.

Quadro 1 – Distribuição das parcelas de acordo com o número de medições e idade

Número de Medições/Idade (meses)	24	36	48	60	72	84	96
4	77	136	177	187	110	51	10
5	113	135	153	153	153	40	18
6	20	23	23	23	23	22	4
7	13	13	13	13	13	13	13

As características médias das parcelas utilizadas no presente trabalho são apresentadas na Tabela 9.

Tabela 9 – Características médias das 376 parcelas utilizadas neste estudo.

Idade	Parcelas	<i>Dap</i>			<i>Ht</i>			<i>Hd</i>			<i>B</i>		
		Min	Méd	Máx	Min	Méd	Máx	Min	Méd	Máx	Min	Méd	Máx
24	223	5	10,3	16	6,4	12,8	21,6	4,3	8,6	14,7	8,5	13,5	21,0
36	307	5	13,5	19,4	8,8	18,5	26,3	8,8	14,4	20,7	12,7	19,4	26,0
48	366	5	15,3	22,6	9,9	22,1	31,4	12,0	18,7	25,1	18	23,5	31,0
60	376	5	16,4	25,5	10,7	24,6	35,0	15,0	21,5	27,7	19,9	26,1	32,6
72	299	5	17,3	26,5	11,6	26,3	38,6	16,6	23,9	30,5	19,9	28,0	35,4
84	126	5,4	18,3	28,1	13,1	28,1	38,8	18,2	26,5	33,3	21,7	29,8	37,9
96	45	5,5	19,0	27,3	13,3	29,8	46,3	20,0	28,5	36,4	25,4	31,5	38,2

Idade: classes de idade em meses; *N*: número de parcelas medidas na classe de idade; *dap*: diâmetro a 1,3 de altura, em cm; *Ht*: altura total, em m; *Hd*: altura média das árvores dominantes, em m; *B*: área basal, em m² ha⁻¹; e Mín, Méd e Máx: valores característicos mínimos, médios e máximos, respectivamente.

2.4.2. Processamento dos dados

Todo o processamento dos dados e geração de resultados foram realizados na linguagem R (R DEVELOPMENT CORE TEAM, 2005).

A manipulação dos dados foi feita utilizando os pacotes da linguagem R: *spldf* (GROTHENDIECK, 2014), *data.table* (DOWLE et al., 2014), *plyr* (WICKHAM, 2011) e *Fgmutils* (FRAGA FILHO et al., 2015).

O ajuste dos modelos não lineares foi realizado utilizando funções do pacote da linguagem R denominado *minpack.lm* (ELZHOV et al., 2013). Para ajustar modelos lineares e não lineares mistos, foi utilizado o pacote *nlme* (PINHEIRO et al., 2014), outras rotinas específicas para preparação dos dados para ajuste nos diferentes níveis, assim como rotinas necessárias para a consistência dos dados foram desenvolvidas pelo autor na referida linguagem.

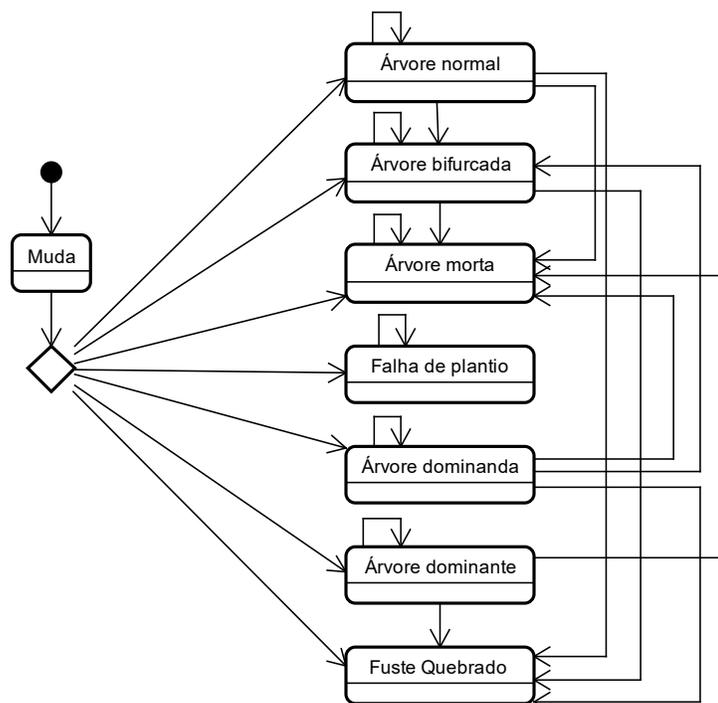
2.4.3. Consistência da base de dados

Procedeu-se com as operações de consistência da base de dados com o objetivo de encontrar desvios e problemas com a qualidade dos dados, buscando minimizar a inclusão de erros no processo de modelagem de crescimento e produção. Tais operações de consistência foram realizadas em diferentes níveis,

contudo, explorando o nível mais elementar para efeitos de inventário florestal contínuo: o de árvore individual.

De forma complementar, para melhorar o entendimento da quantificação de características qualitativas, considera-se definir estados atuais e estados associados, que podem ocorrer ao longo de uma rotação, apresentados na Figura 14.

Figura 14 – Estados associados a covas, árvores e fustes em parcelas permanentes instaladas em povoamentos equiâneos.



Fonte: Adaptado de Campos e Leite (2013).

Na presente base de dados a informação qualitativa referente aos estados dos fustes, representados pela variável CODIGOS na Figura 15.

Figura 15 – Estados dos fustes representados pelo campo CODIGO

A - Dominada
 B - Bifurcada
 D - Dominante
 F - Falha
 I - Inclinada
 M - Morta
 N - Árvore Normal
 Q - Quebrada

Foram realizados os seguintes procedimentos de consistência:

- Avaliação da distribuição de frequência de cada estado ao longo do tempo;
- Identificação de árvores com DAP e HT = 0 e CODIGO diferentes de M, F ou A.
- Identificar árvores com $DAP2 < DAP1$ entre medições consecutivas i e $i+1$, sendo i o ano da medição;
- Identificação de número de fustes mortos (CODIGO = M) ao longo das medições;
- Gráficos de interação mostrando a média de uma variável ao longo das medições.

A geração de dados para modelos de crescimento e produção envolve a criação de estruturas de dados com valores de algumas variáveis de interesse inseridas em pares de colunas ij , onde i representa a variável medida no período 1 e j a mesma variável medida no período 2.

O que pode ser observado no esquema representado na Figura 16.

Figura 16 – Dados pareados gerados a partir de dados de inventário.

Dado do inventário					Dado pareado								
Ano	ARV	I	DAP	HT	ARV	Ano1	Ano2	I1	I2	DAP1	DAP2	HT1	HT2
2003	1	1	12.3	15.1	1	2003	2004	1	2	12.3	13.2	15.1	17.2
2004	1	2	13.2	17.2	1	2004	2005	2	3	13.2	14.1	17.2	19.1
2005	1	3	14.1	19.1									

(a)

(b)

em que: I é a idade da árvore; DAP é o diâmetro a 1,3 m e HT a altura total;

Na Figura 16 observa-se três medições de uma árvore identificada como ARV 1 em (a) e os dados pareados criados a partir de (a) expressos em (b). Esta estrutura de dados definida para diferentes níveis de modelagem denominados dados pareados.

Nos modelos em nível de árvore individual os dados pareados são de cada árvore identificada na base de dados unicamente pareados em medição i e medição $i+1$. Salienta-se que é necessário que exista mais de uma medição para cada árvore, fato que deve ser consistido antes da definição dos dados pareados para o modelo de árvore individual.

Para os outros níveis de modelagem de crescimento e produção, os dados pareados também são definidos. Para modelos em nível de povoamento, os dados

pareados referem-se a características de unidades de manejo, como parcelas permanentes. Em geral, modelo em nível de povoamento considera a necessidade das seguintes variáveis: área basal do povoamento na idade atual e futura ($m^2 ha^{-1}$); a idade atual e futura, o índice de local; o volume de madeira com casca do povoamento, na idade atual e futura ($m^3 ha^{-1}$).

Para os modelos de distribuição de diâmetros, o número de árvores observadas por centro de classe de diâmetro (com amplitude definida) são definidos como observações dentro do registro da parcela para cada idade medida.

2.4.3.1. Avaliação da distribuição de frequência de cada estado ao longo do tempo

Na Tabela 10 é apresentada a quantidade de árvores em cada ano de inventário, agrupadas por código.

Tabela 10 – Avaliação da distribuição de frequência de cada estado ao longo do tempo.

Código/Ano	2006	2007	2008	2009	2010	2011	2012
Código em branco	465	661	70	0	0	0	0
A	272	149	118	99	76	73	156
B	85	44	95	64	37	42	46
D	2471	2374	2214	2054	1855	1453	1152
F	7150	6193	4462	3046	1380	996	1632
I	1380	1228	1320	1040	594	221	286
M	195	161	159	327	85	20	172
N	21605	19633	18147	17368	15160	16837	21687
Q	148	259	271	237	144	87	161
0	0	0	310	0	0	0	0
C	0	0	1	0	0	0	0
T	0	0	0	0	2	3	46

Observa-se que existem árvores com estado indefinido (campo CODIGO em branco, ou seja, sem código) nos anos de 2006 (465), 2007 (661) e 2008 (70). Verificou-se que todas têm DAP e Altura = 0, o que caracteriza que são registros que não existem, podendo ser eliminados. Ainda é possível observar que surgiram dois novos códigos na tabela, diferentes daqueles apresentados: 0 e C. Após consulta com o responsável pela base de dados, os mesmos foram caracterizados como erros e foram removidos.

2.4.3.1. Identificação de árvores com DAP e HT = 0 e CODIGO diferentes de M, F ou A.

Na Tabela 11 são apresentadas as quantidades de árvores com DAP e HT iguais a zero, porém que não estão no estado MORTA, FALHA ou DOMINADA.

Tabela 11 – Identificação de árvores com DAP e HT = 0 e CODIGO diferentes de M, F ou A, agrupados por código/ano

Código/Ano	2006	2007	2008	2009	2010	2011	2012
A	272	149	118	99	75	73	156
B	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0
F	7150	6193	4462	3046	1379	993	1632
I	0	0	0	1	2	0	0
M	188	94	9	327	65	12	172
N	0	0	4	3	0	14	0
Q	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0

Observa-se que existem árvores Inclinadas e Normais, fato que não deveria ocorrer, o que pode caracterizar problema na medição, registro, armazenamento ou até mesmo classificação qualitativa incorreta do estado das árvores em campo. Tal situação pode demandar uma análise mais profunda dependendo do número de observações.

É possível observar na Figura 17 um detalhamento para uma parte das árvores cujo estado é Normal presentes no resumo da Tabela 11.

Figura 17 – Estrato de árvores com estado NORMAL com DAP = 0

Talhão 77 , parcela= 1 , COVA= 48 , FUSTE= 2												
ANO_IFC	PROJETO	COD_TALHAO	CICLO	ROT	IDADE	MG	PARCELA	COVA	FUSTE	FILA	DAP	HT CODIGOS
1	2006	PROJETO 0001	77	NA	2 117,7000	1	1	48	2	6	15,4	0 N
2	2008	PROJETO 0001	77	1	2 147,1211	1	1	48	2	6	0,0	0 N

Talhão 84 , parcela= 2 , COVA= 43 , FUSTE= 1												
ANO_IFC	PROJETO	COD_TALHAO	CICLO	ROT	IDADE	MG	PARCELA	COVA	FUSTE	FILA	DAP	HT CODIGOS
1	2006	PROJETO 0001	84	NA	2 118,8000	1	2	43	1	6	11,5	0 N
2	2008	PROJETO 0001	84	1	2 147,3840	1	2	43	1	6	0,0	0 N
3	2009	PROJETO 0001	84	1	2 159,0801	1	2	43	1	6	0,0	0 F

Talhão 3 , parcela= 3 , COVA= 3 , FUSTE= 2												
ANO_IFC	PROJETO	COD_TALHAO	CICLO	ROT	IDADE	MG	PARCELA	COVA	FUSTE	FILA	DAP	HT CODIGOS
1	2006	PROJETO 0002	3	NA	2 99,0000	2	3	3	2	1	7,0	0 N
2	2007	PROJETO 0002	3	1	2 105,9220	2	3	3	2	1	7,3	0 N
3	2008	PROJETO 0002	3	1	2 117,5524	2	3	3	2	1	0,0	0 N

Talhão 41 , parcela= 4 , COVA= 7 , FUSTE= 1												
ANO_IFC	PROJETO	COD_TALHAO	CICLO	ROT	IDADE	MG	PARCELA	COVA	FUSTE	FILA	DAP	HT CODIGOS
1	2007	PROJETO 0003	41	1	2 84,27105	3	4	7	1	1	31,4	36,6 D
2	2008	PROJETO 0003	41	1	2 93,07598	3	4	7	1	1	32,7	39,1 D
3	2009	PROJETO 0003	41	1	2 104,08214	3	4	7	1	1	0,0	0,0 N

Talhão 41 , parcela= 5 , COVA= 9 , FUSTE= 1												
ANO_IFC	PROJETO	COD_TALHAO	CICLO	ROT	IDADE	MG	PARCELA	COVA	FUSTE	FILA	DAP	HT CODIGOS
1	2007	PROJETO 0003	41	1	2 84,27105	3	5	9	1	2	24,0	30,8 D
2	2008	PROJETO 0003	41	1	2 93,07598	3	5	9	1	2	24,8	31,9 D
3	2009	PROJETO 0003	41	1	2 104,08214	3	5	9	1	2	0,0	0,0 N

Talhão 41 , parcela= 5 , COVA= 55 , FUSTE= 1												
ANO_IFC	PROJETO	COD_TALHAO	CICLO	ROT	IDADE	MG	PARCELA	COVA	FUSTE	FILA	DAP	HT CODIGOS
1	2007	PROJETO 0003	41	1	2 84,27105	3	5	55	1	7	25,2	31,9 D
2	2008	PROJETO 0003	41	1	2 93,07598	3	5	55	1	7	26,5	29,2 D
3	2009	PROJETO 0003	41	1	2 104,08214	3	5	55	1	7	0,0	0,0 N

Talhão 21 , parcela= 6 , COVA= 9 , FUSTE= 2												
ANO_IFC	PROJETO	COD_TALHAO	CICLO	ROT	IDADE	MG	PARCELA	COVA	FUSTE	FILA	DAP	HT CODIGOS
1	2006	PROJETO 0004	21	NA	1 43,50000	4	6	9	2	2	6,8	13,7 N
2	2007	PROJETO 0004	21	2	1 55,22793	4	6	9	2	2	6,8	14,1 N
3	2008	PROJETO 0004	21	2	1 65,90554	4	6	9	2	2	0,0	0,0 N
4	2009	PROJETO 0004	21	2	1 78,42300	4	6	9	2	2	0,0	0,0 M

Para estes casos é necessário identificar a causa, podendo ser erro cometido durante a gravação ou digitação do registro da medição, e até mesmo a manipulação de dados.

2.4.3.1. Árvores com DAP2 < DAP1 entre medições consecutivas i e i+1, sendo i o ano da medição

O comportamento esperado do DAP em medições imediatamente subsequentes, ou seja, consecutivas, é que seu valor seja crescente para uma mesma árvore, ou seja, ela cresce em diâmetro. O mesmo comportamento é esperado para altura. Contudo, erros de digitação, de medição em campo podem ocorrer, incluindo árvores cujo DAP medido em uma idade posterior é menor que o DAP de uma idade inferior.

O resultado desta consistência é apresentado na Tabela 12.

Tabela 12 – Árvores com DAP2 < DAP1 entre medições consecutivas.

FUSTE		CODIGOS	ANO1	ANO2	I1	I2	DAP1	DAP2	HT1	HT2	Δ
29	1	D	2008	2009	93	104	26,4	12,6	37,3	20,02	-13,8
54	1	D	2008	2009	93	104	26,1	13,1	37,3	34,3	-13,0
12	1	N	2007	2008	79	90	23,5	11,5	32,5	23	-12,0
21	1	D	2008	2009	93	104	24,9	16,2	29,8	24,04	-8,7
37	1	D	2008	2009	93	104	28,5	22	36,4	28,46	-6,5
5	1	N	2011	2012	47	57	15,6	10,7	15,99	18,06	-4,9
7	1	N	2009	2010	103	114	18,3	13,8	31,3	34,9	-4,5
13	1	D	2009	2010	51	62	20,1	15,7	27,3	31,6	-4,4
40	1	N	2006	2007	78	89	18,5	14,8	26,5	27,9	-3,7
4	1	N	2011	2012	67	80	18,7	15,2	27,48	25,08	-3,5
41	1	N	2007	2008	89	100	18,6	15,3	28,2	28,3	-3,3
12	1	N	2011	2012	24	34	13,8	10,6	16,13	19,61	-3,2
42	1	N	2007	2008	89	100	15,6	12,9	27,9	21,3	-2,7
11	1	N	2009	2010	63	74	22,5	19,9	26,9	28,9	-2,6
24	1	N	2011	2012	46	56	13,92	11,4	16,65	21,15	-2,5
18	1	N	2008	2009	90	103	18,6	16,1	29	31,7	-2,5
22	1	T	2011	2012	46	56	16,2	13,7	17,71	23,96	-2,5
17	1	T	2011	2012	46	56	17,52	15,1	18,36	25	-2,4
15	1	N	2008	2009	31	42	14,4	12,1	13,4	17,9	-2,3
20	1	Q	2011	2012	47	56	13,68	11,4	16,14	19,89	-2,3

Dependendo da magnitude da diferença, por exemplo, Δ (DAP2 – DAP1) de até é 2 cm, pode-se considerar um erro de medição e optar por fazer a correção manual ou ajustar algum modelo de crescimento com dados da parcela para resolver o problema. Para diferenças maiores, como a apresentada no fuste 29, faz-se necessário avaliar profundamente.

Outro método que pode ser empregado é avaliar o percentual (%) de árvores em uma mesma parcela na mesma idade com valores Δ negativos. Caso o percentual seja alto, é possível que algum erro sistemático tenha sido cometido entre medições, por exemplo, troca da ordem das filas entre medições, de modo que na medição n , por exemplo, o caminhamento na parcela tenha sido realizado a partir do fuste 1 na fileira 1 até 15, mas na medição $n+1$ a ordem tenha sido invertida, ou seja, o caminhamento tenha iniciado pela fileira 15.

2.4.3.1. Número de fustes mortos (CODIGO = M) ao longo das medições

A Figura 18 apresenta um resumo do número de árvores mortas por parcela em medições diferentes.

Figura 18 – Identificação de número de fustes mortos (CODIGO = M) em cada medição

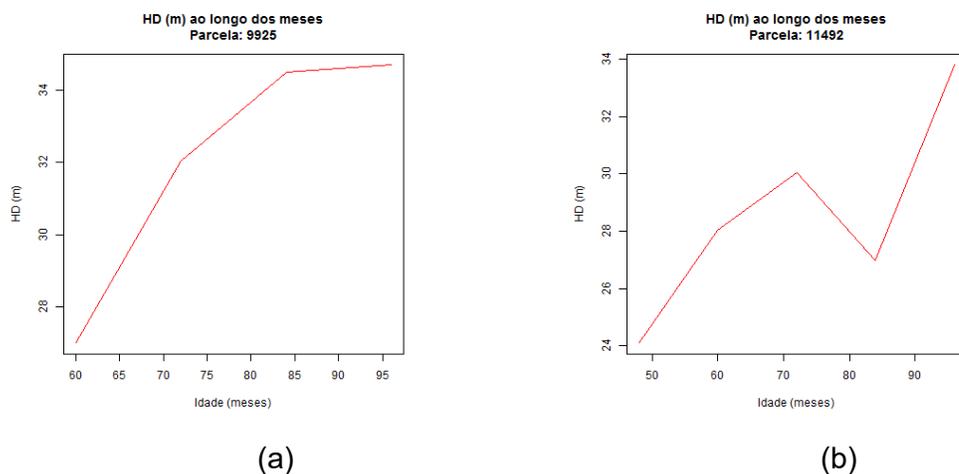
	2006	2007	2008	2009	2010	2011	2012
01	8	0	13	14	0	0	0
02	1	0	0	45	0	0	0
03	8	0	19	22	0	0	0
04	0	19	0	0	0	0	0
05	0	14	0	0	0	0	0
06	30	0	0	0	0	0	0
07	10	15	0	0	0	0	0
08	0	0	3	13	14	0	0
09	0	0	0	12	14	0	0
10	0	0	17	17	17	0	0
11	0	0	0	0	0	1	0
12	0	0	0	0	0	1	0

Observa-se em alguns casos um número alto de mortalidades, o que pode indicar um sinistro florestal, ou um erro sistemático de medição, esse último menos provável.

2.4.3.1. Gráficos de interação com a média de uma variável ao longo das medições

Gráficos de interação são um instrumento útil por permitir avaliar graficamente o comportamento médio de uma variável de interesse ao longo de medições consecutivas, como pode ser observado na Figura 19.

Figura 19 – Altura dominante de duas parcelas



Na Figura 19 (b) nota-se um comportamento irregular da variável altura dominante próximo aos 85 meses, o que indica a necessidade de avaliação.

2.4.4. Avaliação dos resultados

A avaliação das alternativas foi realizada pelas seguintes estatísticas:

a) Coeficiente de correlação, dado pela expressão (PIETILÄ et al., 2010; MÔRA, 2015):

$$r_{y\hat{y}} = \frac{n^{-1}(\sum(\hat{y}_i - \hat{y}_m)(y_i - \bar{y}))}{\sqrt{(n^{-1}\sum(\hat{y}_i - \hat{y}_m)^2)(n^{-1}\sum(y_i - \bar{y})^2)}} \quad (34)$$

b) BIAS(%) (UZOH; OLIVER, 2008):

$$BIAS = \frac{\sum_{i=1}^n (\hat{Y} - Y)}{n} \quad (35)$$

$$BIAS(\%) = \frac{BIAS}{\sum_{i=1}^n Y / n} \cdot 100 \quad (36)$$

c) Raiz quadrada do erro quadrático médio (RMSE %) (*Root Mean Square Error*) (ISLAM et al., 2009; ZHAO et al., 2013; ARIAS-RODIL et al., 2015):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n}} \quad (37)$$

$$RMSE(\%) = \frac{RMSE}{\sum_{i=1}^n Y / n} \cdot 100 \quad (38)$$

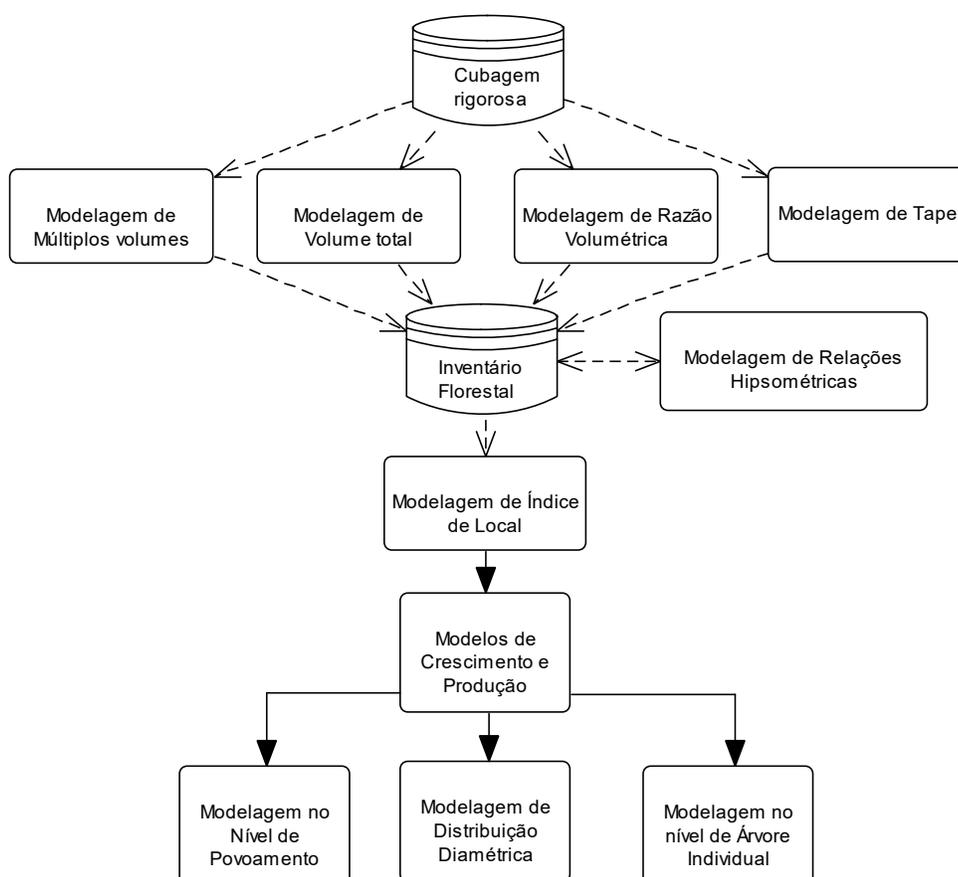
De forma a complementar a análise, foram elaborados gráficos relacionando os volumes observados com os volumes estimados pelas alternativas e gráficos de resíduos percentuais, obtidos pela expressão (MENDONÇA, 2010; MARTINS et al., 2014):

$$res(\%) = \left(\frac{y_i - \hat{y}_i}{y_i} \right) \cdot 100 \quad (39)$$

2.4.5. Capacidade produtiva, altura total e volume por árvore

Para proceder com o ajuste de modelos de crescimento e produção, faz-se necessária a disponibilidade de uma base de dados de inventário florestal, que dentre outras informações, contenha as alturas totais (observadas e estimadas), com a classificação de capacidade produtiva realizada, bem como o volume calculado, dependendo do nível de modelagem empregado. A Figura 20 apresenta o fluxo de dados até que os mesmos possam ser utilizados pelos modelos de crescimento e produção.

Figura 20 – Preparação de dados para uso nos modelos de crescimento e produção.



A modelagem de relações hipsométricas foi realizada usando uma abordagem com efeitos aleatórios, considerando a parcela como variável aleatória para agrupamento. Em seguida foi realizada a classificação da capacidade produtiva dos povoamentos pelo método da curva guia. A cálculo dos volumes individuais deu-se por uma equação de volume obtida previamente, sendo todas as etapas realizadas detalhadas a seguir.

A capacidade produtiva foi definida por índices de local estabelecidos pelo método da curva-guia para uma idade-índice de 60 meses. Os modelos da Tabela 13 foram ajustados, sem qualquer estratificação.

Tabela 13 – Modelos para classificação de capacidade produtiva

Identificação	Função	Função Curva-Guia	
Schumacher Linear	$\ln(Hd) = \beta_0 + \beta_1 \cdot \left(\frac{1}{I}\right) + \varepsilon_i$	$\ln(Hd) = \ln(S) + \beta_1 \cdot \left(\frac{1}{I} - \frac{1}{I_i}\right)$	(40)
Schumacher Não-Linear	$Hd = \beta_0 \cdot e^{\left(\frac{\beta_1}{I}\right)} \cdot \varepsilon_i$	$Hd = S \cdot e^{\beta \left(\frac{1}{I} - \frac{1}{I_i}\right)}$	(41)
Chapman- Richards	$Hd = \beta_0 \cdot \left[1 - e^{-kI}\right]^{\beta_1} \cdot \varepsilon_i$	$Hd = S \cdot \left(\frac{1 - e^{-kI}}{1 - e^{-kI_i}}\right)^{\left(\frac{1}{1-m}\right)}$	(42)
Curtis	$Hd = \beta_0 + \beta_1 \cdot \left(\frac{1}{I}\right) + \varepsilon_i$	$Hd = S + \beta_1 \cdot \left(\frac{1}{I} - \frac{1}{I_i}\right)$	(43)
Clutter-Jones	$Hd = \beta_0 \cdot \left[1 + \beta_1 I^{\beta_2}\right]^{\beta_3} \cdot \varepsilon_i$	$Hd = S \cdot \left(\frac{1 + \beta_1 I^{\beta_2}}{1 + \beta_1 I_i^{\beta_2}}\right)^{\frac{1}{\theta}}$	(44)

em que: Hd = altura dominante; Ln = logaritmo neperiano; S = índice de sítio na idade base; β_i , k, θ e m = parâmetros dos modelos; I = idade do indivíduo no momento da medição da altura; I_i = idade-índice (60 meses); ε_i = erro aleatório.

A escolha do modelo foi realizada com base na RMSE (%) e na consistência entre as curvas de índices de local e as alturas dominantes observadas (CAMPOS; LEITE, 2013).

As equações resultantes dos ajustes dos modelos avaliados para classificar a capacidade produtiva são apresentadas na Tabela 14.

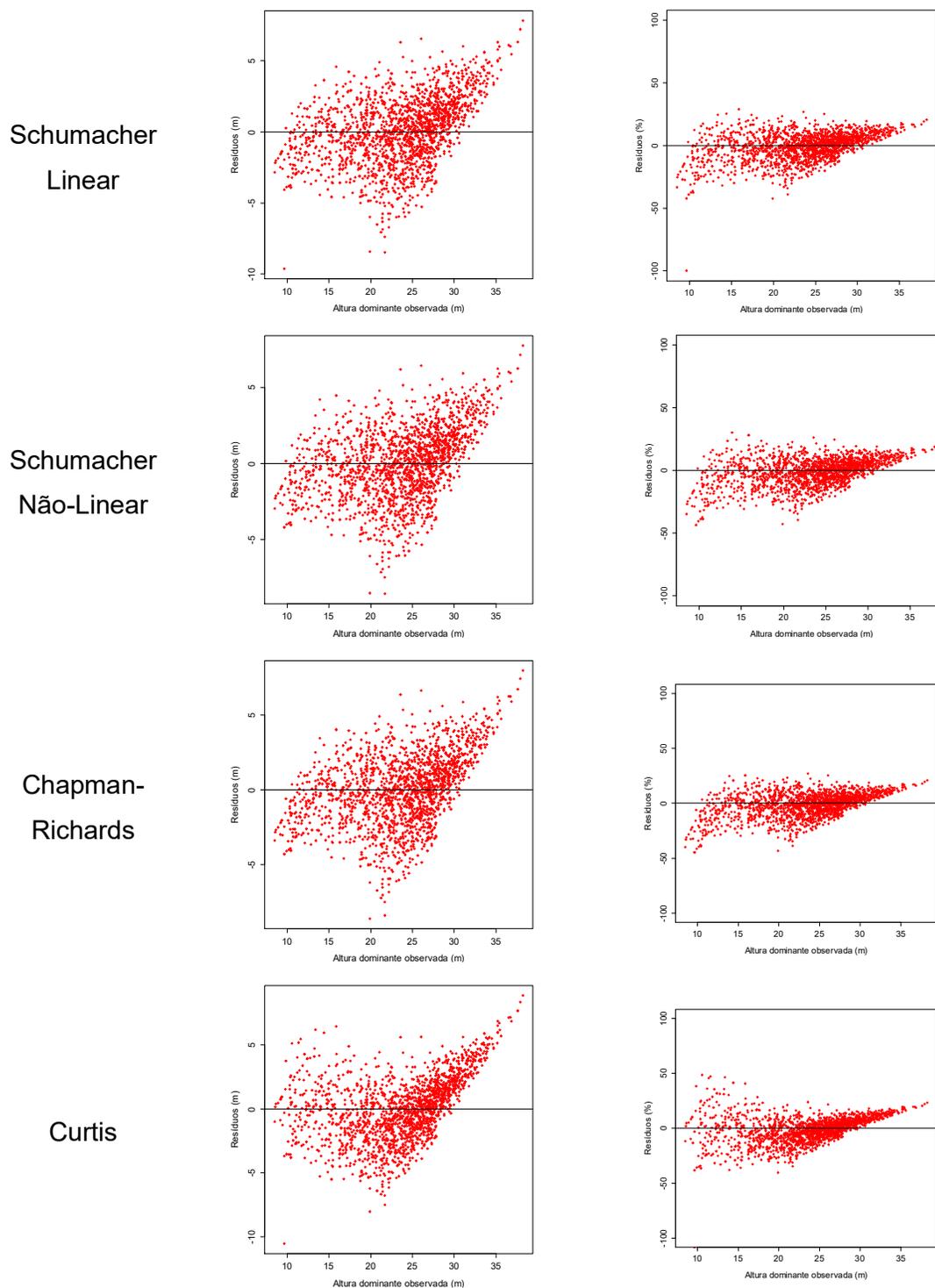
Tabela 14 – Equações ajustadas para os modelos de classificação de capacidade produtiva

Identificação	Equação	R ²	RMSE	RMSE (%)	
Schumacher Linear	$\ln(Hd) = 3.721 - 27,46 \cdot \left(\frac{1}{I}\right)$	0,8272	2,346	9,96	(45)
Schumacher Não-Linear	$Hd = 41.35701 \cdot e^{\left(\frac{-27.24102}{I}\right)}$	0,8347	2,295	9,74	(46)
Chapman- Richards	$Hd = 32,65 \cdot \left[1 - e^{-0,033 \cdot I}\right]^{1,505}$	0,8276	2,343	9,95	(47)
Curtis	$Hd = 35,538 + -552,695 \cdot \left(\frac{1}{I}\right)$	0,8115	2,450	10,41	(48)
Clutter- Jones	$Hd = 41,23 \cdot \left[1 + 1,14 \cdot I^{-1,017}\right]^{-25,54}$	0,8342	2,298	9,76	(49)

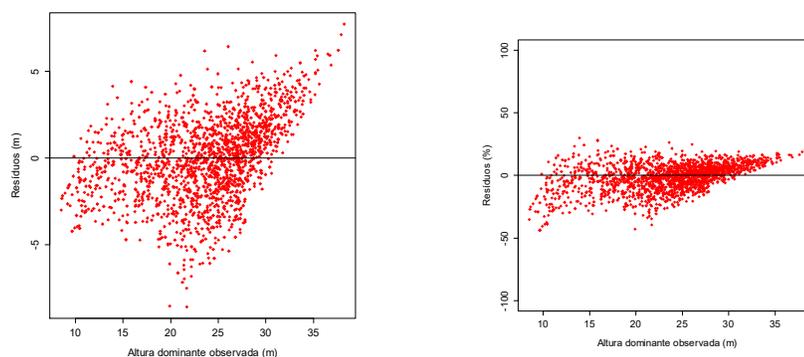
Como pode ser observado o modelo Schumacher não linear, representado pela equação (46) apresentou maior coeficiente de determinação e menor erro padrão relativo. Observa-se na Figura 21 que os resíduos apresentam tendência e comportamento similar.

Nota-se uma tendência dos modelos em subestimar arvores maiores que 30 m de altura.

Figura 21 – Distribuição dos resíduos absolutos (esquerda) e percentuais (direita) para os modelos de classificação de produtividade avaliados.



Clutter-Jones



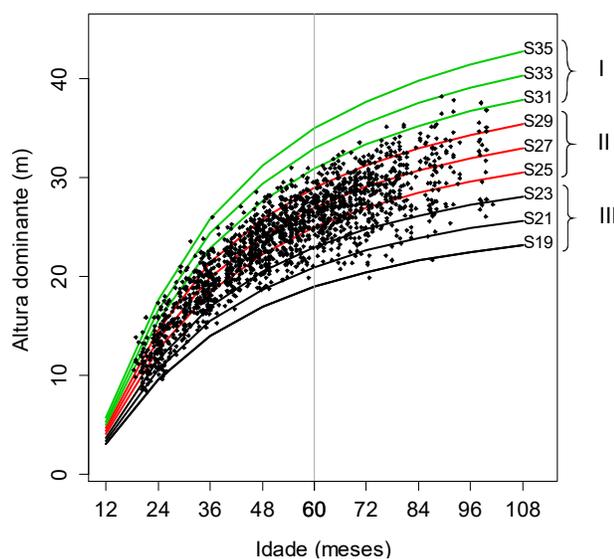
Fonte: O autor.

Com base no modelo definido e da sua equação, considerando-se a idade-índice de 60 meses, procedeu-se a construção das curvas de índices de local, utilizando a seguinte expressão:

$$Hd = S \cdot e^{\left(-27,24102 \left(\frac{1}{I} - \frac{1}{60}\right)\right)}$$

Obteve-se a curva-guia (S = 27) e o feixe de índices de local apresentados na Figura 22.

Figura 22 – Curvas de índice de local, para povoamento de eucalipto na respectiva idade índice de 60 meses.



De acordo com as curvas geradas, foram estabelecidos os limites para a classificação das parcelas nas classes de produtividade, conforme apresentado na Tabela 15.

Tabela 15 – Classificação de capacidade produtiva obtida.

Classe de Capacidade Produtiva	Limites de Inclusão de Hd (m) aos 60 meses
I - Alta (S = 33)	30 – 36
II - Média (S = 27)	24 – 30
III - Baixa (S = 21)	18 – 24

Cabe destacar que é na classe de produtividade média (S = 27) onde está a maior parte das parcelas utilizadas no ajuste (81% das parcelas), seguido pela classe de produtividade baixa (S = 21) com 14% das parcelas e a alta (S = 33) com 5% do total de parcelas.

2.4.5.1. Altura total

Para a obtenção das alturas totais dos fustes cujas alturas não foram medidas realizou-se a estimativa com redes neurais artificiais (RNA) treinadas do tipo *perceptron* de múltiplas camadas, comumente conhecidas como MLP (*Multilayer Perceptron*), compostos por duas camadas de neurônios artificiais que realizam o processamento dos dados (camada intermediária e camada de saída) e uma camada de neurônios artificiais que utilizam os dados (camada de entrada) e os entrega à camada intermediária. O pacote *nnet* (VENABLES; RIPLEY, 2002) da linguagem R (R DEVELOPMENT CORE TEAM, 2005) foi utilizado para obtenção das redes neurais artificiais.

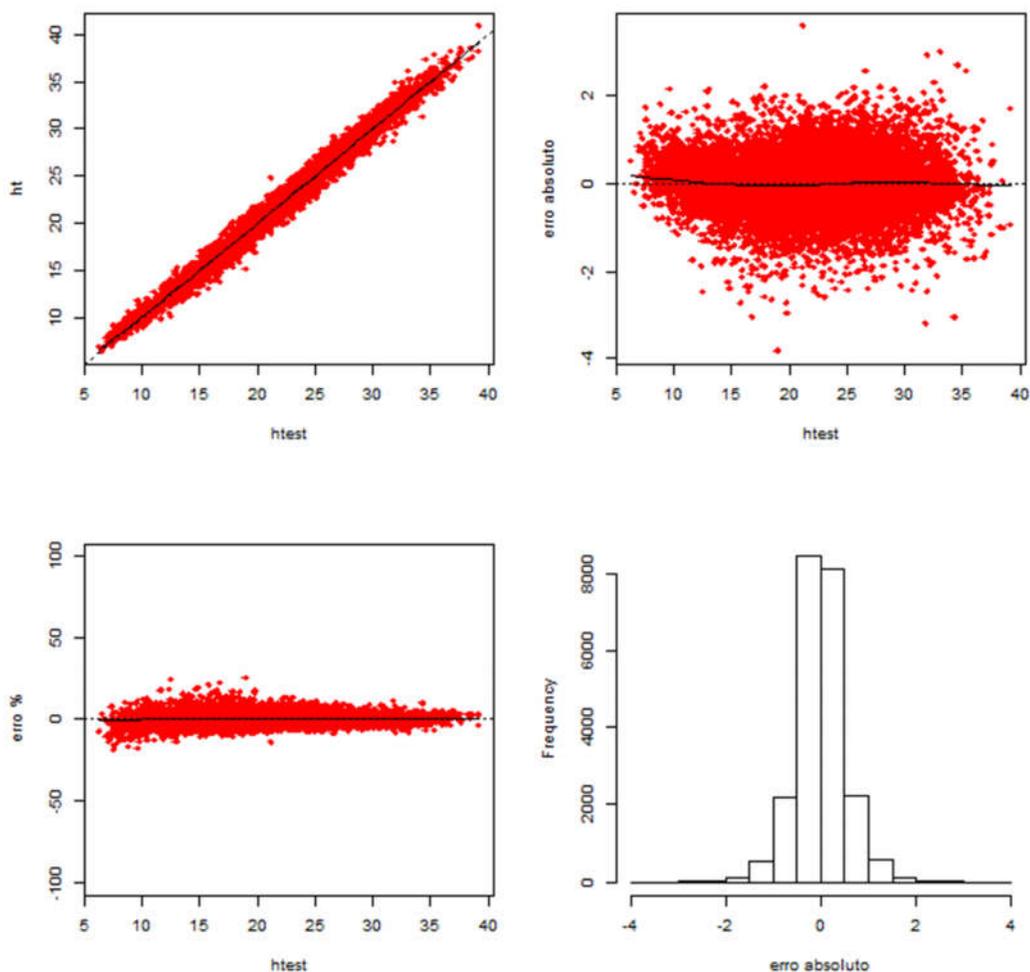
A RNA foi selecionada pela avaliação da correlação entre a altura observada e a altura estimada pelas redes, na raiz quadrada do erro quadrático médio (em porcentagem), no BIAS (%), na análise gráfica dispersão dos erros percentuais, histograma de frequência dos erros percentuais. A rede selecionada foi aplicada aos dados das demais árvores para estimar sua altura total. Todas as possíveis combinações de variáveis de entrada foram testadas, sendo a combinação que melhor expressou a relação hipsométrica: $H = f(Hd, dap, Idade, material\ genético)$. A arquitetura da rede resultante foi 4-8-1, sendo quatro neurônios na camada de entrada, oito na camada intermediária e uma saída (altura total). O procedimento apresentado é utilizado nos processos de inventário florestal nas empresas florestais. As estatísticas de ajuste são apresentadas na Tabela 16.

Tabela 16 – Estatísticas de ajuste do modelo hipsométrico utilizado para estimar altura total não medidas dos fustes

Correlação	RMSE	RMSE (%)	BIAS	BIAS (%)
0,995	0,5301	2,23	1,35 ⁻¹³	5,87 ⁻¹³

Os gráficos com as alturas observadas (y) e estimadas (x), resíduos absolutos e percentuais, e o histograma de resíduos é apresentado na Figura 23.

Figura 23 – Gráficos de alturas totais observadas e estimadas, resíduos absolutos e percentuais e histograma de resíduos.



2.4.5.2 Volume

O volume total com casca (V) das árvores individuais foi obtido por meio da seguinte equação de volume ajustada previamente:

$$V = e^{-10.139975493 + 1.868359307 \ln(DAP) + 1.0777866 \ln(Ht)}$$

Realizado o cálculo do volume individual pela equação de volume apresentada, obteve-se os volumes individuais cujo resumo é apresentado na Tabela 17.

Tabela 17 – Estatísticas do volume individual apresentado por idade

Idade (meses)	Volume		
	Mínimo	Médio	Máximo
24	0,0065	0,0532	0,1601
36	0,0086	0,1236	0,3211
48	0,0117	0,1905	0,5096
60	0,0115	0,2448	0,7140
72	0,0117	0,2924	0,8335
84	0,0157	0,3470	0,9650
96	0,0170	0,3989	0,9829

Idade em meses, volume expresso em m³;

3. REFERÊNCIAS BIBLIOGRÁFICAS

AMADOR, J. P. **Modelos mistos no ajuste de curvas de crescimento de *Ricinus communis* L.**, 2010. 141f. Tese (Doutorado em Agronomia) – Universidade Federal de Santa Maria, Santa Maria, 2010.

ANDRIOTTI, J. L. S. **Técnicas estatísticas aplicáveis a tratamento de informações oriundas de procedimentos laboratoriais.** Porto Alegre, 2005.

ARAÚJO JÚNIOR, C. A.; NOGUEIRA, G. S.; OLIVEIRA, M. L. R. DE; et al. Projeção da distribuição diamétrica de povoamentos de eucalipto em diferentes amplitudes de classe. **Pesq. agropec. bras., Brasília**, v. 45, n. 11, p. 1275–1281, 2010. SciELO Brasil.

ARIAS-RODIL, M.; CASTEDO-DORADO, F.; CÁMARA-OBREGÓN, A.; DIÉGUEZ-ARANDA, U. Fitting and calibrating a multilevel mixed-effects stem taper model for *Maritime* Pine in NW Spain. **PloS one**, v. 10, n. 12, p. e0143521, 2015. Public Library of Science. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143521>>. Acesso em: 31/12/2015.

BAILEY, R. L.; DELL, T. R. Quantifying diameter distributions with the Weibull function. **Forest Science**, v. 19, n. 2, p. 97–104, 1973. Society of American Foresters.

BARBOSA, M. **Uma abordagem para análise de dados com medidas repetidas utilizando modelos lineares mistos**, 2009. 118f. Dissertação (Mestrado em Agronomia) - Universidade de São Paulo, 2009.

BASTIANI, F. DE. **Influência local em modelos espaciais lineares com distribuição da família de contornos elípticos**, 2012. 137f. Dissertação (Mestrado em Estatística) – Universidade Federal de Pernambuco, Recife, 2012.

BATES, D.; MÄCHLER, M.; BOLKER, B.; WALKER, S. Fitting linear mixed-effects models using lme4. **arXiv preprint arXiv:1406.5823**, 2014.

BATINI, C.; SCANNAPIECA, M. **Data Quality: Concepts, Methodologies and Techniques.** Roma: Springer-Verlag Berlin Heidelberg, 2006.

BATISTA, J. L. F. **A função Weibull como modelo para a distribuição de diâmetros de espécies arbóreas tropicais**, 1989. 116f. Dissertação (Mestrado em Ciências Florestais) – Escola Superior de Agricultura “Luiz de Queiroz”, da Universidade de São Paulo, Piracicaba, 1989.

BINOTI, D. H. B.; LEITE, H. G.; GUIMARÃES, D. P.; et al. Eficiência das funções weibull e hiperbólica para descrição de distribuições diamétricas de povoamentos de *Tectona grandis*. **Revista Árvore**, v. 35, n. 2, p. 299–306, 2011. Sociedade de Investigações Florestais. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-

67622011000200014&lng=en&nrm=iso&tlng=pt>. Acesso em: 14/5/2014.

BINOTI, D. H. B.; LEITE, H. G.; NOGUEIRA, G. S.; et al. Uso da função Weibull de três parâmetros em um modelo de distribuição diamétrica para plantios de eucalipto submetidos a desbaste. **Revista Árvore**, v. 34, n. 1, p. 147–156, 2010. SciELO Brasil.

BOHORA, S. B.; CAO, Q. V. Prediction of tree diameter growth using quantile regression and mixed-effects models. **Forest Ecology and Management**, v. 319, p. 62–66, 2014. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112714000917>>. Acesso em: 31/12/2015.

BROWN, H.; PRESCOTT, R. **Applied mixed models in medicine**. 2º ed. Chichester: John Wiley & Sons, 2006.

BUCKMAN, R. E. **Growth and yield of red pine in Minnesota**. U.S. Dept. of Agriculture, 1962.

BURKHART, H. E. Stand modelling for radiata pine in New Zealand. **New Zealand Journal of Forestry**, 1977. Disponível em: <<http://agris.fao.org/agris-search/search.do?recordID=NZ19800496000>>. Acesso em: 14/5/2014.

CALAMA, R.; MONTERO, G. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. **Silva Fennica**, v. 39, n. 1, p. 37–54, 2005.

CALEGARIO, N.; CALEGARIO, C. L. L.; MAESTRI, R.; DANIELS, R. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não lineares generalizados. **Scientia forestalis**, v. 12, n. 69, p. 38–50, 2005.

CALEGARIO, N.; MAESTRI, R.; LEAL, C. L.; DANIELS, R. F. Estimativa do crescimento de povoamentos de *Eucalyptus* baseada na teoria dos modelos não lineares em multinível de efeito misto. **Ciência Florestal**, v. 15, n. 3, p. 285–292, 2005.

CAMARINHA FILHO, J. A. **Modelos lineares mistos: estruturas de matrizes de variâncias e covariâncias e seleção de modelos**, 2002. 85f. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2002.

CAMARINHA FILHO, J. A. **Nota Metodológica sobre modelos lineares mistos**. Curitiba, 2003.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração Florestal: Perguntas e Respostas**. 4º ed. Viçosa, MG, 2013.

CAMPOS, J. C. C.; ROSAS, M. P.; LEITE, H. G. Comparação de alternativas de determinação da idade técnica de corte em plantações de eucalipto. **Revista**

Árvore, v. 29, n. 1, p. 37–49, 1996.

CARVALHO, S. DE P. C. E.; CALEGARIO, N.; SILVA, F. F. E.; et al. Modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* clonal. **CERNE**, v. 17, n. 4, p. 541–548, 2011. Universidade Federal de Lavras.

CASTRO, R. V. O. **Modelagem do crescimento em nível de árvores individuais utilizando redes neurais e autômatos celulares**, 2011. 80f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Viçosa, Viçosa, 2011.

CASTRO, R. V. O.; SOARES, C. P. B.; MARTINS, F. B.; LEITE, H. G. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos. **Pesquisa Agropecuária Brasileira**, v. 48, n. 3, p. 287–295, 2013. SciELO Brasil.

CHEN, W.; WANG, Z.; ZHOU, J. Large-scale L-BFGS using MapReduce. In: Z. Ghahramani; M. Welling; C. Cortes; N. D. Lawrence; K. Q. Weinberger (Orgs.); *Advances in Neural Information Processing Systems*. **Anais...** . p.1332–1340, 2014. Curran Associates, Inc.

CLUTTER, J. L. Compatible growth and yield models for loblolly pine. **Forest Science**, v. 9, n. 3, p. 354–371, 1963. Society of American Foresters.

COSTA, S. C. DA. **Modelos lineares generalizados mistos para dados longitudinais**, 2003. 110f. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura Luiz de Queiroz, 2003.

COSTA, T. R. DA. **Modelos Lineares Mistos: Uma aplicação na produção de leite de vacas da raça SINDI**, 2010. 64f. Dissertação (Mestrado em Biometria e Estatística Aplicada) – Universidade Federal Rural de Pernambuco, 2010. Recife, 2010.

COUTO, H. T. Z. DO; BASTOS, N. L. M.; LACERDA, J. S. Comparação de dois modelos de prognose de madeira de *Eucalyptus* para alto fuste e talhadia. **Revista Scientia Forestalis**, v. 41-42, p. 27–35, 1989.

CRECENTE-CAMPO, F.; TOMÉ, M.; SOARES, P.; DIÉGUEZ-ARANDA, U. A generalized nonlinear mixed-effects height–diameter model for *Eucalyptus globulus* L. in northwestern Spain. **Forest Ecology and Management**, v. 259, n. 5, p. 943–952, 2010. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112709008676>>. Acesso em: 31/12/2015.

CRUZ, J. P. DA; LEITE, H. G.; SOARES, C. P. B.; et al. Modelos de crescimento e produção para plantios comerciais jovens de *Tectona Grandis* em Tangará da Serra, Mato Grosso. **Revista Árvore**, v. 32, n. 5, p. 821–828, 2008. SciELO Brasil.

DALLA LANA, M.; LINS, C. F.; BRANDÃO, S.; et al. Distribuição diamétrica de *Eschweilera Ovata* em um fragmento de floresta ombrófila densa –Igarassu, PE.

FLORESTA, v. 43, n. 1, p. 59–68, 2013.

DEMIDENKO, E. **Mixed models: theory and applications with R**. 2º ed. John Wiley & Sons, 2013.

DIAMANTOPOULOU, M. J.; ÖZÇELİK, R.; CRECENTE-CAMPO, F.; ELER, Ü. Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. **Biosystems Engineering**, v. 133, p. 33–45, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1537511015000367>>. Acesso em: 7/1/2016.

DONG, L.; LIU, Z.; LI, F.; JIANG, L. Modelling primary branch growth based on a multilevel nonlinear mixed effects model: a *Pinus koraiensis* plantation case study in north-east China. **Southern Forests: a Journal of Forest Science**, 2015. Taylor & Francis. Disponível em: <<http://www.tandfonline.com/doi/abs/10.2989/20702620.2014.1001676#.VoWYfhfcwOk.mendeley>>. Acesso em: 31/12/2015.

DOWLE, M.; SHORT, T.; LIANOGLU, S.; SRINIVASAN, A. data.table: Extension of data.frame. Disponível em: <<http://cran.r-project.org/package=data.table>>. Acesso em: 31/12/2015.

EISFELD, R. DE L.; SANQUETTA, C. R.; ARCE, J. E.; MAESTRI, R.; WEBER, K. S. Modelagem do crescimento e da produção de *Pinus taeda* L. por meio de função probabilística. **Floresta**, v. 35, n. 2, 2005.

ERCANLI, I.; GUNLU, A.; BAŞKENT, E. Z. Mixed effect models for predicting breast height diameter from stump diameter of Oriental beech in Göldağ. **Scientia Agricola**, v. 72, n. 3, p. 245–251, 2015. Scientia Agricola. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-90162015000300245&lng=en&nrm=iso&tlng=en>. Acesso em: 31/12/2015.

FANG, Z. **A simultaneous system of linear and nonlinear mixed effects models for forest growth and yield prediction**, 1999. 171f. Dissertation. (PhD in Forest Science). University of Georgia. Athens, 1999.

FAUSTO, M. A.; CARNEIRO, M.; ANTUNES, C. M. DE F.; PINTO, J. A.; COLOSIMO, E. A. O modelo de regressão linear misto para dados longitudinais: uma aplicação na análise de dados antropométricos desbalanceados. **Cadernos de Saúde Pública**, 2008.

FIGURA, M. A. **A distribuição de Weibull na descrição da estrutura diamétrica de Eucalyptus grandis: um enfoque sobre o método dos momentos.**, 2010. 98f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2010.

FRAGA FILHO, C. V.; SIMIQUELI, A. P. .; ALTOE, W. . Fgmutils: Forest Growth Model Utilities. Disponível em: <<https://cran.r-project.org/web/packages/Fgmutils/index.html>>. Acesso em: 31/12/2015.

FU, L.; SUN, H.; SHARMA, R. P.; et al. Nonlinear mixed-effects crown width models for individual trees of Chinese fir (*Cunninghamia lanceolata*) in south-central China. **Forest Ecology and Management**, v. 302, p. 210–220, 2013. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112713001837>>. Acesso em: 31/12/2015.

GAŁECKI, A.; BURZYKOWSKI, T. **Linear Mixed-Effects Models Using R - A Step-by-Step Approach**. New York: Springer, 2013.

GARBER, S. M.; MAGUIRE, D. A. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. **Forest Ecology and Management**, v. 179, n. 1, p. 507–522, 2003. Elsevier.

GERTNER, G. Prediction bias and response surface curvature. **Forest science**, v. 37, n. 3, p. 755–765, 1991. Society of American Foresters.

GORGOSO, J. J.; ROJO, A.; CÁMARA-OBREGÓN, A.; DIÉGUEZ-ARANDA, U. A comparison of estimation methods for fitting Weibull, Johnson's SB and beta functions to *Pinus pinaster*, *Pinus radiata* and *Pinus sylvestris* stands in northwest Spain. **Forest Systems**, v. 21, n. 3, p. 446–459, 2012. Disponível em: <<http://revistas.inia.es/index.php/fs/article/view/2736>>. Acesso em: 27/9/2014.

GORGOSO-VARELA, J. J.; ROJO-ALBORECA, A. Use of Gumbel and Weibull functions to model extreme values of diameter distributions in forest stands. **Annals of Forest Science**, v. 71, n. 7, p. 741–750, 2014. Disponível em: <<http://link.springer.com/10.1007/s13595-014-0369-1>>. Acesso em: 27/9/2014.

GOUVEIA, J. F. **Modelos volumétricos mistos aplicados em clones de eucalyptus no polo gesseiro do Araripe-PE**, 2013. 131f. Tese (Doutorado em Biometria e Estatística Aplicada) – Universidade Federal Rural de Pernambuco, Recife, 2013.

GRÉGOIRE, T. G.; SCHABENBERGER, O.; BARRETT, J. P. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. **Canadian Journal of Forest Research**, v. 25, n. 1, p. 137–156, 1995.

GROTHENDIECK, G. sqldf: Perform SQL Selects on R Data Frames. Disponível em: <<https://cran.r-project.org/web/packages/sqldf/index.html>>. Acesso em: 31/12/2015.

GUIMARÃES, D. P. **Desenvolvimento de um modelo de distribuição diamétrica de passo invariante para prognose e projeção da estrutura de povoamentos de eucalipto**, 1994. 160 f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 1994.

HASENAUER, H. Princípios para a modelagem de ecossistemas florestais. **Ciência & Ambiente, Santa Maria**, v. 20, p. 53–69, 2000.

HERSHBERGER, S. L.; MOSKOWITZ, D. S. **Modeling intraindividual variability**

with repeated measures data: Methods and applications. London: Psychology Press, 2002.

ISLAM, M. N.; KURTTILA, M.; MEHTÄTALO, L.; HAARA, A. Analyzing the effects of inventory errors on holding-level forest plans: the case of measurement error in the basal area of the dominated tree species. **Silva Fennica**, v. 43, n. 1, p. 71–85, 2009. Finnish Society of Forest Science.

LEITE, H. G. **Ajuste de um modelo de estimação de frequência e produção por classe de diâmetro, para povoamentos de Eucalyptus saligna Smith**, 1990. 97f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 1990.

LEITE, H. G.; ALCÂNTARA, A. E. M. DE; BINOTI, D. H. B.; et al. Comparação entre predição e projeção da distribuição de diâmetros de povoamentos de eucalipto submetidos a desbastes. **Revista Árvore**, v. 37, n. 2, p. 321–328, 2013. Sociedade de Investigações Florestais. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000200013&lng=en&nrm=iso&tlng=pt>. Acesso em: 11/1/2016.

LIMA, C. R. DE A. **Gestão da qualidade dos dados e informações dos sistemas de informação em saúde: subsídios para a construção de uma metodologia adequada ao Brasil**, 2010. Escola Nacional de Saúde Pública Sérgio Arouca.

MACHADO, S. DO A.; AUGUSTYNCZIK, A. L. D.; NASCIMENTO, R. G. M.; et al. Funções de distribuição diamétrica em um fragmento de Floresta Ombrófila Mista. **Ciência Rural**, v. 39, n. 8, p. 2428–2434, 2009. SciELO Brasil.

MACKINNEY, A. L.; SCHUMACHER, F. X.; CHAIKEN, L. E. Construction of yield tables for nonnormal loblolly pine stands. **Journal of Agricultural Research**, v. 54, n. 7, p. 531–545, 1937.

MAGALHÃES, J. G. DE S. **Prognose e Regulação de Florestas de eucalipto utilizando Modelo de Crescimento e Inteligência Computacional**, 2014. 66f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Viçosa, Viçosa, 2014.

MALAVERRI, J. E. G.; MEDEIROS, C. B. Data Quality in Agriculture Applications. **GeolInfo. Anais...** . p.128–139, 2012.

MARTINS, F. B. **Modelagem de crescimento em nível de árvore individual para plantios comerciais de eucalipto**, 2011. 143f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2011.

MARTINS, F. B.; SOARES, C. P. B.; SILVA, G. F. DA. Individual tree growth models for eucalyptus in northern Brazil. **Scientia Agricola**, v. 71, n. 3, p. 212–225, 2014. SciELO Brasil.

MEHTÄTALO, L.; PELTOLA, H.; KILPELÄINEN, A.; IKONEN, V.-P. The Response of

Basal Area Growth of Scots Pine to Thinning: A Longitudinal Analysis of Tree-Specific Series Using a Nonlinear Mixed-Effects Model. **Forest Science**, v. 60, n. 4, p. 636–644, 2014. Society of American Foresters. Disponível em: <<http://www.ingentaconnect.com/content/saf/fs/2014/00000060/00000004/art00004>>. Acesso em: 31/12/2015.

MENDONÇA, A. R. DE E. **Modelagem não linear do crescimento e da produção de plantações florestais**, 2010. 84f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras, 2010.

MÔRA, R. **Funções de afilamento de forma variável e modelagem de efeitos mistos em fustes de *Pinus taeda* e *Eucalyptus saligna***, 2015. 277f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2015.

MUNRO, D. D. Forest growth models—a prognosis. **Growth models for tree and stand simulation**, , n. 30, p. 7–21, 1974. Research Note 30. Department of Forest Yield Research, Royal College of Forestry, Stockholm.

NEWTON, P. F.; LEI, Y.; ZHANG, S. Y. Stand-level diameter distribution yield model for black spruce plantations. **Forest Ecology and Management**, v. 209, n. 3, p. 181–192, 2005. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112705000290>>. Acesso em: 7/1/2016.

NUNES, A. M. R. **O modelo linear misto multinível na análise do efeito do desbaste de pinheiros na recuperação ecológica de uma pedreira calcária**, 2010. 104f. Dissertação (Mestrado em Bioestatística) – Universidade de Lisboa. Lisboa, Portugal, 2010.

OLIVEIRA, J. D. C. **Modelagem de Crescimento e Produção de povoamentos de eucalipto empregando mínimos quadrados generalizados**, 2012. 75f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Viçosa, Viçosa, 2012.

OLIVEIRA NETO, S. N. DE; REIS, G. G. DOS; REIS, M. DAS G. F.; LEITE, H. G.; NEVES, J. C. L. Crescimento e distribuição diamétrica de *Eucalyptus camaldulensis* em diferentes espaçamentos e níveis de adubação na região de Cerrado de Minas Gerais. **FLORESTA**, v. 40, n. 4, 2010.

PATTERSON, H. .; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v. 58, p. 545–554, 1971.

PEDRO, C.; SOARES, B.; LEITE, H. G.; et al. Especificação de um modelo de crescimento e produção florestal. **Revista Árvore**, v. 28, n. 6, p. 831–837, 2004. SciELO Brasil.

PIETILÄ, I.; KANGAS, A.; MÄKINEN, A.; MEHTÄTALO, L. Influence of growth prediction errors on the expected losses from forest decisions. **Silva Fennica**, v. 44, n. 5, p. 829–843, 2010.

PINHEIRO, J.; BATES, D. **Mixed-effects models in S and S-PLUS**. Springer Science & Business Media, 2006.

PINHEIRO, J.; BATES, D.; DEBROY, S.; SARKAR, D. nlme: Linear and Nonlinear Mixed Effects Models. Disponível em: <<http://cran.r-project.org/package=nlme>>. Acesso em: 31/12/2015.

PLÁCIDO, A. C. DE; BARTOSZECK, S.; MACHADO, S. DO A.; FIGUEIREDO FILHO, A.; OLIVEIRA, E. B. A distribuição diamétrica para bracingais em diferentes idades, sítios e densidades na região metropolitana de Curitiba. **Floresta**, v. 34, n. 3, 2004.

QUEIROZ, G. DE. **Métodos e interpretação de modelos estatísticos de análise de medidas repetidas: uma aplicação a ensaio clínico**, 2012. 132 f. Dissertação (Mestrado em Saúde Pública) – Escola Nacional de Saúde Pública Sergio Arouca. Rio de Janeiro, 2012.

R DEVELOPMENT CORE TEAM, R. D. C. R: A language and environment for statistical computing. **R foundation for Statistical Computing**, 2005. Vienna, Austria: sn.

RESENDE, M. D. V. DE; PRATES, D. F.; JESUS, A.; YAMADA, C. K. Estimção de componentes de variância e predição de valores genéticos pelo método da máxima verossimilhança restrita (REML) e melhor predição linear não viciada (BLUP) em *Pinus*. **Boletim de Pesquisa Florestal**, v. 32, n. 33, p. 18–45, 1996.

SALLES, T. T.; LEITE, H. G.; OLIVEIRA NETO, S. N.; et al. Modelo de Clutter na modelagem de crescimento e produção de eucalipto em sistemas de integração lavoura-pecuária-floresta. **Pesquisa Agropecuária Brasileira**, v. 47, n. 2, p. 253–260, 2012. SciELO Brasil.

SANTOS, E. M. DOS. **Crescimento e produção de plantios de paricá (*Schizolobium amazonicum* Huber ex. Ducke) sob diferentes espaçamentos**, 29. fev. 2012. Universidade Federal do Espírito Santo. Disponível em: <<http://www.bibliotecaflorestal.ufv.br//handle/123456789/4094>>. Acesso em: 14/5/2014.

SCHNEIDER, P. R.; FINGER, C. A. G.; BERNETT, L. G.; SCHNEIDER, P. S. P.; FLEIG, F. D. Estimativa dos parâmetros da função de densidade probabilística de 9

SILVA, G. F. DA; LEITE, H. G.; SOARES, C. P. B.; SILVA, M. L. DA. Influência de estimativas de produção de madeira em processos de regulação florestal utilizando programação linear. **Ciência Florestal**, v. 13, n. 1, p. 57–72, 2003.

SOARES, C. P. B.; GEZAN, S. A.; SILVA, G. F. DA; CASTRO, R. V. O. Individual-tree growth and mortality models for a semideciduous Atlantic forest in Brazil. **Australian Journal of Basic and Applied Sciences**, v. 9, p. 542–552, 2015.

SOARES, T. S.; LEITE, H. G.; SOARES, C. P. B.; VALE, A. B. DO. Projection of

yield and distribution diametrical of eucalyptus stands with different forms of weibull function. **Revista Árvore**, v. 35, n. 5, p. 1027–1032, 2011. SciELO Brasil.

SOARES, T. S.; LEITE, H. G.; VALE, A. B. DO; SOARES, C. P. B.; SILVA, G. F. DA. Avaliação de um modelo de passo invariante na predição da estrutura de um povoamento de *Eucalyptus sp.* **Revista Árvore**, v. 31, n. 2, p. 275–283, 2007. SciELO Brasil.

TIMILSINA, N.; STAUDHAMMER, C. L. Individual Tree-Based Diameter Growth Model of Slash Pine in Florida Using Nonlinear Mixed Modeling. **Forest Science**, v. 59, n. 1, p. 27–37, 2013. Society of American Foresters. Disponível em: <<http://www.ingentaconnect.com/content/saf/fs/2013/00000059/00000001/art00003>>. Acesso em: 31/12/2015.

TREVIZOL JÚNIOR, T. L. **Análise de um modelo compatível de crescimento e produção de *Eucalyptus grandis* (W. Hill ex Maiden)**, 1985. 74f. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, 1985.

UZOH, F. C. C. C.; OLIVER, W. W. Individual tree diameter increment model for managed even-aged stands of ponderosa pine throughout the western United States using a multilevel linear mixed effects model. **Forest ecology and management**, v. 256, n. 3, p. 438–445, 2008. Elsevier. Disponível em: <<http://www.treesearch.fs.fed.us/pubs/40984#.VoV-xYlcjD4.mendeley>>. Acesso em: 31/12/2015.

VANCLAY, J. K. Modelling forest growth and yield: applications to mixed tropical forests. **School of Environmental Science and Management Papers**, p. 537, 1994.

VENABLES, W. N.; RIPLEY, B. D. **Modern Applied Statistics with S**. 4^o ed. New York: Springer, 2002.

VISMARA, E. DE S.; MEHTÄTALO, L.; BATISTA, J. L. F. Linear mixed-effects models and calibration applied to volume models in two rotations of *Eucalyptus grandis* plantations. **Canadian Journal of Forest Research**, p. 132–141, 2015. NRC Research Press. Disponível em: <<http://www.nrcresearchpress.com/doi/abs/10.1139/cjfr-2014-0435#.VoWA8-JYWug>>. Acesso em: 31/12/2015.

WICKHAM, H. The split-apply-combine strategy for data analysis. **Journal of Statistical Software**, v. 40, n. 1, p. 1–29, 2011. Disponível em: <<http://www.jstatsoft.org/index.php/jss/article/view/v040i01/v040i01.pdf>>. Acesso em: 8/10/2015.

XU, H.; SUN, Y.; WANG, X.; et al. Nonlinear mixed-effects (NLME) diameter growth models for individual China-fir (*Cunninghamia lanceolata*) trees in southeast China. **PLoS ONE**, v. 9, n. 8, p. e104012, 2014. Public Library of Science. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104012>>. Acesso em: 31/12/2015.

YANG, Y.; HUANG, S.; MENG, S. Stem Profile Model Calibrated by Nonlinear Mixed-Effects Modeling. *Applied Statistics in Agriculture*. **Anais...** . p.141–155, 2009. Disponível em: <<http://newprairiepress.org/agstatconference/2009/proceedings/11>>. Acesso em: 2/1/2016.

ZHAO, L.; LI, C.; TANG, S. Individual-tree diameter growth model for fir plantations based on multi-level linear mixed effects models across southeast China. **Journal of Forest Research**, v. 18, n. 4, p. 305–315, 2013.

ZHU, C.; BYRD, R. H.; LU, P.; NOCEDAL, J. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. **ACM Transactions on Mathematical Software (TOMS)**, v. 23, n. 4, p. 550–560, 1997. ACM.

ZUUR, A.; IENO, E. N.; WALKER, N.; SAVELIEV, A. A.; SMITH, G. M. **Mixed effects models and extensions in ecology with R**. Springer Science & Business Media, 2009.

CAPÍTULO 1 – MODELAGEM DO CRESCIMENTO E PRODUÇÃO DE EUCALIPTO, EM NÍVEL DE POVOAMENTO, COM E SEM EFEITO ALEATÓRIO

1. INTRODUÇÃO

O processo decisório relacionado ao planejamento florestal é tarefa fundamental para gestores, pesquisadores e analistas de inventário florestal, sendo a modelagem estatística a principal abordagem utilizada para a tomada de decisão. Este processo deve responder a perguntas sobre: como, onde, quanto, o que e para quem produzir, o que torna necessário buscar conhecimento a respeito do crescimento e produção das florestas, dando ênfase à variável volume, uma vez que esta constitui a informação de produção florestal de uma região, a partir do volume individual (CALEGARIO et al., 2004; GARCÍA et al., 2011; CAO, 2014; HEVIA et al., 2015). Visando obter tais informações, faz-se necessário obter o estoque em crescimento por meio de medição e estimação de características das árvores e da área onde as plantações estão localizadas (BINOTI et al., 2013; XU et al., 2014).

Muitas são as abordagens utilizadas para quantificar o volume atual e futuro de madeira de povoamentos, com destaque para análise de regressão com variáveis como diâmetro, altura e idade das árvores relacionando-as com características como espaçamento, capacidade produtiva e outras. Contudo, modelos em nível de povoamento não fornecem explicações diretas sobre a variação de tamanho das árvores do povoamento – realizam a estimação do crescimento e a produção utilizando variáveis do povoamento tais como idade, área basal e índice de local, sendo o crescimento estimado por unidade de área independente da classe de tamanho da árvore (CAMPOS; LEITE, 2013).

O avanço dos recursos de computação resultou um maior poder de processamento, o que permitiu o uso de técnicas estatísticas mais complexas, dentre as quais se destacam os modelos de regressão não lineares generalizados, generalizados de efeito aleatório e de efeito aleatório (CARVALHO et al., 2011). Os mesmos autores citam trabalhos que envolvem a aplicação de tais modelos, como Budhathoki et al. (2008), Calegario et al. (2005a, b), Fang e Bailey (2001), Gregoire et al. (1995) e Pilar-Díaz e Couto (1999). Outras abordagens como o uso de redes neurais artificiais (CASTRO et al., 2013; BINOTI et al., 2013, 2015), aprendizado de máquina e técnicas de sensoriamento remoto tem ganhado espaço (GÖRGENS et al., 2015).

No Brasil, Mendonça (2010) propôs dois sistemas não lineares com adição de covariáveis para predição de área basal e do crescimento e da produção do volume total de povoamentos. Oliveira (2012) realizou ajustes de modelo de crescimento e produção em nível de povoamento pelo método dos mínimos quadrados generalizados, obtendo resultados similares à estimativa de produção pelo método dos mínimos quadrados ordinários, com o modelo de Clutter. Destaca-se ainda o uso de modelos lineares de efeitos aleatórios em multinível na estimativa de crescimento de povoamentos homogêneos de *Eucalyptus* em dois níveis de modelagem aleatória: unidades amostrais e árvores individuais dentro das unidades amostrais (CALEGARIO et al., 2005). Hall e Bailey (2001) apresentaram um método para estimar a altura total com dados longitudinais agrupados por localização, que permite a estimação da altura futura de árvores individuais e altura média de parcelas utilizando modelos não lineares mistos.

Um modelo de projeção de volume exige um conjunto de dados representativo para a etapa de ajuste dos dados, que apesar de sua estrutura hierárquica natural, representada pelas diferentes fazendas (ou projetos), talhões, parcelas, os modelos de regressão tradicionais não as consideram, não permitindo a variação dos parâmetros entre um ou mais grupos hierárquicos (VISMARA et al., 2015).

O uso de modelos mistos permite identificar as diferentes fontes estocásticas de variabilidade que não são explicadas pela parte fixa do modelo, dividindo-se a variância residual em diferentes componentes (CALAMA; MONTERO, 2005).

Na modelagem de efeitos aleatórios, os parâmetros de um modelo ajustado podem variar entre os grupos, sendo assim, o modelo pode ser calibrado para um determinado grupo afim de melhorar a qualidade das predições dentro do referido grupo (VISMARA et al., 2015).

Diante do exposto, o objetivo deste capítulo foi avaliar e comparar duas estratégias para a projeção de volume com casca para povoamentos de eucalipto, com e sem efeitos aleatórios.

2. MATERIAL E MÉTODOS

2.1. Dados utilizados

Os dados utilizados neste capítulo são descritos na seção 3.4 da tese. Foram selecionadas variáveis agrupadas por parcela e idade, e os volumes das árvores foi totalizado desta forma, obtendo-se a produção em m^3 . Em seguida, o volume do povoamento foi obtido em $m^3 ha^{-1}$. A base de dados com as variáveis pareadas na idade 1 e idade 2 foi criada com a função *createPairedData* do pacote *Fgmutils* (FRAGA FILHO et al., 2015), criado durante o desenvolvimento deste trabalho e disponibilizado nos repositórios da linguagem R.

2.2. Divisão da base de dados para ajuste e validação dos modelos

Com o objetivo de avaliar diferentes grupos na calibração de modelos mistos, os dados foram divididos utilizando-se dois critérios: no primeiro (critério 1), o objetivo foi avaliar o comportamento dos modelos para idades futuras, definindo a variável parcela como grupo. No segundo critério (critério 2), foram definidos grupos formados por variáveis relacionadas ao povoamento, disponíveis na base de dados do inventário.

De acordo com o critério 1, os dados foram divididos em dois conjuntos: o primeiro para ajuste, e o segundo para validação dos modelos. Do conjunto de ajuste participaram todas as parcelas com idade futura (idade 2) até 60 meses, e no segundo, de projeção (ou validação), foram selecionadas as mesmas parcelas do grupo de ajuste, porém com idade futura igual ou superior a 72 meses.

De acordo com o critério 2, o conjunto de parcelas foi dividido aleatoriamente também em dois grupos, sendo 70% das parcelas utilizadas para ajuste dos modelos e 30% para validação, ou seja, de acordo com este critério.

2.3. Alternativas de modelagem em nível de povoamento

Para realizar a projeção do volume em nível de povoamento, foram avaliadas duas estratégias, sendo a primeira delas o modelo de Clutter (1963), denominado a partir de agora de estratégia 1. Este modelo é composto pelas seguintes expressões:

$$\text{Ln}B_2 = \text{Ln}B_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) + \alpha_1 \left(1 - \frac{l_1}{l_2} \right) S + \varepsilon \quad (1)$$

$$\text{Ln}V_2 = \beta_0 + \frac{\beta_1}{l_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon \quad (2)$$

em que: B_2 é a área basal do povoamento na idade futura ($\text{m}^2 \text{ha}^{-1}$); B_1 é a área basal do povoamento na idade atual ($\text{m}^2 \text{ha}^{-1}$); l_1 é a idade atual, (meses); l_2 , a idade futura (meses); S é o índice de local; V_2 é o volume de madeira com casca do povoamento, na idade futura ($\text{m}^3 \text{ha}^{-1}$); com α_1 e α_2 , que são os coeficientes do modelo de área basal; $\beta_1 \dots \beta_4$ são os coeficientes do modelo volumétrico; e ε é o erro aleatório.

A segunda estratégia avaliada, denominada aqui de estratégia 2, se baseia no modelo proposto por Mendonça (2010), composto pelas seguintes expressões:

$$B_2 = B_1 + \frac{[-(\beta_{00} + \beta_{01}S)l_2 e^{-(\beta_{00} + \beta_{01}S)l_2} - e^{-(\beta_{00} + \beta_{01}S)l_2} + e^{-(\beta_{00} + \beta_{01}S)l_2} (\beta_{10} + \beta_{11}S)]}{\beta_{00} + \beta_{01}S} + \frac{[(\beta_{00} + \beta_{01}S)l_1 e^{-(\beta_{00} + \beta_{01}S)l_1} + e^{-(\beta_{00} + \beta_{01}S)l_1} - e^{-(\beta_{00} + \beta_{01}S)l_1} (\beta_{10} + \beta_{11}S)]}{\beta_{00} + \beta_{01}S} + \varepsilon_i \quad (3)$$

$$V_2 = \frac{\beta_{00} + \beta_{01}S}{1 + \exp\{[(\beta_{10} + \beta_{11}B_2) - l_2] / \beta_2\}} + \varepsilon_i \quad (4)$$

Tendo como referência as estratégias apresentadas, foram propostos dois conjuntos de alternativas de modelagem com e sem efeitos aleatórios, com o objetivo de verificar se a inclusão de efeitos aleatórios pode trazer ganhos em termos de acurácia na estimação do volume de povoamentos de eucalipto.

O primeiro conjunto, refere-se às alternativas de modelos mistos apresentadas na Tabela 1, que foram propostas para realização da calibração para

cada parcela, ou seja, utilizando a variável parcela como grupo de efeito aleatório. As alternativas 1_0 (estratégia 1) e 7_0 (estratégia 2) consideram apenas efeitos fixos e as alternativas 2_1 e 2_2 (estratégia 1) e 7_1 (estratégia 2) consideram efeitos aleatórios. Tais alternativas foram ajustadas e validadas com o conjunto de dados estabelecido no critério 1.

No caso das estratégias 1 e 2 de modelagem, testou-se o uso de efeitos aleatórios para a expressão de área basal (expressões 1 e 3), mas não foram encontrados resultados satisfatórios. Por este motivo, não foram propostas alternativas de modelos mistos envolvendo essas expressões. Em função disso, na Tabela 1 são apresentadas comparações de modelagem, com e sem efeitos aleatórios, levando-se em conta apenas a expressão de volume.

Tabela 1 – Alternativas de modelagem avaliadas para projeção do volume em nível de povoamento utilizando a parcela como grupo de efeitos aleatórios.

Alternativa	Modelo	
1_0	$\text{Ln}V_2 = \beta_0 + \frac{\beta_1}{I_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon$	(5)
2_1	$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1 + b_1}{I_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon$, com a variável parcela como variável de agrupamento de efeitos aleatórios	(6)
2_2	$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1}{I_2} + \beta_2 S + (\beta_3 + b_3) \text{Ln}B_2 + \varepsilon$, com a variável parcela como variável de agrupamento de efeitos aleatórios e $\text{Ln}B_2$ como inclinação aleatória	(7)
7_0	$V_2 = \frac{\beta_{00} + \beta_{01} S}{1 + \exp\{[(\beta_{10} + \beta_{11} B_2) - I_2] / \beta_2\}} + \varepsilon_i$	(8)
7_1	$V_2 = \frac{\beta_{00} + \beta_{01} S}{1 + \exp\{[(b_0 + \beta_{10}) + (b_1 + \beta_{11}) B_2] - I_2\} / \beta_2}$ + ε_i , com a variável parcela como variável de agrupamento de efeitos aleatórios e área basal na idade futura como inclinação aleatória	(9)

em que: S é o índice de local; B_2 área basal projetada para a idade futura; β - parâmetros fixos do modelo, b parâmetros aleatórios do modelo misto.

Objetivando avaliar alternativas além das propostas na Tabela 1, que não utilizassem a parcela como grupo de efeitos aleatórios na calibração, foram definidas outras variáveis e a partir delas, alternativas. Como opções ao uso da parcela,

apresentam-se o regime de manejo, espaçamento, o material genético, sendo estas variáveis candidatas a formarem grupos de efeitos aleatórios. Contudo, a base de dados utilizada neste trabalho apresenta apenas um regime de manejo e um tipo de espaçamento. A variável material genético foi considerada como variável de agrupamento de efeitos aleatórios (Tabela 2), e outras três variáveis foram consideradas para calibração das novas alternativas de modelagem mista.

O genótipo totaliza 20 grupos de efeitos aleatórios, enquanto a variável classeB1 é formada por valores da área basal na idade atual ($\text{m}^2 \text{ha}^{-1}$), agrupada em classes com amplitude de $1 \text{ m}^2 \text{ha}^{-1}$, produzindo 30 grupos. Considerando as classes formadas pela idade atual e futura (em meses), com amplitude de 2 meses, foram formados 36 grupos denominados classidade1, e classidade2, respectivamente.

Tabela 2 – Variáveis representando grupos de efeitos aleatórios utilizadas nas novas alternativas para calibração de modelos mistos.

Grupo proposto de efeitos aleatórios	Variável agrupada em classes	Amplitude da classe criada
<i>Material genético</i>	-	-
<i>classeB1</i>	Área basal na idade atual ($\text{m}^2 \text{ha}^{-1}$)	$1 \text{ m}^2 \text{ha}^{-1}$
<i>classidade1</i>	Idade atual (meses)	2 meses
<i>classidade2</i>	Idade futura (meses)	2 meses

Considerando os novos grupos de efeitos aleatórios apresentados na Tabela 2, foram propostas alternativas de modelagem de efeitos aleatórios apresentadas na Tabela 3. Dentre as onze alternativas apresentadas, as alternativas 2, 3, 4, 5, 6, 10 e 11 foram baseadas no modelo de Clutter (1963) (estratégia 1) e levam em conta efeitos aleatórios. Por outro lado, a alternativa 1 (expressão 10), também baseada no modelo de Clutter (1963), só considera efeitos fixos. Para a estratégia 2, baseada no modelo proposto por Mendonça (2010), foram propostas as alternativas 8 e 9, ambas com efeitos aleatórios, e a alternativa 7, apenas com efeitos fixos.

Tabela 3 – Alternativas de modelagem avaliadas para projeção de volume em nível de povoamento considerando os grupos de efeitos aleatórios propostos na Tabela 2.

Alternativa	Modelo
1	$\text{Ln}V_2 = \beta_0 + \frac{\beta_1}{I_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon \quad (10)$
2	$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1}{I_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon$, com a variável material genético como variável de agrupamento de efeitos aleatórios (11)
3	$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1}{I_2} + (\beta_2 + b_2)S + \beta_3 \text{Ln}B_2 + \varepsilon$, com a variável classe de área basal na idade atual (classeB1) como variável de agrupamento de efeitos aleatórios e S como inclinação aleatória (12)
4	$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1 + b_1}{I_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon$, com a variável classe de área basal na idade atual (classeB1) como variável de agrupamento de efeitos aleatórios e I_2^{-1} como inclinação aleatória (13)
5	$\text{Ln}V_2 = \beta_0 + \frac{\beta_1}{I_2} + \beta_2 S + (\beta_3 + b_3) \text{Ln}B_2 + \varepsilon$, com a variável classe de área basal na idade atual (classeB1) como variável de agrupamento de efeitos aleatórios e $\text{Ln}B_2$ como inclinação aleatória (14)
6	$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1}{I_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + \varepsilon$, com a variável classe de Idade 2 como variável de agrupamento de efeitos aleatórios (15)
7	$V_2 = \frac{\beta_{00} + \beta_{01}S}{1 + \exp\{[(\beta_{10} + \beta_{11}B_2) - I_2] / \beta_2\}} + \varepsilon_i \quad (16)$
8	$V_2 = \frac{\beta_{00} + b_0 + \beta_{01}S}{1 + \exp\{[(\beta_{10} + \beta_{11}B_2) - I_2] / \beta_2\}} + \varepsilon_i$, com a variável classeldade2 como variável de agrupamento de efeitos aleatórios (17)

$$V_2 = \frac{\beta_{00} + b_0 + \beta_{01}S}{1 + \exp\{[(\beta_{10} + \beta_{11}B_2) - I_2] / \beta_2\}} + \varepsilon_i, \text{ com a variável} \quad (18)$$

9 classeldade1 como variável de agrupamento de efeitos aleatórios

$$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1}{I_2} + \beta_2S + \beta_3\text{Ln}B_2 + b_1\hat{H}_2 + \varepsilon, \text{ com a variável} \quad (19)$$

10 classeldade1 como variável de agrupamento de efeitos aleatórios e altura média da parcela estimada na idade 2 e sítio como inclinações aleatórias.

$$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1}{I_2} + \beta_2S + \beta_3\text{Ln}B_2 + b_1\hat{H}_2 + b_2\ln(V1) + \varepsilon, \text{ com a} \quad (20)$$

11 variável classeldade1 como variável de agrupamento de efeitos aleatórios, altura média da parcela estimada na idade futura e log do volume m^3ha^{-1} na idade1 como variáveis de inclinação aleatória.

em que: I_1 é a idade atual, em meses; I_2 , a idade futura; S é o índice de local; V_2 é o volume de madeira com casca do povoamento, na idade futura ($\text{m}^3 \text{ha}^{-1}$); \hat{H}_2 = altura média estimada da parcela na idade 2 obtida por Pienaar e Schiver (1981); β - parâmetros fixos do modelo, b parâmetros aleatórios do modelo misto.

Os dados organizados de acordo com o critério 2, foram utilizados no ajuste e na validação das alternativas apresentadas na Tabela 3. As estimativas dos parâmetros do modelo de Clutter (1963) (alternativa 1) foram obtidas pelo métodos dos mínimos quadrados em dois estágios, utilizando a função *lm* do software R (R DEVELOPMENT CORE TEAM, 2005). Para as alternativas mistas, foram utilizadas as funções *lme* para ajuste de modelos lineares mistos e *nlme* para modelos não lineares mistos, disponibilizadas no pacote *nlme* (PINHEIRO et al., 2014) do software R. A função *gnls* do mesmo pacote foi utilizada para o ajuste da alternativa 7.

2.4. Avaliação dos resultados

A avaliação das alternativas foi realizada pelas seguintes estatísticas:

a) Coeficiente de correlação (PIETILÄ et al., 2010; MÔRA, 2015):

$$r_{\hat{y}y} = \frac{n^{-1}(\sum(\hat{y}_i - \hat{y}_m)(y_i - \bar{y}))}{\sqrt{(n^{-1}\sum(\hat{y}_i - \hat{y}_m)^2)(n^{-1}\sum(y_i - \bar{y})^2)}} \quad (21)$$

b) BIAS(%) (UZOH; OLIVER, 2008):

$$BIAS = \frac{\sum_{i=1}^n (\hat{Y} - Y)}{n} \quad (22)$$

$$BIAS(\%) = \frac{BIAS}{\sum_{i=1}^n Y / n} \cdot 100 \quad (23)$$

c) Raiz quadrada do erro quadrático médio relativo (RMSE %), (ISLAM et al., 2009; ZHAO et al., 2013; ARIAS-RODIL et al., 2015):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n}} \quad (24)$$

$$RMSE(\%) = \frac{RMSE}{\sum_{i=1}^n Y / n} \cdot 100 \quad (25)$$

De forma a complementar a análise, foram elaborados gráficos relacionando os volumes observados com os volumes estimados pelas onze alternativas e gráficos de resíduos percentuais, obtidos pela expressão (MENDONÇA, 2010; MARTINS et al., 2014):

$$res(\%) = \left(\frac{y_i - \hat{y}_i}{y_i} \right) \cdot 100 \quad (26)$$

3. RESULTADOS E DISCUSSÃO

Nessa seção os resultados são apresentados de acordo com os critérios empregados na divisão dos dados em grupos.

3.1. Critério 1 – Alternativas usando a variável parcela como grupo de efeito aleatórios

Com base nas estatísticas das equações apresentadas na Tabela 4, verifica-se que as alternativas avaliadas apresentaram desempenho semelhante em termos de acurácia, com vantagem para a alternativa 7_1 quando se considera as estatísticas $r_{y\hat{y}}$ e RMSE (%) nas classes de produtividade média (S = 27) e alta (S = 33). Em todas as classes de produtividade, esta alternativa também mostrou ser a menos viesada, o que pode ser constatado pelos valores de BIAS (%) mais próximos de zero.

Tabela 4 – Estatísticas utilizadas para avaliar o ajuste das alternativas que utilizam a parcela como grupo de efeitos aleatórios, apresentadas por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9485	9,62	0,6948
	S = 27	0,9487	8,99	-0,8188
	S = 33	0,9751	6,55	1,1939
2_1	S = 21	0,9773	6,43	0,3150
	S = 27	0,9785	5,92	-0,5129
	S = 33	0,9887	4,28	0,3260
2_2	S = 21	0,9821	5,69	0,2673
	S = 27	0,9808	5,64	-0,4976
	S = 33	0,9894	4,06	0,5260
7_0	S = 21	0,9506	9,40	0,8781
	S = 27	0,9514	8,65	-0,2067
	S = 33	0,977	5,85	1,2171
7_1	S = 21	0,9816	5,76	0,1168
	S = 27	0,9866	4,59	0,0092
	S = 33	0,9909	3,71	-0,0439

Ao realizar a validação dos modelos, isto é, aplicando-os no grupo de dados com as mesmas parcelas utilizadas no ajuste, porém com idade futura superior a 60 meses (critério 1 de validação), verifica-se o mesmo comportamento das estatísticas de ajuste para a alternativa 7_1 (estratégia 2). Entretanto, a alternativa 7_0

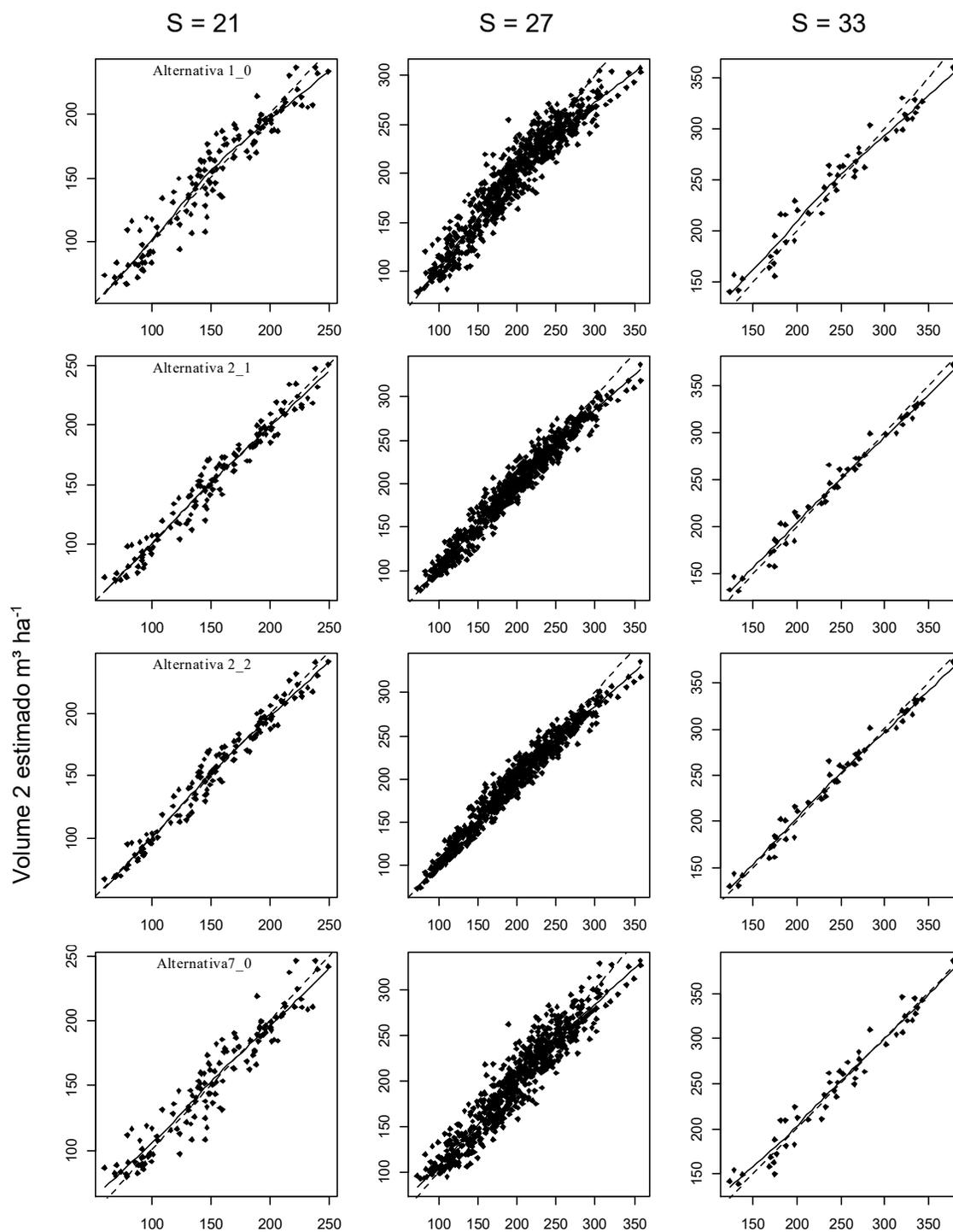
apresenta resultados mais acurados (RMSE (%) e $r_{\hat{y}y}$) e menos viesados (BIAS) na classe de produtividade Média (S = 27), porém, com maior viés na classe de produtividade Alta (S = 33), quando comparada à alternativa 7_1 (Tabela 5). As alternativas 1_0, 2_1 e 2_2 apresentaram resultados próximos para todas as estatísticas consideradas, assim como ocorreu no ajuste, com ligeira vantagem em termos de acurácia e ausência de viés para a alternativa 2_2.

Tabela 5 – Estatísticas utilizadas para avaliar as alternativas que utilizaram a parcela como grupo de efeitos aleatórios, apenas com parcelas usadas no ajuste, apresentadas por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}y}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9438	9,15	-5,5330
	S = 27	0,9358	11,12	-7,5976
	S = 33	0,9873	9,62	-5,9731
2_1	S = 21	0,9435	8,18	-5,1626
	S = 27	0,9597	9,50	-7,0135
	S = 33	0,996	8,24	-5,5657
2_2	S = 21	0,9608	7,80	-5,0451
	S = 27	0,9614	9,20	-6,4938
	S = 33	0,9954	7,83	-4,5947
7_0	S = 21	0,9513	5,61	1,3222
	S = 27	0,9346	6,41	0,4151
	S = 33	0,9883	4,58	2,6142
7_1	S = 21	0,9435	6,70	1,3544
	S = 27	0,9491	7,01	-1,0427
	S = 33	0,9829	7,80	-0,7443

Corroborando os resultados encontrados na Tabela 4 e na Tabela 5, verifica-se na Figura 1 e na Figura 2 desvios menores para a alternativa 7_1 no ajuste, e um menor viés, especialmente para a classe de produtividade média (S = 27). As alternativas 1_0, 2_1 e 2_2 apresentaram tendência de subestimação para parcelas com maior volume.

Figura 1 – Volumes estimados *versus* observados de acordo com o critério 1 de separação da base de dados, para as diferentes classes de qualidade de local.



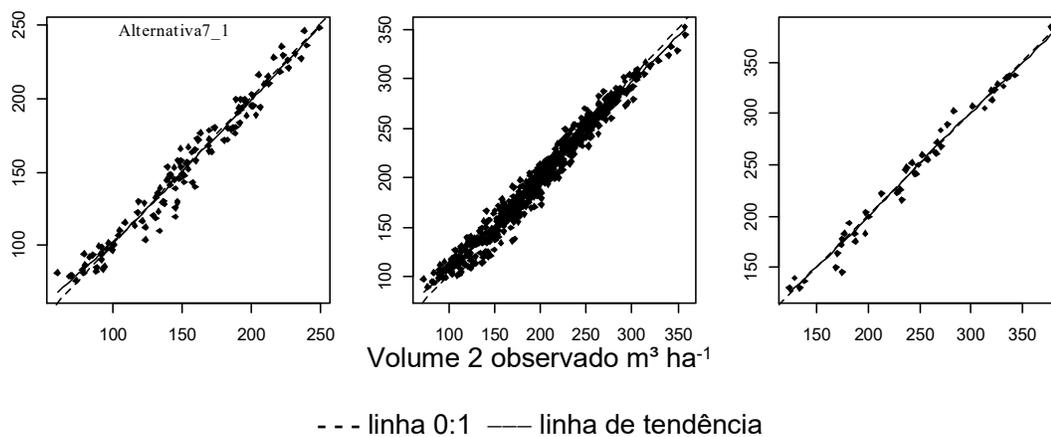
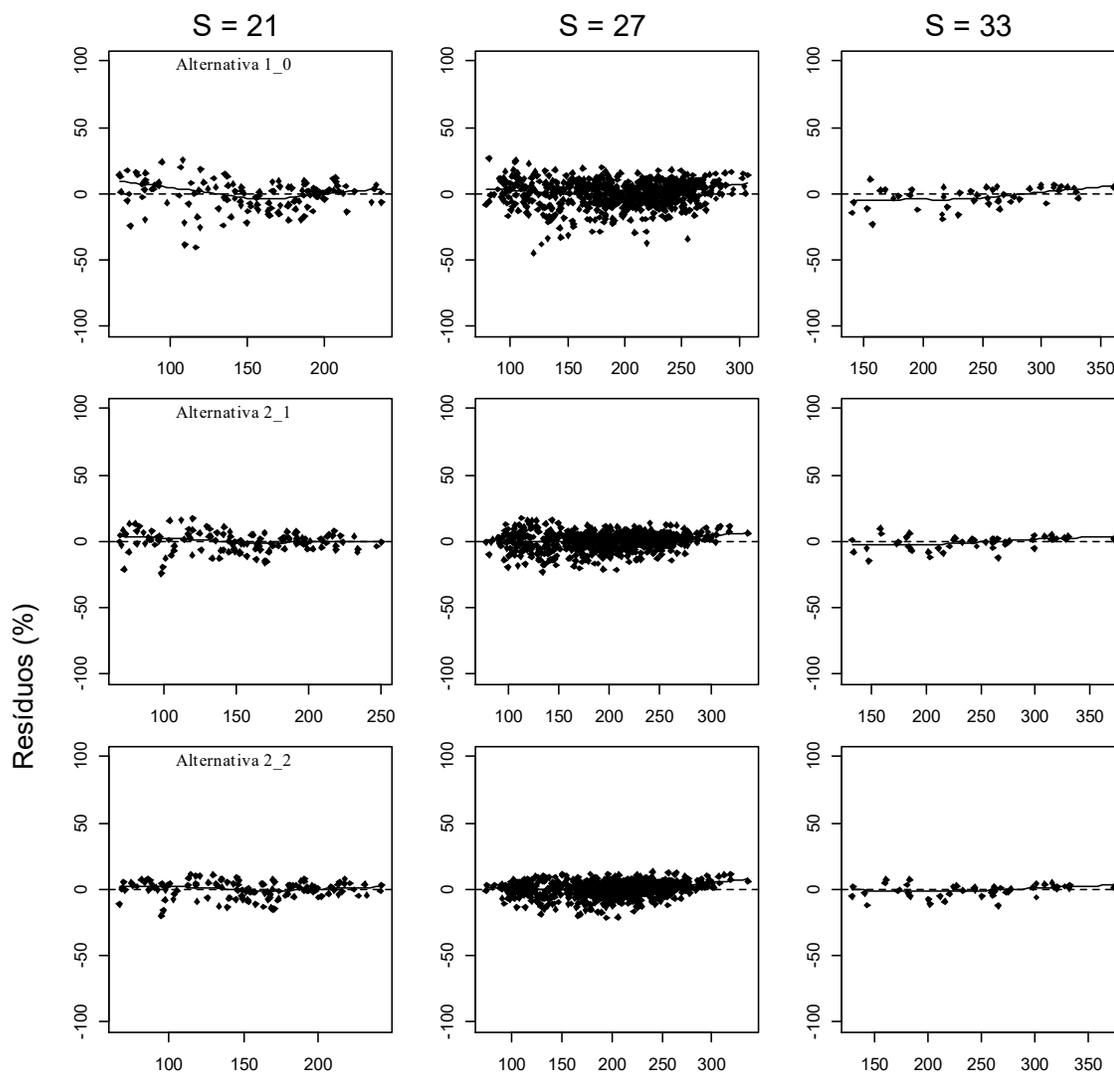
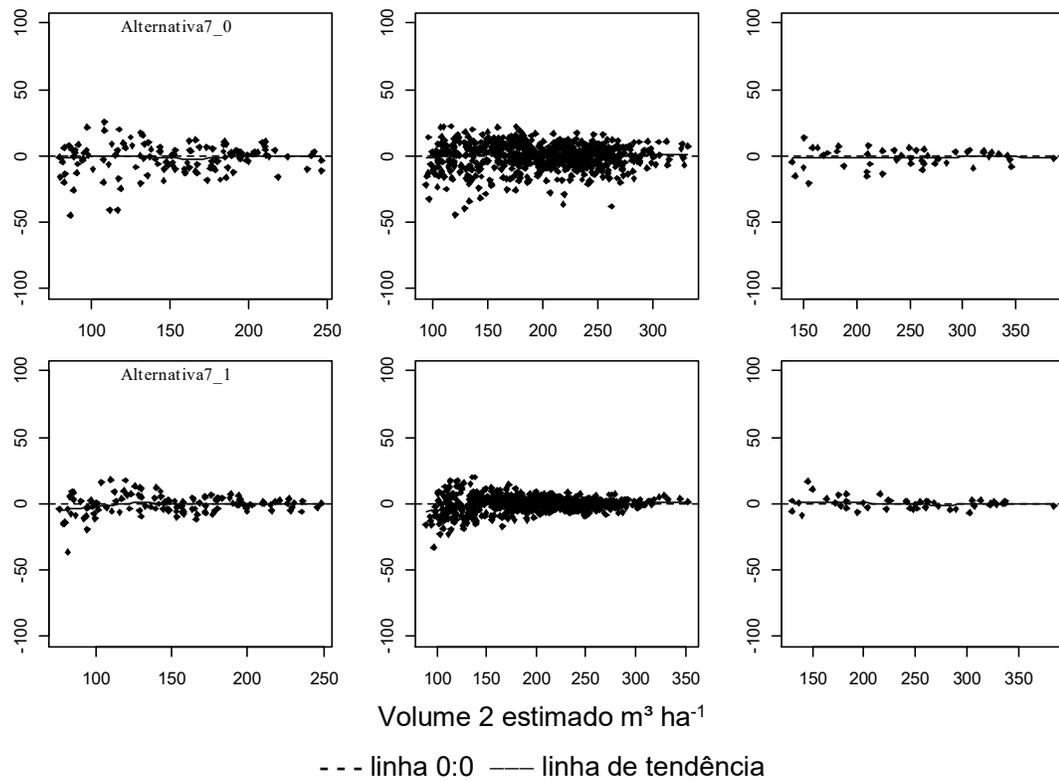


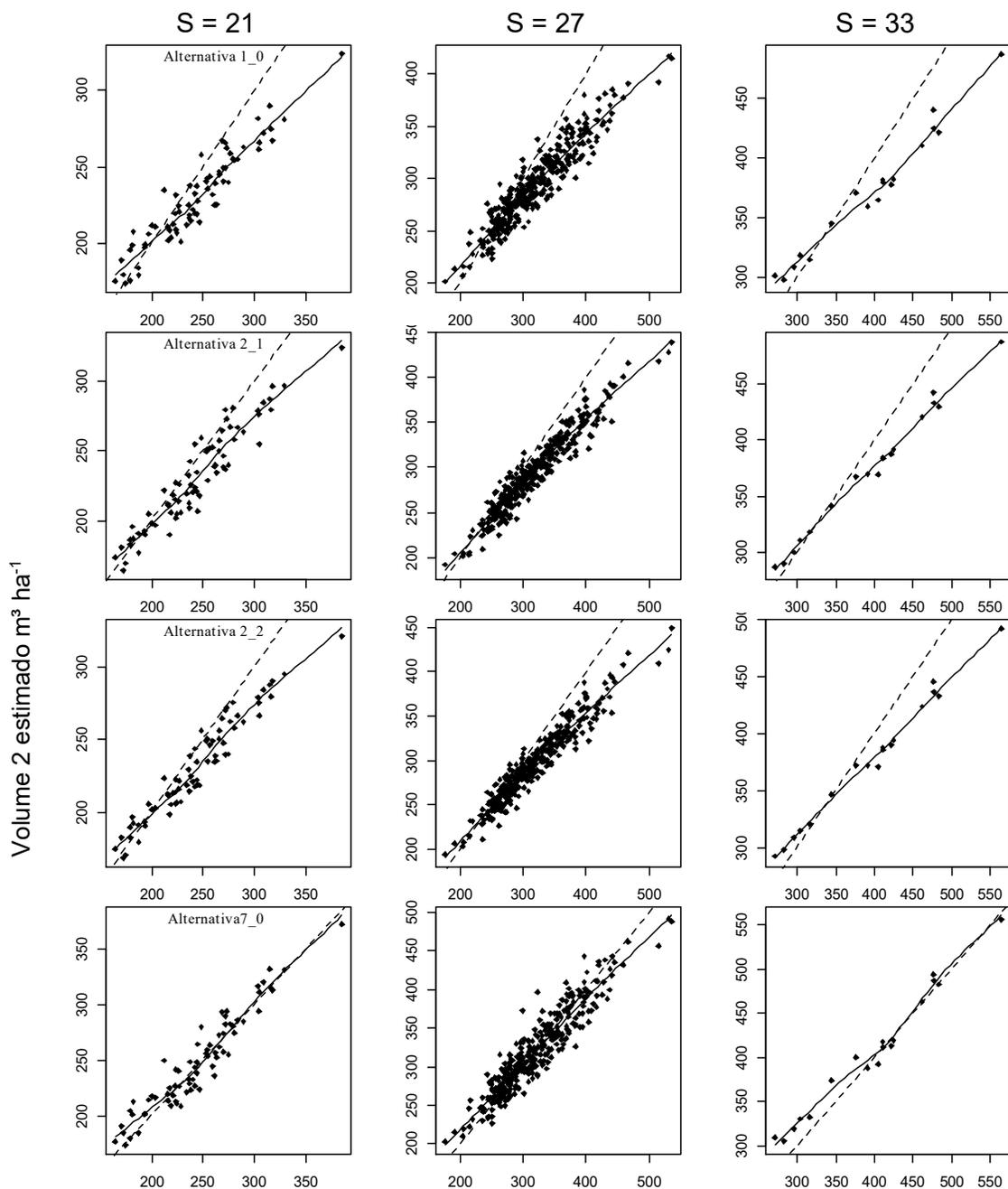
Figura 2 – Distribuição dos resíduos percentuais das equações referentes às alternativas avaliadas de acordo com o critério 1 de separação da base de dados, para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em $\text{m}^3 \text{ha}^{-1}$ e no eixo Y o erro em percentagem.





A tendência de subestimação revelada pela estatística Bias (%) (Tabela 5), pode ser confirmada nos gráficos apresentados nas Figuras 3 e 4, em todas as alternativas e em todas as classes de produtividade, notando-se um viés menos acentuado para a alternativa 7_0. Nota-se também que, para todas as alternativas e classes de produtividade, este viés é mais acentuado nos povoamentos com maiores volumes. No caso dos povoamentos com menores volumes, nota-se uma leve tendência de superestimação.

Figura 3 – Volumes totais por parcela estimados *versus* observados, obtidos pelas equações ajustadas seguindo o critério 1, isto é, aplicadas em uma base de dados independente do ajuste, com as mesmas parcelas com idade futura superior a 60 meses, em cada classe de produtividade.



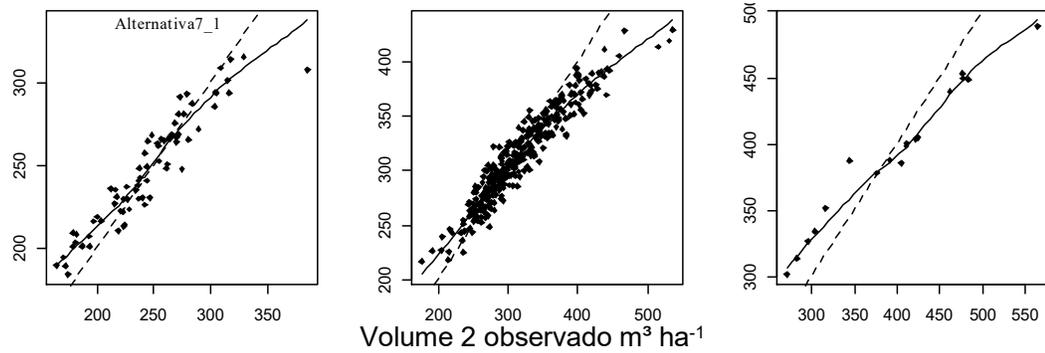
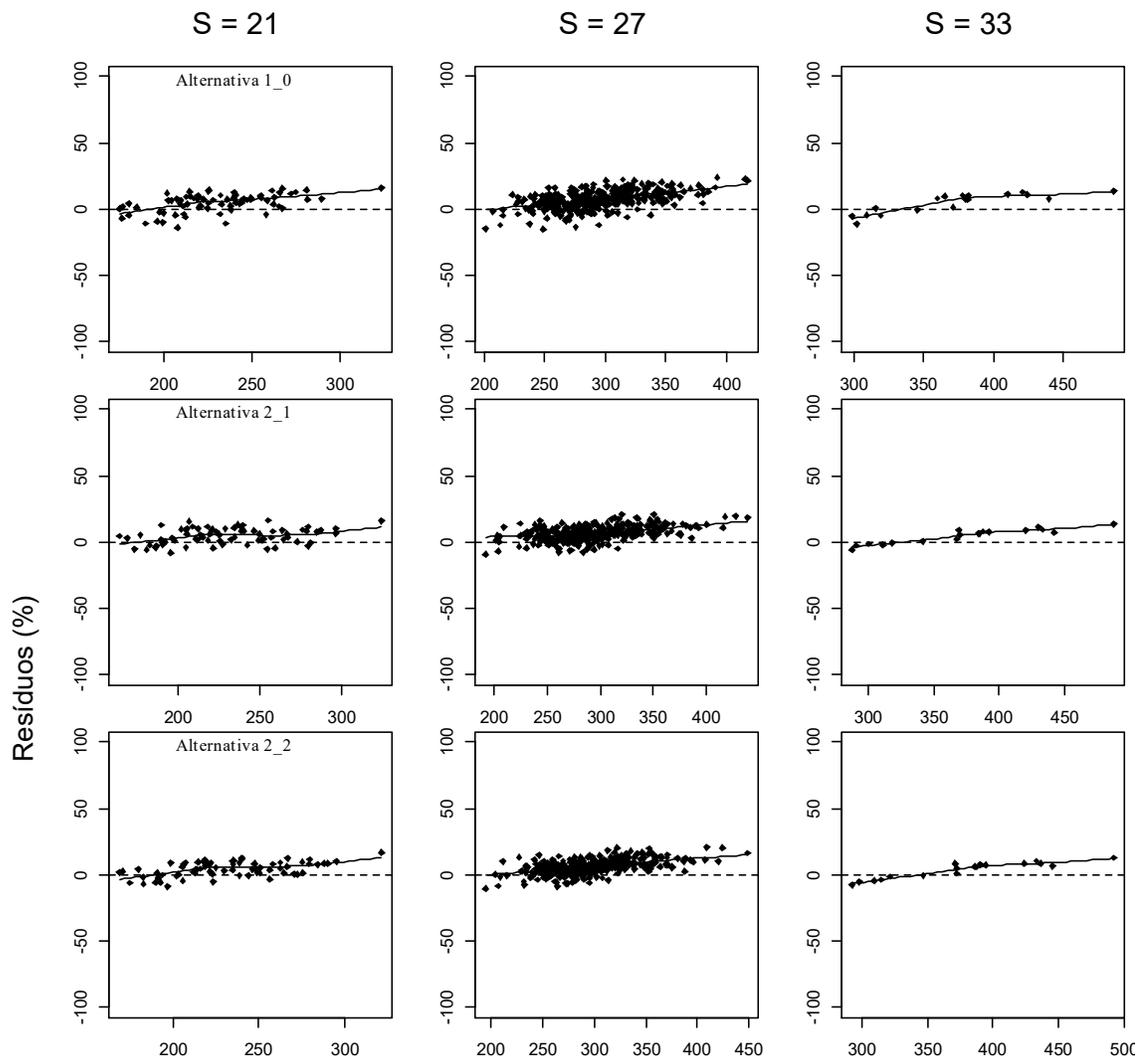
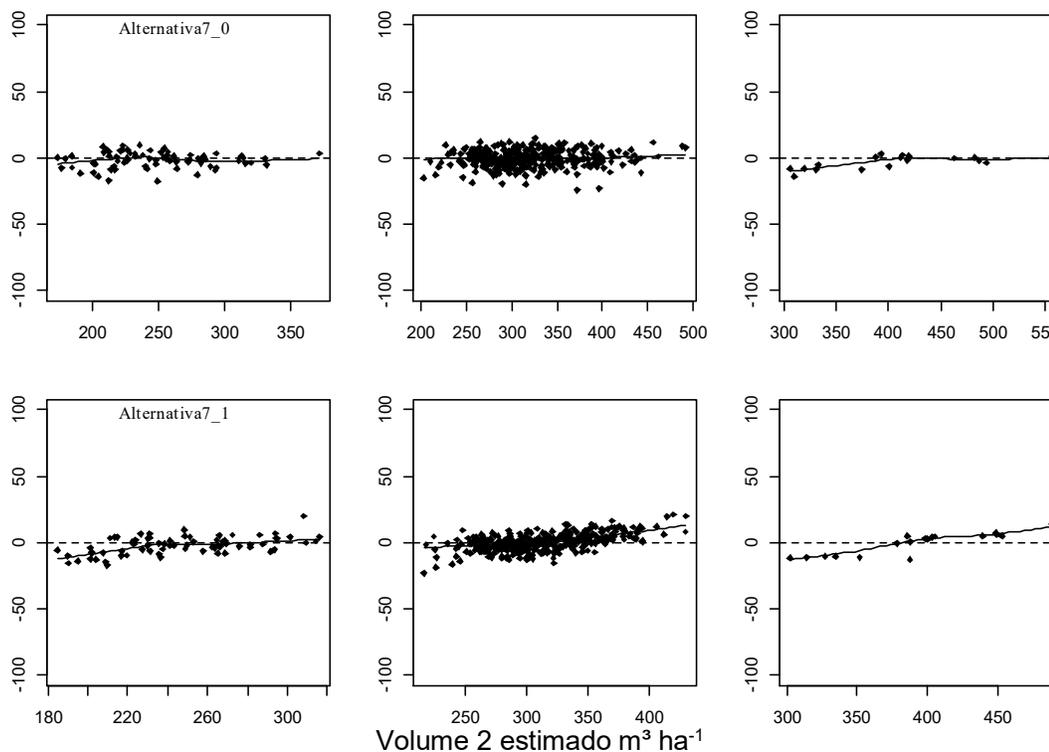


Figura 4 – Distribuição dos resíduos percentuais em função dos volumes totais estimados, obtidos pela aplicação das alternativas aplicadas seguindo o critério 1, isto é, em uma base de dados independente do ajuste, com as mesmas parcelas com idade futura superior a 60 meses em cada classe de produtividade para os modelos em nível de povoamento.





3.2. Critério 2 – Alternativas usando outras variáveis como grupos de efeitos aleatórios

De acordo com o apresentado na Tabela 6, as alternativas propostas utilizando modelos mistos apresentaram melhores ajustes em todas as classes de produtividade. Cabe destacar as alternativas 3 e 11, na classe II ($S = 27$), podendo-se observar para elas um RMSE (%) de 4,77 e 4,57, respectivamente, o que representa uma melhoria significativa para a mesma estatística obtida pelo modelo de Clutter (1963) na forma original (alternativa 1 da Tabela 6) na mesma classe de produtividade. Destaca-se ainda o $r_{y\hat{y}}$ superior, contudo, com uma tendência maior de superestimação na classe de produtividade Média para a alternativa 3.

Considerando agora a classe de produtividade Alta ($S = 33$), a alternativa 4 para projeção de volume adaptada de Clutter (1963), com Classe de B_1 como variável de agrupamento de efeitos aleatórios e I_2^{-1} como inclinação aleatória, mostrou-se superior quando comparada às mesmas classes de produtividade da alternativa de referência (alternativa 1), contudo, com uma tendência maior de superestimação na classe Alta. Na classe de produtividade Baixa ($S = 21$) a alternativa 11 apresentou RMSE (%) igual a 4,43, valor inferior ao de todas as alternativas na mesma classe.

Na classe de produtividade Baixa ($S = 21$), observa-se uma leve superioridade de $r_{y\hat{y}}$ na alternativa 8. Nas demais estatísticas, percebe-se que os valores são idênticos para o RMSE (%), com uma leve piora na classe de produtividade Baixa considerando a alternativa 9. Destaca-se o valor de BIAS (%) de -0,0371 na classe de produtividade média para a alternativa 8, porém próxima da alternativa 7 na mesma classe de produtividade.

Em todas as classes de produtividade, a alternativa 11 apresentou $r_{y\hat{y}}$ variando de 0,9905 a 0,9921, RMSE (%) entre 4,42 e 4,57 e BIAS (%) com variação de -0,2869 a 1,3930. Na alternativa 4, a variação de $r_{y\hat{y}}$ foi entre 0,9750 e 0,9930, e de 4,36 a 7,36 de RMSE (%). A variação de BIAS (%) nesta alternativa foi de -0,4493 a 1,4510.

Tabela 6 – Estatísticas utilizadas para avaliar o ajuste dos onze modelos de projeção de volume, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1*	S = 21	0,9720	8,24	0,3830
	S = 27	0,9660	8,91	-0,8730
	S = 33	0,9890	5,76	0,9711
2	S = 21	0,9719	8,37	-0,5814
	S = 27	0,9685	8,25	0,1745
	S = 33	0,9888	8,55	-0,7855
3	S = 21	0,9800	7,01	-0,1040
	S = 27	0,9920	4,77	1,9900
	S = 33	0,9740	7,46	-0,4760
4	S = 21	0,9820	6,48	0,0346
	S = 27	0,9750	7,36	-0,4493
	S = 33	0,9930	4,36	1,4510
5	S = 21	0,9800	6,97	-0,1280
	S = 27	0,9750	7,44	-0,4690
	S = 33	0,9920	4,78	2,0200
6	S = 21	0,9800	6,84	0,5170
	S = 27	0,9730	7,73	-0,6800
	S = 33	0,9910	4,91	1,7990
7	S = 21	0,9790	7,14	1,2230
	S = 27	0,9720	7,79	-0,2190
	S = 33	0,9910	4,71	0,8910
8	S = 21	0,9800	7,06	0,9985
	S = 27	0,9720	7,79	-0,0371
	S = 33	0,9910	4,65	0,9974
	S = 21	0,9790	7,15	0,5670

9	S = 27	0,9720	7,87	-0,4600
	S = 33	0,9910	4,44	0,4430
10	S = 21	0,9897	5,01	0,3689
	S = 27	0,9871	5,35	-0,4496
	S = 33	0,9908	4,90	1,8714
11	S = 21	0,9919	4,43	0,1957
	S = 27	0,9905	4,56	-0,2869
	S = 33	0,9921	4,42	1,3930

Comparando-se as alternativas 11 e 7, verifica-se superioridade da alternativa 11 em todas as classes de produtividade, exceto para o BIAS (%) nas classes de produtividade Média e Alta, com tendência de subestimação na classe Média de -0,2869 comparada a -0,2190 na mesma classe na alternativa 7, e menor superestimação na classe Alta, com valor de 0,891 na alternativa 7, e 1,3930 na alternativa 11.

Os coeficientes fixos dos modelos representados nas onze alternativas foram todos significativos ($p < 0,01$) (Tabela 7).

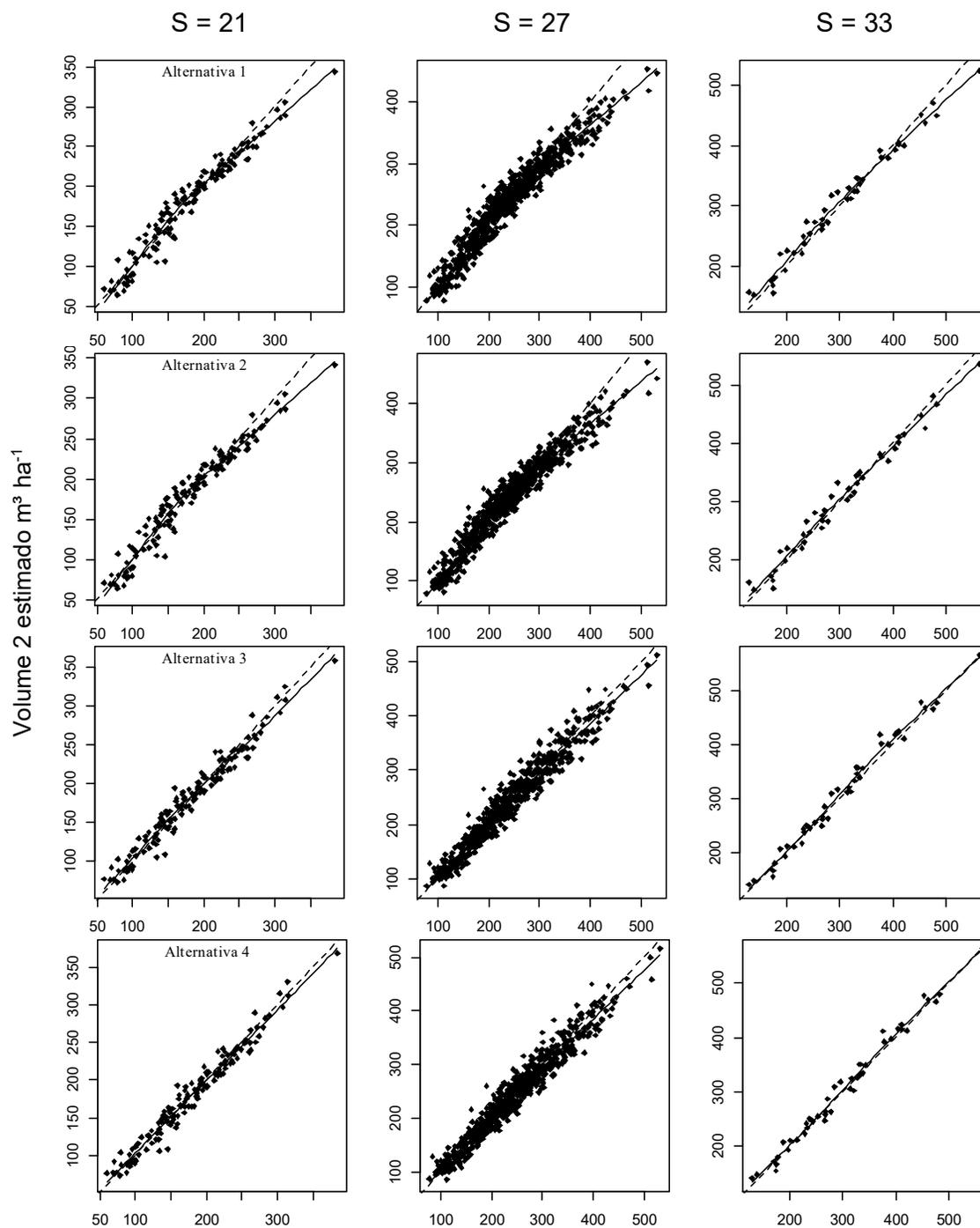
Tabela 7 – Coeficientes do ajuste das alternativas propostas.

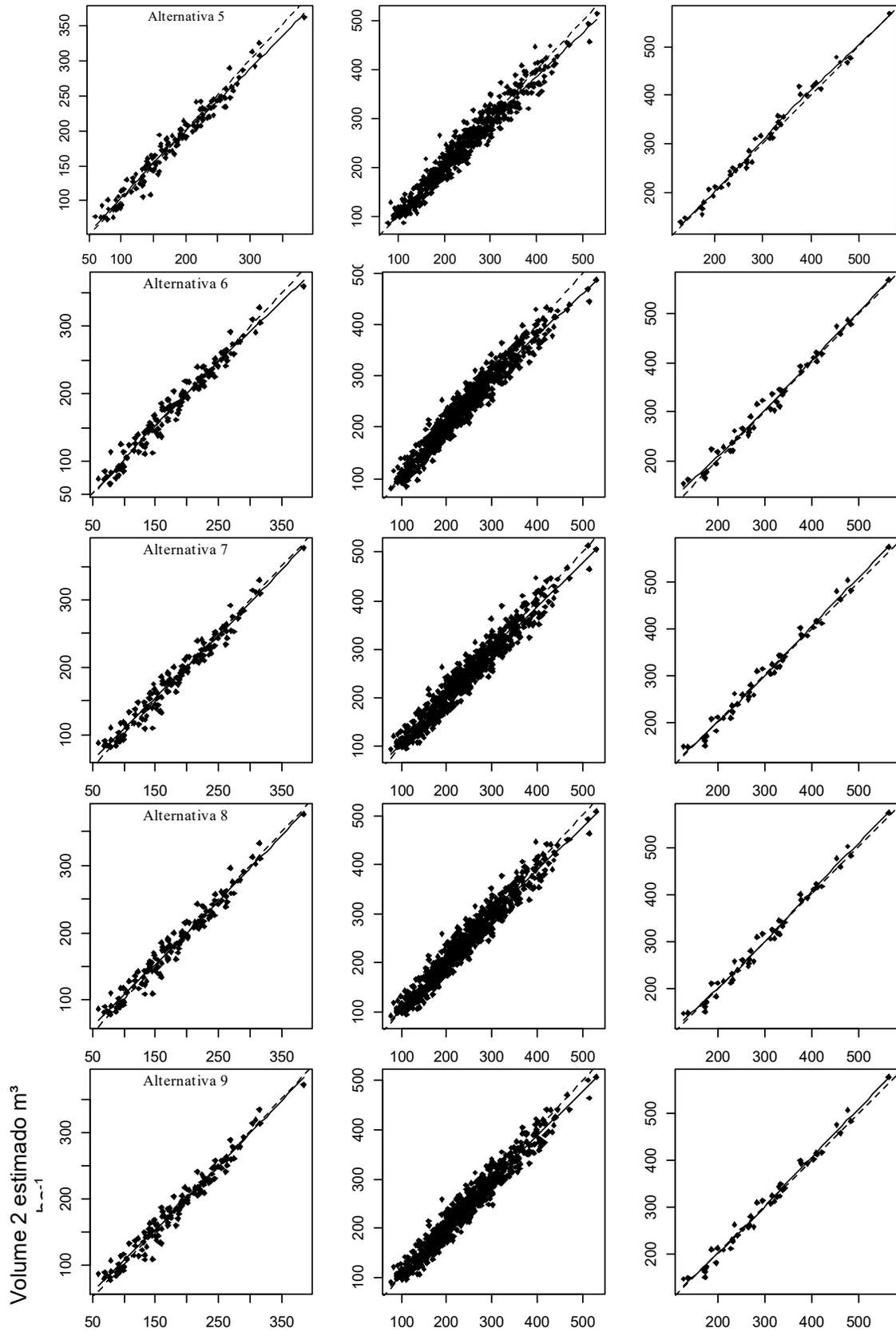
Alternativa	Coeficientes				
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
1	1,3727**	-13,0218**	0,0228**	1,211**	-
2	1.4720**	-13.9058**	0.023242**	1.1846**	-
3	3,0744**	-5,9277**	0,0216**	0,6228**	-
4	2,9530**	-9,2778**	0,0216**	0,6704**	-
5	2,8759**	-6,1663**	0,0216**	0,6805**	-
6	1,1926**	-11,9008**	0,0227**	1,2699**	-
7	417,6031**	18,8122**	425,9958**	-10,6078**	130,5650**
8	413,6702**	18,3562**	635,0153**	-17,2602**	195,9035**
9	403,0818**	17,7815**	709,4137**	-19,7930**	222,7860**
10	1,6660**	23,0791**	0,0218**	1,0334**	-
11	2,9354**	13,1711**	0,0070**	0,7499**	-

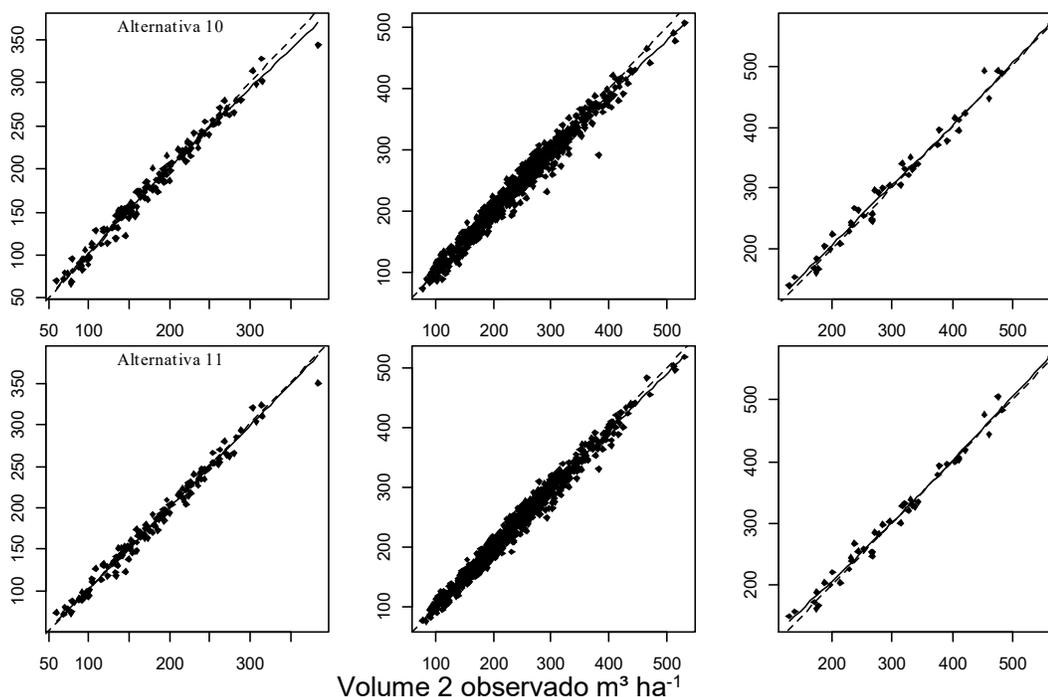
** - significativo a 1% pelo Test T para modelos fixos, e Teste de Wald para modelos mistos

Pode-se observar na Figura 5, que os valores estimados pelas equações das onze alternativas ajustadas para a projeção de volume estão concentradas próximas à linha 0:1, o que reflete o bom desempenho dos modelos nas três classes de produtividade. Destacam-se as alternativas 10 e 11, que apresentam maior acurácia.

Figura 5 – Volumes estimados *versus* observados pelas equações ajustadas referentes aos onze modelos, em cada classe de produtividade.



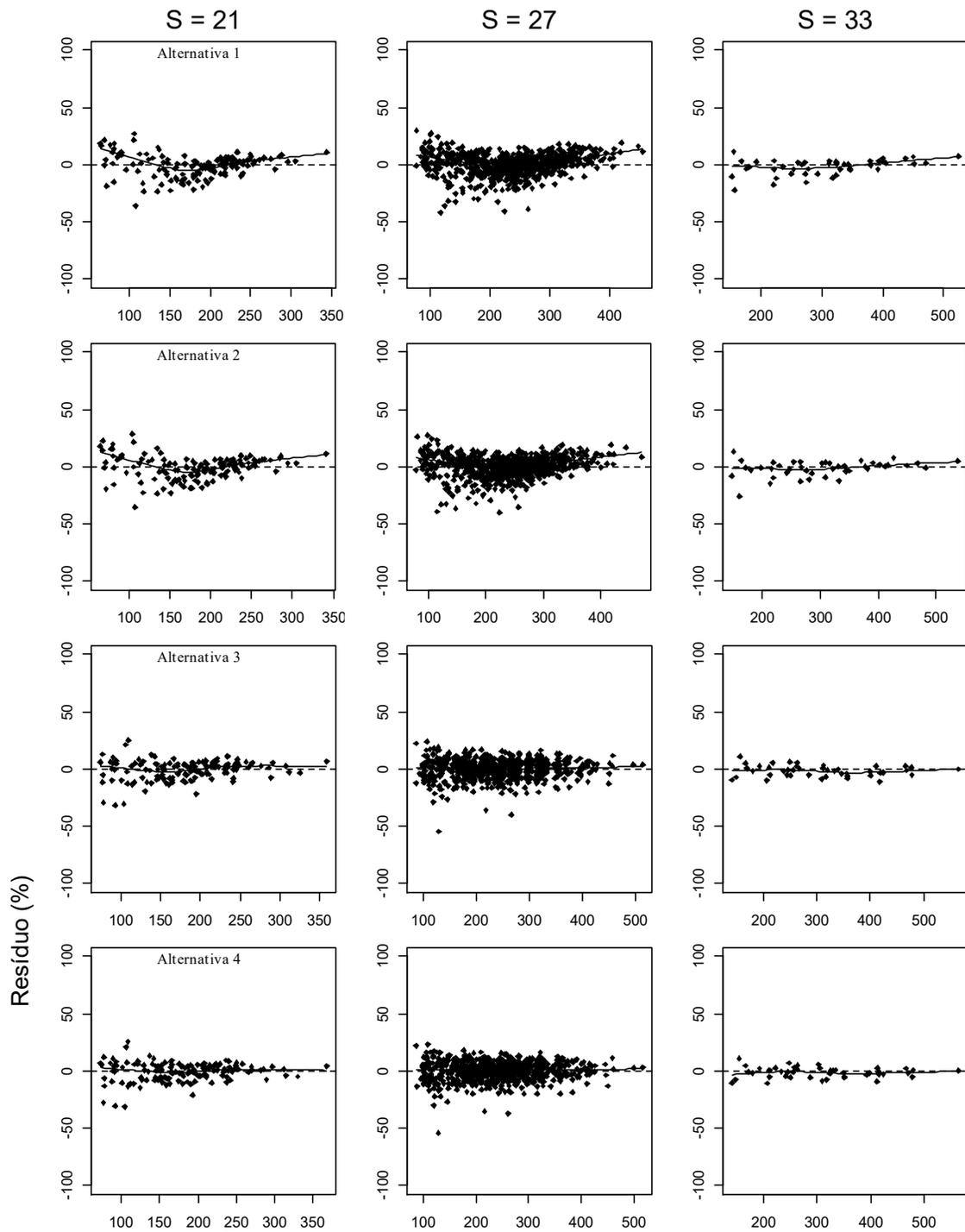


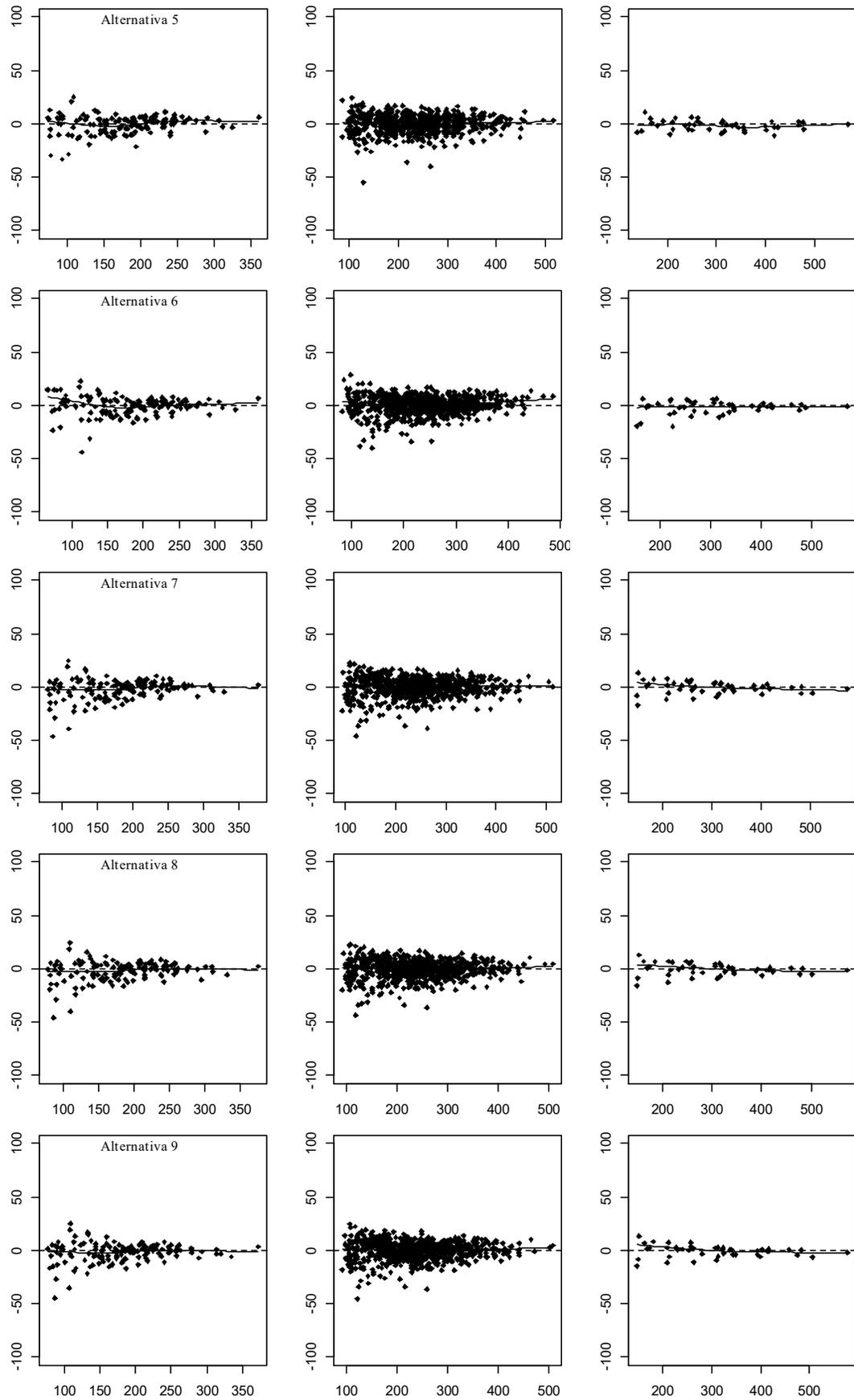


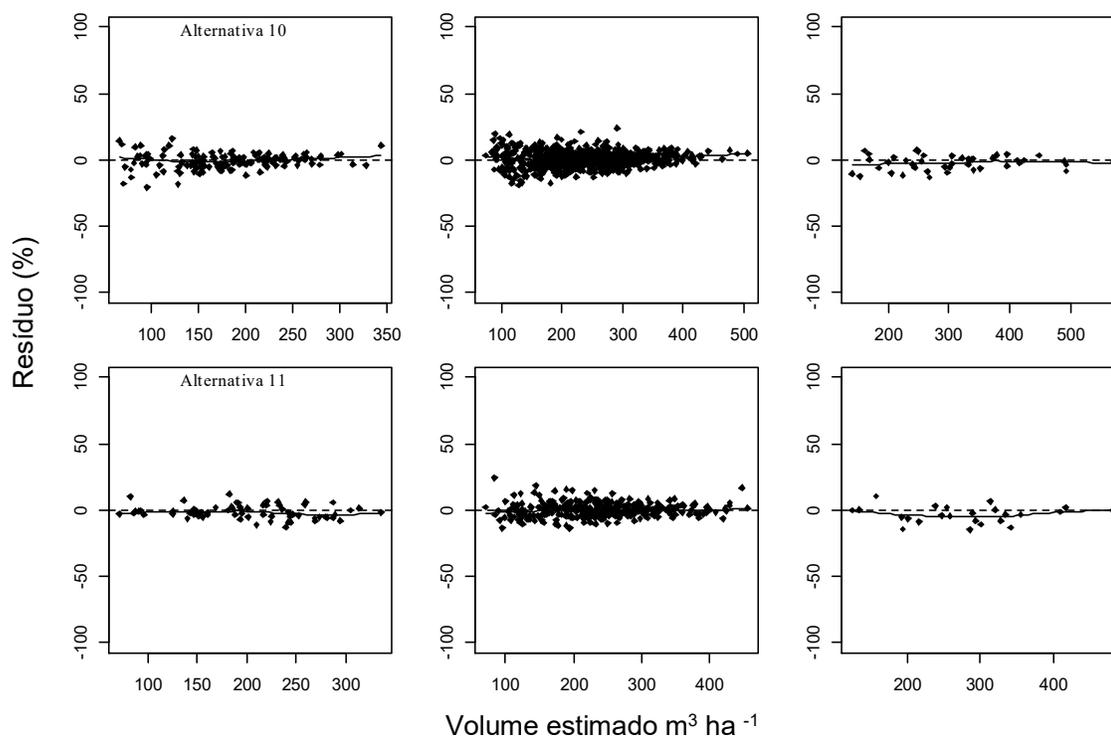
Os resíduos percentuais referentes à projeção de volume (Figura 6) demonstram tendência similar, exceto nas alternativas 10 e 11, nas quais os resíduos estão agrupados em torno de 20%. Na alternativa 1, observa-se uma tendência de supestimação para povoamentos de menor e maior volume, além de pontos de subestimação que chegam a 50% nas classes Baixa e Média, tendência contrária à apresentada por Mendonça (2010). Uma tendência similar é apresentada na alternativa 6. Observa-se que para as alternativas 3, 4, 5 e 7 a maior parte dos resíduos ficou bem distribuída, demonstrando um ajuste acurado destas alternativas.

Diante do exposto, considerando as estatísticas e a análise gráfica dos resíduos, pode-se afirmar que as alternativas 7, 10 e 11 conseguiram estimar com mais acurácia o volume nas três classes de produtividade, com destaque para o modelo 11, que apresentou estatísticas similares no ajuste e na validação com uma base de dados independente, e apresentou resíduos dentro do limite $\pm 20\%$.

Figura 6 – Distribuição dos resíduos percentuais onze alternativas de projeção de volume em função dos volumes estimados, em cada classe de produtividade.







Com o objetivo de avaliar a capacidade de generalização das alternativas propostas, foi conduzida uma validação, aplicando-se os modelos ajustados em um conjunto de parcelas independentes daquelas utilizadas no ajuste. O resultado é apresentado na Tabela 8.

Tabela 8 – Estatísticas utilizadas para avaliar os modelos ajustados em uma base de dados de validação, com parcelas independentes daquelas usadas no ajuste.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1*	S = 21	0,9600	8,87	-0,9290
	S = 27	0,9760	8,12	-0,9640
	S = 33	0,9800	8,37	4,9000
2	S = 21	0,9573	9,16	-1,1465
	S = 27	0,9743	8,34	-1,0642
	S = 33	0,9824	7,26	3,1489
3	S = 21	0,9670	8,06	-0,8710
	S = 27	0,9810	6,78	-0,5560
	S = 33	0,9890	6,70	4,6140
4	S = 21	0,9670	7,94	-0,2730
	S = 27	0,9810	6,65	-0,4810
	S = 33	0,9880	6,75	4,4890
5	S = 21	0,9670	8,06	-0,9150
	S = 27	0,9810	6,77	-0,5550
	S = 33	0,9890	6,70	4,6260
	S = 21	0,9630	8,48	-0,7550

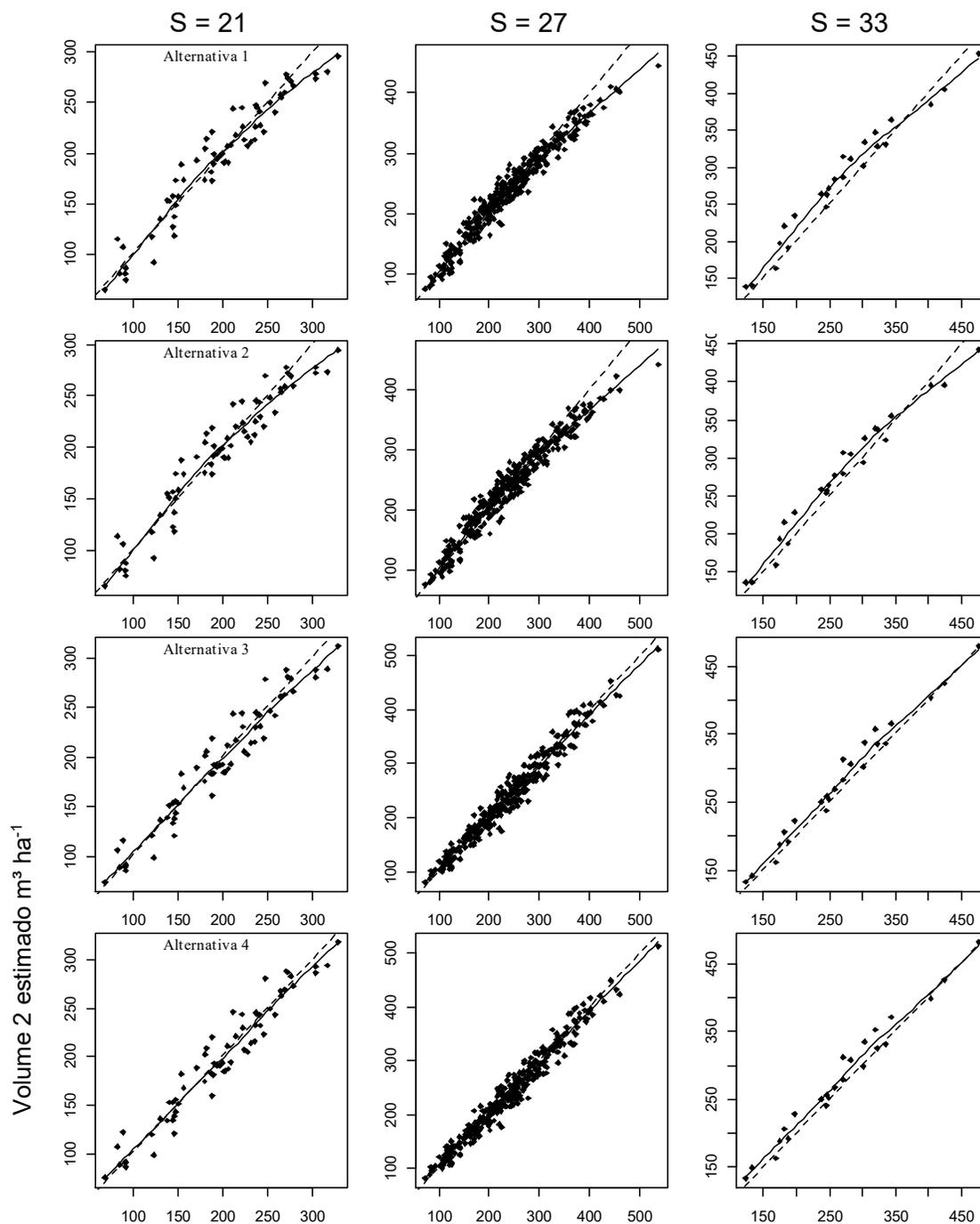
6	S = 27	0,9780	7,32	-0,8500
	S = 33	0,9850	8,00	5,5310
7	S = 21	0,9660	8,08	0,0642
	S = 27	0,9800	6,84	-0,2613
8	S = 33	0,9870	6,21	3,3755
	S = 21	0,9730	7,97	-3,3290
9	S = 27	0,9810	6,76	-0,2810
	S = 33	0,9840	7,63	4,0790
10	S = 21	0,9750	7,52	-2,8900
	S = 27	0,9810	6,80	-0,2500
11	S = 33	0,9830	7,62	3,5600
	S = 21	0,9852	5,39	0,6208
10	S = 27	0,9859	5,82	-0,6046
	S = 33	0,9883	6,74	4,5398
11	S = 21	0,9858	5,63	1,6071
	S = 27	0,9880	5,33	-0,4757
	S = 33	0,9849	6,66	3,4785

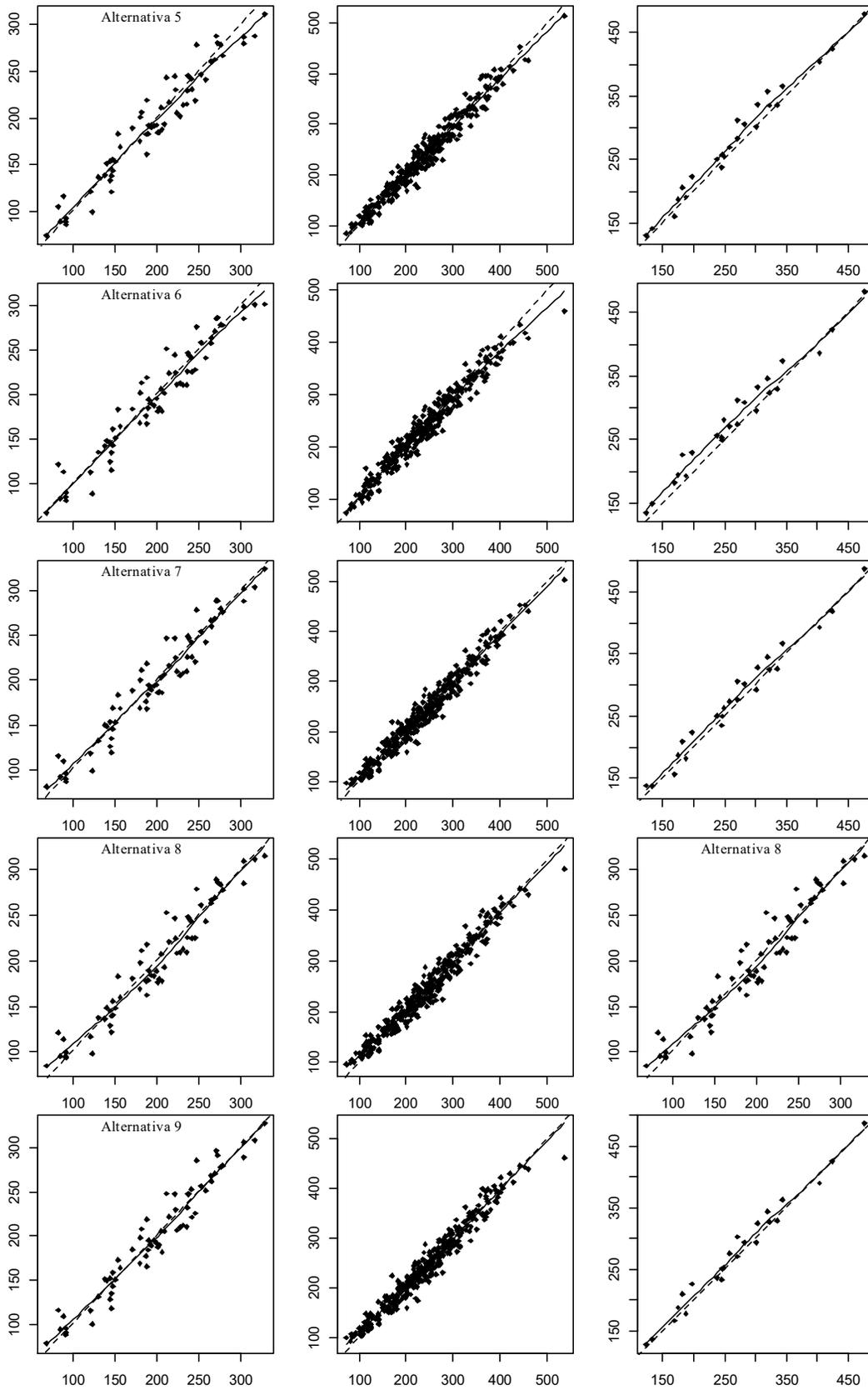
Observa-se que, em geral, as estatísticas obtidas com a aplicação dos modelos ajustados apresentaram tendências próximas quando comparadas às estatísticas obtidas no ajuste. Considerando o RMSE (%), cabe destacar, para a classe de produtividade Baixa, a alternativa 10 com 5,3905, e para a classe de produtividade Média a alternativa 11 com RMSE (%) de 5,3280.

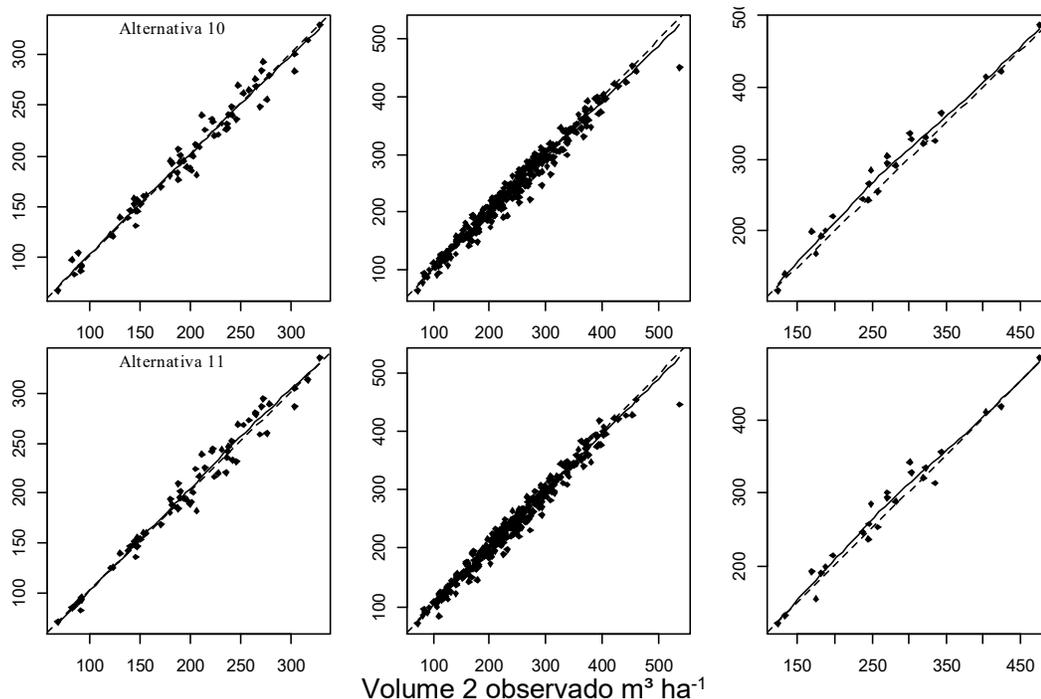
Analisando o $r_{\hat{y}y}$, observa-se que a na classe de produtividade Média, a alternativa 11 apresenta 0,9880, superior às demais alternativas na mesma classe. Na classe Alta, observa-se na alternativa 3 uma correlação de 0,9890, e na classe Baixa 0,9858 de $r_{\hat{y}y}$ para a alternativa 11, superior aos valores observados nas mesmas classes nas demais alternativas.

Na Figura 7 observa-se uma tendência de subestimação nas parcelas com volume acima de $400 \text{ m}^3 \text{ ha}^{-1}$ em todas as alternativas, sendo mais acentuada nas alternativas 1, 2 e 6. Verifica-se na classe de produtividade Alta (S = 33) da alternativa 1, uma tendência de superestimar parcelas com menor volume. É possível verificar uma menor dispersão entre volumes observados e estimados nas alternativas mistas 10 e 11, com apenas uma parcela apresentando tendência de subestimação nestes casos.

Figura 7 – Volumes totais por parcela estimados *versus* observados, obtidos pelas equações ajustadas seguindo o critério 2, isto é, aplicadas em uma base de dados independente do ajuste, com as mesmas parcelas com idade futura superior a 60 meses, em cada classe de produtividade.

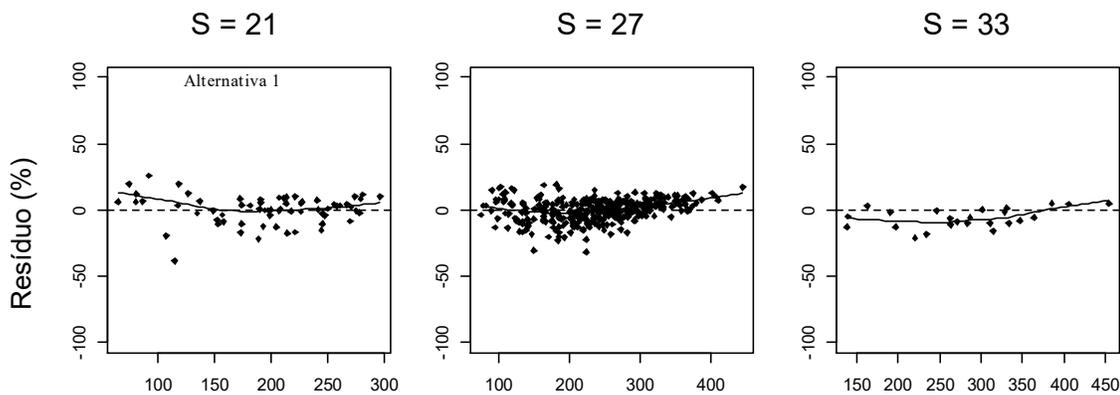


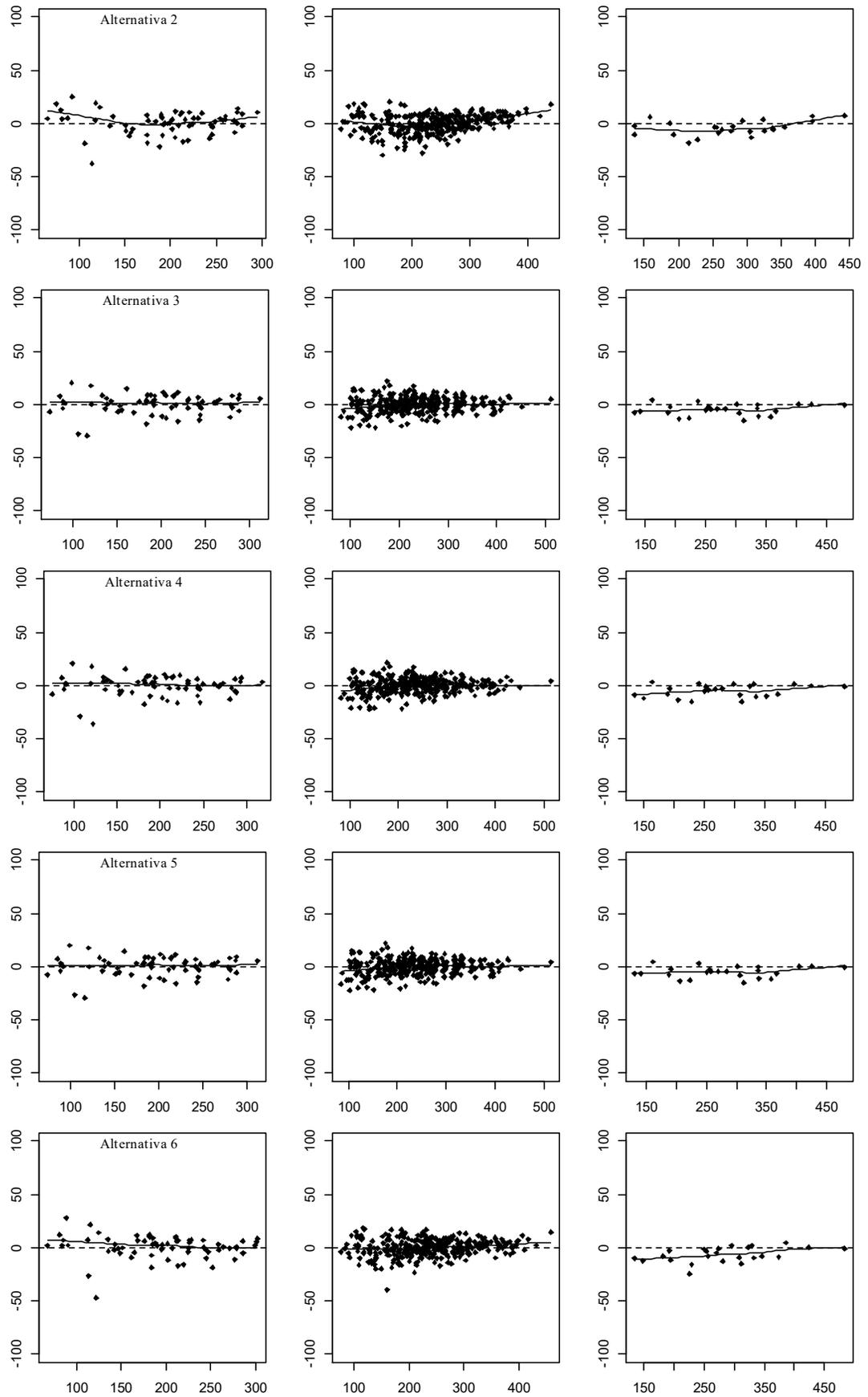


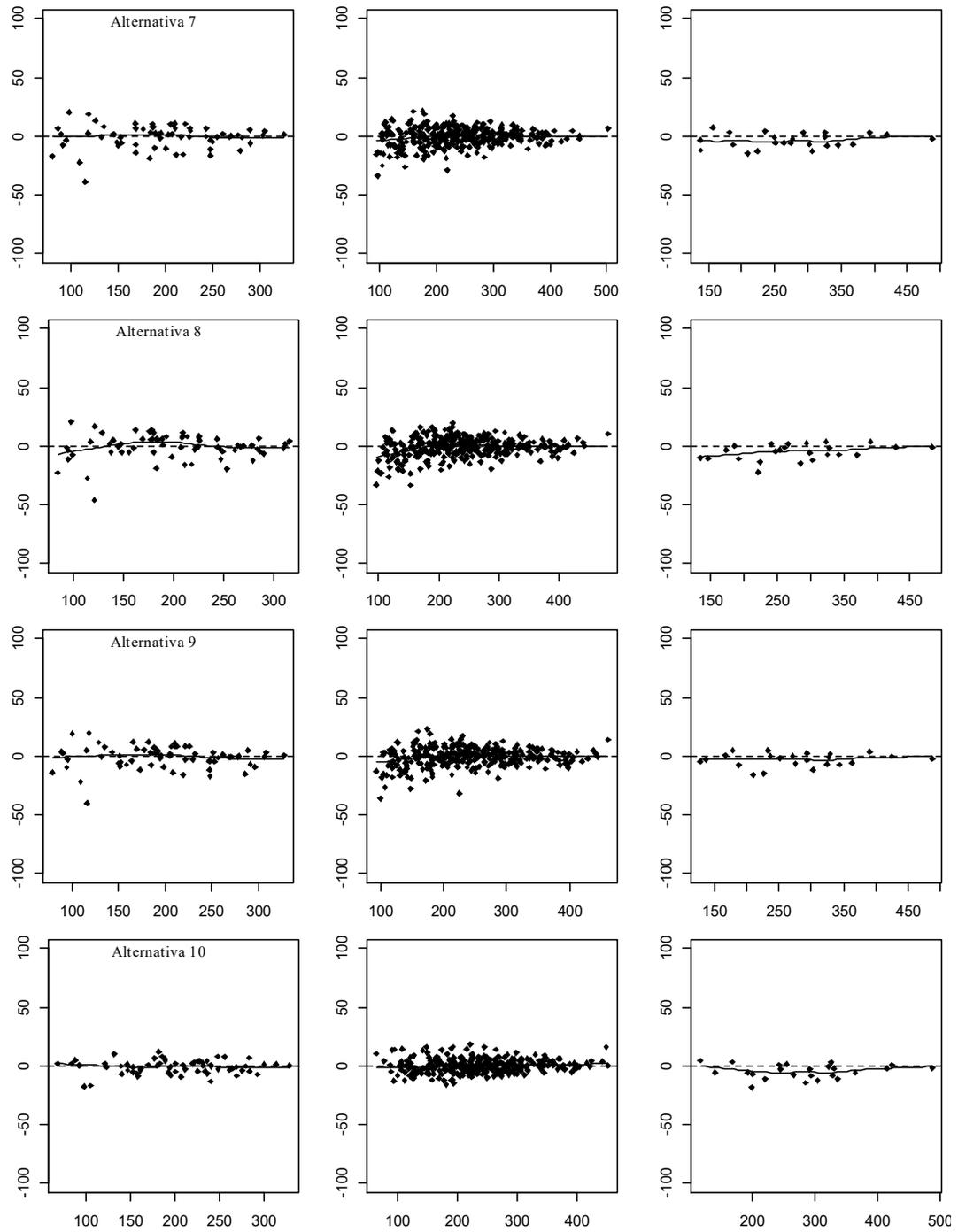


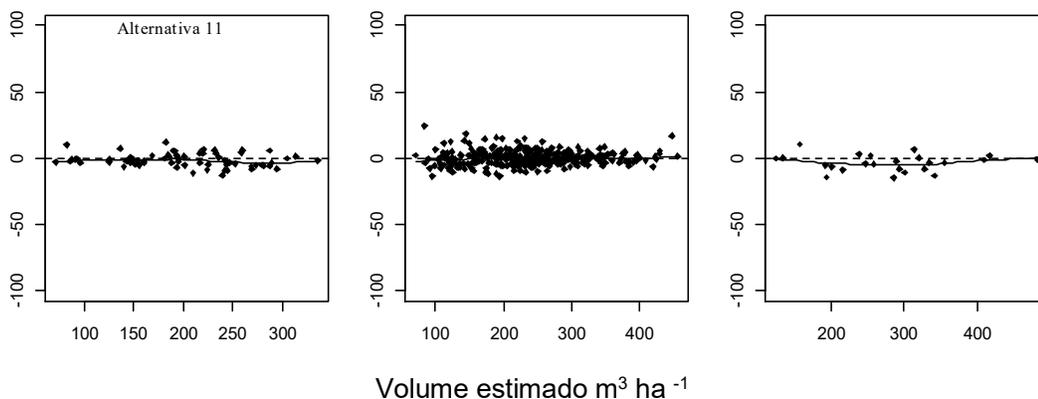
A dispersão residual relativa (Figura 8) é mais acentuada nas alternativas 1, 2 e 6, corroborando com as tendências apresentadas. Observa-se ainda que nas classes de produtividade Alta destas alternativas, uma leve tendência de superestimar parcelas com volumes abaixo de $350 \text{ m}^3 \text{ ha}^{-1}$. Também por esta figura verifica-se ausência de viés pela concentração da nuvem de pontos entre $\pm 15\%$.

Figura 8 – Distribuição dos resíduos percentuais em função dos volumes totais estimados, obtidos pela aplicação das alternativas aplicadas seguindo o critério 2, isto é, em uma base de dados independente do ajuste, com as mesmas parcelas com idade futura superior a 60 meses em cada classe de produtividade para os modelos em nível de povoamento.









3.3. Comparação das alternativas ajustadas e validadas para os critérios definidos

Comparando-se as alternativas de modelagem de efeitos fixos que usaram dados do critério 1, observa-se que, embora semelhantes nas estatísticas de ajuste, a alternativa 1_0 apresenta um viés mais acentuado na classe de produtividade Média ($S = 27$), com tendência de subestimação nas parcelas com maiores volumes. Além disso, verifica-se uma tendência de superestimação na classe de produtividade Alta ($S = 33$). A alternativa 7_0, baseada na modelagem da estratégia 2, apresenta-se como a mais acurada e menos viesada no ajuste das alternativas com modelagem de efeitos fixos.

Na validação, a alternativa 1_0 apresenta tendência acentuada de subestimação de parcelas com maior volume em todas as classes de produtividade, e a acurácia diminui. Embora apresente tendência semelhante na classe de produtividade Média ($S = 27$), o viés é menos acentuado na alternativa 7_0 comparado à alternativa 1_0.

Ao realizar a mesma comparação entre modelos de efeitos fixos (alternativa 1 e 7) para dados do critério 2, pelas estatísticas de ajuste e análise gráfica dos resíduos, conclui-se que a alternativa 7 se apresenta mais acurada e menos viesada no ajuste, exceto na classe Baixa ($S = 21$), com uma leve tendência de subestimar parcelas com maiores volumes.

Comparando o desempenho dos modelos ajustados e validados com base apenas em efeitos fixos, de acordo com os critérios 1 e 2, verifica-se que as alternativas são acuradas, com a alternativa 7 é ligeiramente mais acurada quando comparada às alternativas 1_0, 7_0 e 1, com RMSE relativo no ajuste de 7,79 na classe de produtividade Média ($S = 27$), apresentando uma melhoria de 10% em relação a alternativa 7_0, que apresentou RMSE relativo de 8,65, e com viés

semelhante ao apresentado na alternativa 7_0 na mesma classe. Na validação, a alternativa 7_0 é mais acurada em todas as classes de produtividade, embora o ganho em acurácia na classe de produtividade Média (S = 27) tenha sido de 6%, na classe de produtividade Baixa (S = 21) representou 30%, e na Alta (S = 33) de 26%. O Bias (%) é menor na classe de produtividade Média (S = 27) na alternativa 7, e com tendência de subestimação, comparado com a mesma da alternativa 7_0, porém com tendência levemente superior para superestimação.

Dentre as alternativas que usaram modelagem mista, e foram ajustadas e validadas com dados do critério 1, a alternativa 7_1, baseada na estratégia 2, apresenta-se como a mais acurada e menos tendenciosa nas classes de produtividade Média (S = 27) e Alta (S = 33). Na classe de produtividade Baixa (S = 21), a alternativa 2_2 foi a mais acurada, semelhante à apresentada pela alternativa 7_1. Em termos de tendência, a alternativa 2_2 apresentou um menor viés que a alternativa 7_1. Pela análise gráfica dos resíduos, ao avaliar os volumes estimados *versus* observados, observa-se que nas classes de produtividade Média (S = 27) e Alta (S = 33), a tendência de subestimação é maior na alternativa 2_2 quando comparada à alternativa 7_1, fato corroborado pelo gráfico de resíduos relativos, nos quais para todas as classes de produtividade, a dispersão da nuvem de pontos é menor na alternativa 7_1.

Na comparação das alternativas mistas ajustadas e validadas com dados de ambos os critérios, verifica-se que no ajuste, a alternativa 11 (critério 2) foi a mais acurada em todas as classes de produtividade, embora na classe de produtividade média tenha sido semelhante à alternativa 7_1, que apresenta viés reduzido (Bias relativo de 0,0092). Na classe de produtividade Alta (S = 33), a alternativa 7_1 mantém a superioridade em termos de ausência de tendência, e mais acurada, com melhoria de 15% quando comparada à alternativa 11. Cabe destacar que a correlação foi superior em todas as classes de produtividade na alternativa 11.

Ao realizar a análise gráfica dos resíduos, verifica-se que para o ajuste, a alternativa 11 é mais acurada, uma vez que há a nuvem de pontos é menos dispersa, especialmente nas classes de produtividade Média (S = 27) e Baixa (S = 21), indicando a alternativa 11, dentre as ajustadas com modelagem mista, como a mais acurada e menos viesada.

Na validação das alternativas ajustadas por modelagem mista, a acurácia superior da alternativa 11 é confirmada. Verifica-se, entretanto, que tal alternativa é mais viesada nas classes de produtividade Baixa (S = 21) e Alta (S = 33), quando

comparada à alternativa 7_1 e 7, que apresentam BIAS relativo na classe Baixa de 0,06 e 1,3544, e na Alta -0,74 e 3,3755, respectivamente.

A alternativa 11, mostrou ser a mais acurada e menos viesada dentre todas as alternativas apresentadas para ambos os critérios, tanto no ajuste, quanto na validação. No ajuste geral, o RMSE relativo obtido foi de 4,79%, com BIAS (%) de -0.15. Demolinari (2006) obteve S_{yx} (%) (em todas as classes de produtividade) de 4,86 no ajuste do modelo de Clutter (1963) para uma base de dados de 66 parcelas permanentes, menor comparado ao número de parcelas utilizadas neste capítulo (376 parcelas). Tal estatística é obtida de maneira similar ao RMSE (%), portanto, são comparáveis.

Para uma base de dados de eucalipto clonais, Mendonça (2010) encontrou para seu modelo um valor de S_{yx} (%) de 8,31 e para o ajuste de Clutter (1963) 7,13% para a mesma estatística. Também para uma base clonal de eucalipto, Oliveira (2012) obteve o valor de BIAS (%) de 0,57.

4. CONCLUSÕES

Comparando os modelos fixos, o modelo mais acurado foi o de Mendonça (2010) (alternativa 7), independente do critério de divisão de dados.

Utilizando modelagem mista, o modelo que apresentou melhor resultado foi o modelo de Clutter (1963) modificado, com classe de idade atual como grupo de efeitos aleatórios, com altura média da parcela na idade futura e logaritmo do volume na idade atual como inclinação aleatória (alternativa 11).

Verifica-se também que a calibração de modelos mistos utilizando outras variáveis de agrupamento diferentes da parcela, no caso de modelagem em nível de povoamento, pode ser uma opção na aplicação de equações a um conjunto de parcelas diferente das utilizadas no ajuste.

Com os resultados obtidos para as condições em que foi desenvolvido este estudo, é possível afirmar que a modelagem mista para projeção de volume em nível de povoamento foi superior à modelagem de efeitos fixos, comumente empregada neste nível de modelagem.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ARIAS-RODIL, M.; CASTEDO-DORADO, F.; CÁMARA-OBREGÓN, A.; DIÉGUEZ-ARANDA, U. Fitting and calibrating a multilevel mixed-effects stem taper model for *Maritime* Pine in NW Spain. **PloS one**, v. 10, n. 12, p. e0143521, 2015. Public Library of Science. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143521>>. Acesso em: 31/12/2015.

BINOTI, M. L. M. DA S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, v. 37, n. 4, p. 639–645, 2013. Sociedade de Investigações Florestais. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000400007&lng=pt&nrm=iso&tlng=en>. Acesso em: 31/12/2015.

BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **CERNE**, v. 21, n. 1, p. 97–105, 2015. UFLA - Universidade Federal de Lavras. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-77602015000100097&lng=en&nrm=iso&tlng=pt>. Acesso em: 31/12/2015.

CALAMA, R.; MONTERO, G. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. **Silva Fennica**, v. 39, n. 1, p. 37–54, 2005.

CALEGARIO, N.; DANIELS, R. F.; MAESTRI, R.; NEIVA, R. Desenvolvimento de um modelo linear de efeito misto na estimativa do crescimento e produção de povoamentos clonais. **Cerne, Lavras**, v. 10, n. 1, p. 67–86, 2004.

CALEGARIO, N.; MAESTRI, R.; LEAL, C. L.; DANIELS, R. F. Estimativa do crescimento de povoamentos de *Eucalyptus* baseada na teoria dos modelos não lineares em multinível de efeito misto. **Ciência Florestal**, v. 15, n. 3, p. 285–292, 2005.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração Florestal: Perguntas e Respostas**. 4th ed. Viçosa, MG, 2013.

CAO, Q. V. Linking individual-tree and whole-stand models for forest growth and yield prediction. **Forest Ecosystems**, v. 1, n. 1, p. 18, 2014. Springer. Disponível em: <<http://www.forestecosyst.com/content/1/1/18>>. Acesso em: 31/12/2015.

CARVALHO, S. DE P. C. E.; CALEGARIO, N.; SILVA, F. F. E; et al. Modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* clonal. **CERNE**, v. 17, n. 4, p. 541–548, 2011. Universidade Federal de Lavras.

CASTRO, R. V. O.; SOARES, C. P. B.; MARTINS, F. B.; LEITE, H. G. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos. **Pesquisa Agropecuária Brasileira**, v. 48, n. 3, p. 287–295, 2013. SciELO Brasil.

CLUTTER, J. L. Compatible growth and yield models for loblolly pine. **Forest**

Science, v. 9, n. 3, p. 354–371, 1963. Society of American Foresters.

DEMOLINARI, R. DE A. **Crescimento de povoamentos de eucalipto não-desbastados**, 2006. 72f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa. Viçosa, 2006.

FRAGA FILHO, C. V.; SIMIQUELI, A. P. .; ALTOE, W. . Fgmutils: Forest Growth Model Utilities. Disponível em: <<https://cran.r-project.org/web/packages/Fgmutils/index.html>>. Acesso em: 31/12/2015.

GARCÍA, O.; BURKHART, H. E.; AMATEIS, R. L. A biologically-consistent stand growth model for loblolly pine in the Piedmont physiographic region, USA. **Forest Ecology and Management**, v. 262, n. 11, p. 2035–2041, 2011. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112711005494>>. Acesso em: 31/12/2015.

GÖRGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. **Computers and Electronics in Agriculture**, v. 116, p. 221–227, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169915001994>>. Acesso em: 9/12/2015.

HALL, D. B.; BAILEY, R. L. Modeling and Prediction of Forest Growth Variables Based on Multilevel Nonlinear Mixed Models. **Forest Science**, v. 47, n. 3, p. 311–321, 2001. Society of American Foresters. Disponível em: <<http://www.ingentaconnect.com/content/saf/fs/2001/00000047/00000003/art00004>>. Acesso em: 31/12/2015.

HEVIA, A.; CAO, Q. V.; ÁLVAREZ-GONZÁLEZ, J. G.; RUIZ-GONZÁLEZ, A. D.; GADOW, K. VON. Compatibility of whole-stand and individual-tree models using composite estimators and disaggregation. **Forest Ecology and Management**, v. 348, p. 46–56, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112715001747>>. Acesso em: 31/12/2015.

ISLAM, M. N.; KURTTILA, M.; MEHTÄTALO, L.; HAARA, A. Analyzing the effects of inventory errors on holding-level forest plans: the case of measurement error in the basal area of the dominated tree species. **Silva Fennica**, v. 43, n. 1, p. 71–85, 2009. Finnish Society of Forest Science.

MARTINS, F. B.; SOARES, C. P. B.; SILVA, G. F. DA. Individual tree growth models for eucalyptus in northern Brazil. **Scientia Agricola**, v. 71, n. 3, p. 212–225, 2014. SciELO Brasil.

MENDONÇA, A. R. DE E. **Modelagem não linear do crescimento e da produção de plantações florestais**, 2010. 84f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras, 2010.

MÔRA, R. **Funções de afilamento de forma variável e modelagem de efeitos mistos em fustes de *Pinus taeda* e *Eucalyptus saligna***, 2015. 277f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2015.

OLIVEIRA, J. D. C. **Modelagem de Crescimento e Produção de povoamentos de eucalipto empregando mínimos quadrados generalizados**, 2012. 75f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Viçosa, Viçosa, 2012.

PIETILÄ, I.; KANGAS, A.; MÄKINEN, A.; MEHTÄTALO, L. Influence of growth prediction errors on the expected losses from forest decisions. **Silva Fennica**, v. 44, n. 5, p. 829–843, 2010.

PINHEIRO, J.; BATES, D.; DEBROY, S.; SARKAR, D. nlme: Linear and Nonlinear Mixed Effects Models. Disponível em: <<http://cran.r-project.org/package=nlme>>. Acesso em: 31/12/2015.

R DEVELOPMENT CORE TEAM, R. D. C. R: A language and environment for statistical computing. **R foundation for Statistical Computing**, 2005. Vienna, Austria: sn.

UZOH, F. C. C. C.; OLIVER, W. W. Individual tree diameter increment model for managed even-aged stands of ponderosa pine throughout the western United States using a multilevel linear mixed effects model. **Forest ecology and management**, v. 256, n. 3, p. 438–445, 2008. Elsevier. Disponível em: <<http://www.treearch.fs.fed.us/pubs/40984#.VoV-xYIcjD4.mendeley>>. Acesso em: 31/12/2015.

VISMARA, E. DE S.; MEHTÄTALO, L.; BATISTA, J. L. F. Linear mixed-effects models and calibration applied to volume models in two rotations of *Eucalyptus grandis* plantations. **Canadian Journal of Forest Research**, p. 132–141, 2015. NRC Research Press. Disponível em: <<http://www.nrcresearchpress.com/doi/abs/10.1139/cjfr-2014-0435#.VoWA8-JYWug>>. Acesso em: 31/12/2015.

XU, H.; SUN, Y.; WANG, X.; et al. Nonlinear mixed-effects (NLME) diameter growth models for individual China-fir (*Cunninghamia lanceolata*) trees in southeast China. **PLoS ONE**, v. 9, n. 8, p. e104012, 2014. Public Library of Science. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104012>>. Acesso em: 31/12/2015.

ZHAO, L.; LI, C.; TANG, S. Individual-tree diameter growth model for fir plantations based on multi-level linear mixed effects models across southeast China. **Journal of Forest Research**, v. 18, n. 4, p. 305–315, 2013.

CAPÍTULO 2 – MODELAGEM DE CRESCIMENTO E PRODUÇÃO EM NÍVEL DE DISTRIBUIÇÃO DE DIÂMETROS COM E SEM EFEITO ALEATÓRIO

1. INTRODUÇÃO

Modelos de Crescimento e produção florestal são essenciais para todo o planejamento do manejo florestal, já que realizar a previsão do crescimento e produção de povoamentos é um pré-requisito para o planejamento da gestão das florestas em qualquer nível (DIAMANTOPOULOU et al., 2015).

O volume total dos povoamentos é suficiente para muitos propósitos de gestão e planejamento florestal eficaz, muitas vezes requer informações sobre a distribuição do volume por classe de diâmetro e de produtos (BURKHART; TOMÉ, 2012).

Modelos de distribuição de diâmetros não são tão onerosos como os modelos de árvores individuais e nem apresentam pouca flexibilidade como os modelos de povoamento total, tornando-se, portanto, uma alternativa a ser considerada. Seu objetivo é fornecer informações sobre a estrutura do povoamento, permitindo assim a estimativa da produção por classes de diâmetro. Com tais informações é possível realizar o planejamento do corte, de desbastes, os tratamentos silviculturais, estimar multiprodutos como por exemplo, madeira para serraria, celulose e energia (LEITE, 1990).

Modelos de distribuição de diâmetros apresentam como principal característica o uso de uma função de densidade e probabilidade (fdp) para descrição do perfil diamétrico das áreas de interesse (BINOTI, D. H. B. et al., 2015). Apesar da existência de diversas funções estatísticas, é predominante o uso de uma das formas da função Weibull (CAMPOS; LEITE, 2013), principalmente a Weibull de dois parâmetros (BINOTI et al., 2010; DIAMANTOPOULOU et al., 2015). Entre outras características, a boa correlação apresentada por seus parâmetros com características do povoamento tem sido fundamental à utilização dessa função (BINOTI, D. H. B. et al., 2013).

Existem métodos para realizar o ajuste de uma função probabilística, permitindo a estimativa dos parâmetros desta função, sendo os principais, o método dos percentis, o dos momento e da máxima verossimilhança (NORD-LARSEN; CAO, 2006). Destaca-se o método da máxima verossimilhança para ajustamento da função Weibull, pela disponibilidade de rotinas computacionais capazes de otimizar o

processo de estimativa dos parâmetros (ARAÚJO JÚNIOR et al., 2010; CRECENTE-CAMPO et al., 2010; CAMPOS; LEITE, 2013; DIAMANTOPOULOU et al., 2015; BINOTI, D. H. B. et al., 2015) .

Com a estimativa dos parâmetros da distribuição Weibull realizada, modelos de regressão são definidos para estimativa dos parâmetros utilizando variáveis do povoamento, e quando preditos, a distribuição acumulada pode ser calculada e o número de árvores em cada classe de diâmetros previamente definida pode ser estimado (MALTAMO; GOBAKKEN, 2014).

O avanço dos recursos de computação, trazendo maior poder de processamento trouxe consigo o uso de técnicas estatísticas alternativas, dentre as quais destaca-se os modelos de regressão não lineares generalizados, generalizados de efeito aleatório e de efeito aleatório (CARVALHO et al., 2011). Os mesmos autores citam trabalhos que envolvem a aplicação de tais modelos, como Budhathoki et al. (2008), Calegario et al. (2005a, b), Fang e Bailey (2001), Gregoire et al. (1995) e Pilar-Díaz e Couto (1999). Outras abordagens como o uso de redes neurais artificiais (BINOTI, D. H. B. et al., 2013; CASTRO et al., 2013; BINOTI, M. L. M. DA S. et al., 2013, 2015; DIAMANTOPOULOU et al., 2015), aprendizado de máquina e técnicas de sensoriamento remoto tem ganhado espaço (MALTAMO; GOBAKKEN, 2014; XU et al., 2014; GÖRGENS et al., 2015), dentre outros.

Como alternativa aos modelos de regressão empregados na modelagem de distribuição de diâmetros, o uso de modelos mistos permite identificar as diferentes fontes de variabilidade estocástica que não são explicadas pela parte fixa do modelo dividindo-se a variância residual em diferentes componentes (CALAMA; MONTERO, 2005). Neste tipo de modelo, as medidas dos indivíduos não precisam ser igualmente espaçadas e balanceadas e as análises podem ser conduzidas com os dados de indivíduos que tiveram medidas perdidas (parcelas que não foram medidas, árvores que não foram medidas dentro de parcelas, por exemplo) ou que apresentam ausência de informação em algum momento do estudo. Portanto, na ausência de dados, os modelos mistos oferecem mais uma vantagem de utilização frente aos outros tipos de modelos para dados longitudinais e de medidas repetidas (UZOH; OLIVER, 2008).

Na modelagem de efeitos aleatórios, os parâmetros de um modelo ajustado podem variar entre os grupos, sendo assim, o modelo pode ser calibrado para um determinado grupo afim melhorar a qualidade das previsões dentro do referido grupo (VISMARA et al., 2015).

O objetivo deste capítulo foi avaliar e comparar um modelo de distribuição de diâmetros que utiliza a abordagem fixa para ajuste e validação, com um modelo de distribuição de diâmetros que foi ajustado e validado com modelagem mista.

2. MATERIAL E MÉTODOS

2.1. Descrição dos dados

Os dados utilizados neste capítulo são descritos na seção 3.4 da tese. Foram selecionadas variáveis agrupadas por parcela e medição. Diante do volume calculado por classe de diâmetro, procedeu-se com a totalização dos volumes em $\text{m}^3 \text{ha}^{-1}$ por parcela, em cada idade.

A base de dados com as variáveis pareadas na idade 1 e idade 2 foi criada com a função *createPairedData* do pacote *Fgmutils* (FRAGA FILHO et al., 2015), criado durante o desenvolvimento deste trabalho e disponibilizado nos repositórios da linguagem R.

2.2. Ajuste da Função de Densidade e Probabilidade

Para proceder com o ajuste da fdp, os dados dos diâmetros a 1,3m de altura (DAP) observados em cada parcela em cada idade, foram agrupados em classes diamétricas com amplitude de 2,0 cm. Em seguida, para cada parcela e em cada idade, foi ajustada a função de densidade e probabilidade Weibull:

$$f(x) = \frac{\gamma}{\beta} \left(\frac{x}{\beta} \right)^{\gamma-1} e^{-\left(\frac{x}{\beta} \right)^\gamma} \quad (1)$$

em que: β é o parâmetro escala ($\beta > 0$), γ refere-se ao parâmetro de forma, ($\gamma > 0$), x é o centro da classe de diâmetro ($x > 0$).

A escolha da função Weibull de dois parâmetros se deu após vários testes com a mesma função com dois e três parâmetros e com dois e três parâmetros: truncada a esquerda, truncada a direita, truncada a esquerda e a direita. O ajuste da função Weibull foi realizado com o Método da Máxima Verossimilhança (BAILEY; DELL, 1973), com auxílio de rotinas implementadas na linguagem R (FRAGA FILHO; SCHMIDT, 2014).

2.3. Divisão da base de dados para ajuste e validação dos modelos

Com o objetivo de avaliar diferentes grupos na calibração de modelos mistos, os dados foram divididos utilizando-se dois critérios: no primeiro (critério 1), o objetivo foi avaliar o comportamento dos modelos para idades futuras, definindo a variável parcela como grupo. No segundo critério (critério 2), foram definidos grupos formados por variáveis relacionadas ao povoamento, disponíveis na base de dados do inventário.

De acordo com o critério 1, os dados foram divididos em dois conjuntos: o primeiro para ajuste, e o segundo para validação dos modelos. Do conjunto de ajuste, participaram todas as parcelas com idade futura (idade 2) até 60 meses, e no segundo, de projeção (ou validação), foram selecionadas as mesmas parcelas do grupo de ajuste, porém com idade futura igual ou superior a 72 meses.

De acordo com o critério 2, o conjunto de parcelas foi dividido aleatoriamente também em dois grupos, sendo 70% das parcelas utilizadas para ajuste dos modelos e 30% para validação, ou seja, de acordo com este critério, as parcelas do conjunto de ajuste não estavam presentes no conjunto de validação.

2.4. Modelo de Distribuição de diâmetros

O modelo de distribuição de diâmetros utilizado foi (CAMPOS; LEITE, 2013):

$$\beta_2 = \beta_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) d \max_2 + \varepsilon \quad (2)$$

$$\gamma_2 = \gamma_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \varepsilon \quad (3)$$

$$d \max_2 = d \max_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \varepsilon \quad (4)$$

$$N_2 = N_1 \cdot e^{(-\alpha_0)(\frac{I_2^{c_1} - I_1^{c_1}}{I_2^{c_1} - I_1^{c_1}})} + \varepsilon \quad (5)$$

em que: I_1 e I_2 são as idades atual e futura, em meses; γ_1 e γ_2 são os parâmetros de forma da função Weibull nas idades atual e futura, β_1 e β_2 são os parâmetros de escala da função Weibull nas idades atual e futura; $d \max_1$ e $d \max_2$ são os diâmetros máximos nas idades atual e futura, em cm, N_1 e N_2 são o número de árvores por hectare nas idades atual e futura, α_i , com $i=0$ a 1 são parâmetros dos modelos e ε , o erro.

Com o ajuste da função Weibull realizado, realizou-se o ajuste das expressões de (2) a (5) pelo método iterativo *Levenberg-Marquardt*, implementado na função *nlsLM* do pacote *minpack.lm* (ELZHOV et al., 2015) da linguagem R. Os modelos não lineares mistos foram ajustados pela função *nlme* do pacote *nlme* (PINHEIRO et al., 2014).

De posse das estimativas dos parâmetros e da função Weibull, calculou-se a probabilidade estimada do número de árvores por classe de diâmetro, de acordo com a expressão (6):

$$p(x_i, x_s) = \exp\left[-\left(\frac{x_i}{\beta}\right)^\gamma\right] - \exp\left[-\left(\frac{x_s}{\beta}\right)^\gamma\right] \quad (6)$$

em que: x_i é o limite inferior da classe de diâmetro, x_s é o limite superior, β e γ definidos anteriormente.

Com a probabilidade por classe calculada, obteve-se o número de árvores por hectare em cada classe de diâmetro pelo produto deste valor pelo número de árvores estimadas por hectare (expressão 5).

A altura total na idade futura foi obtida pela equação ajustada apresentada na seção 3.4. Em seguida, o volume por classe de diâmetro foi obtido pelo produto da equação de volume apresentada na seção 3.4 da tese e o número de árvores estimado, por classe diamétrica.

2.5. Alternativas de modelagem

Tendo como referência o modelo de distribuição de diâmetros apresentado, foram propostos dois conjuntos de alternativas de modelagem de distribuição de diâmetros com modelos mistos, com o objetivo de avaliar seu desempenho em termos de acurácia e ausência de viés comparado com a abordagem de modelagem fixa.

O primeiro conjunto é apresentado na Tabela 1. A alternativa 1_0 representa a modelagem com efeitos fixos. A alternativa 1_1 utiliza efeitos aleatórios, baseada no modelo de distribuição de diâmetros apresentado, cuja calibração considera a parcela como grupo de efeitos aleatórios. Seu ajuste e validação foi realizado com o conjunto de dados definido no critério 1.

Tabela 1 – Alternativa de modelagem de modelagem de distribuição de diâmetros usando modelos mistos, avaliadas para projeção de volume tendo como grupos de efeitos aleatórios a parcela.

Alternativa	Modelo de distribuição de diâmetros
1_0	$\beta_2 = \beta_1 \left(\frac{l_1}{l_2}\right) + \alpha_0 \left(1 - \frac{l_1}{l_2}\right) d \max_2 + \varepsilon$ $\gamma_2 = \gamma_1 \left(\frac{l_1}{l_2}\right) + \alpha_0 \left(1 - \frac{l_1}{l_2}\right) + \varepsilon$

(7)

$$d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

$$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) d \max_2 + \varepsilon$$

$$1_{-1} \quad \gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon \quad (8)$$

$$d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

em que: a_0 é o coeficiente aleatório estimado para cada parcela.

Com o propósito de avaliar alternativas além das propostas na Tabela 1, que não utilizassem a parcela como grupo de efeitos aleatórios na calibração, foram definidas outras variáveis e a partir delas, alternativas, ajustadas e validadas com o conjunto de dados definidos no critério 2. Como opções ao uso da parcela, apresentam-se o regime de manejo, espaçamento, o material genético, sendo estas variáveis candidatas a formarem grupos de efeitos aleatórios. Contudo, a base de dados utilizada neste trabalho apresenta apenas um regime de manejo e um tipo de espaçamento e o material genético não apresentou diferença significativa, sendo desconsiderado nesta avaliação. As variáveis definidas para serem grupos de efeitos aleatórios são apresentadas na Tabela 2.

Como já mencionado, depois de realizar o ajuste de diferentes tipos da função Weibull, verificou-se que a Weibull de dois parâmetros foi mais aderente aos dados. Diante disso, para cada parcela, e em cada idade, ela foi ajustada, e os parâmetros de forma (γ) e escala (β) foram estimados pelo método da máxima verossimilhança.

Dentre todas as parcelas e idades, o valor mínimo e máximo do parâmetro γ foi obtido, e definida uma amplitude de 1. Assim, foram criadas 28 classes de forma (*classeForma*) e cada parcela, em cada idade, recebeu uma classificação de acordo com o valor apresentado pelo parâmetro de forma. O mesmo foi feito para o parâmetro escala, sendo a variável *classeEscala* criada para representar a classificação da parcela, em determinada idade, quanto ao valor da variável escala.

Deste modo, uma parcela em determinada idade que apresente um valor do parâmetro de forma de 7,032, é classificada como pertencente à classe 7 quanto à forma. A mesma parcela apresenta o valor de 19,71 para o parâmetro escala, então sua classificação quanto à escala é 18.

Outra variável categórica criada na base de dados denominada *classificacaoWeibull*, considera uma classificação simultânea (dupla entrada) de forma e escala. Neste caso e considerando o exemplo anterior, para a parcela seu valor único seria de 7 e 18, representado pela concatenação de ambos, definindo a classificação 7_18 da parcela na respectiva idade.

Tabela 2 – Grupos de efeitos aleatórios criados para ajuste das alternativas propostas utilizando com modelos mistos.

Grupo	Variável agrupada	Amplitude
<i>classeForma</i>	Forma na idade atual (adimensional)	1
<i>classeEscala</i>	Escala na idade atual (adimensional)	1
<i>classificacaoWeibull</i>	Junção dos grupos <i>classeForma</i> e <i>classeEscala</i>	-

Considerando os novos grupos de efeitos aleatórios apresentados na Tabela 2, foram propostas novas alternativas de modelagem de efeitos aleatórios denominado conjunto de alternativas 2 (alternativas de 2 a 4), apresentadas na Tabela 3, ajustadas e validadas com os dados do critério 2.

Tabela 3 – Alternativas de modelagem avaliadas para projeção de volume em nível de distribuição de diâmetros, em que 1 representa a alternativa com efeitos fixos e 2 a 4 alternativas considerando efeitos aleatórios.

Alternativas	Modelo de distribuição de diâmetros
1	$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) d \max_2 + \varepsilon$ $\gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$ $d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$ $N_2 = N_1 \cdot e^{(-\alpha_0)(l_2^{\alpha_0} - l_1^{\alpha_0})} + \varepsilon$
2	$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) d \max_2 + \varepsilon$

$$\gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

$$d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

$$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + w_0) \left(1 - \frac{l_1}{l_2} \right) d \max_2 + \varepsilon$$

$$3 \quad \gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + w_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

$$d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + w_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

$$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + z_0) \left(1 - \frac{l_1}{l_2} \right) d \max_2 + \varepsilon$$

$$4 \quad \gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + z_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

$$d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + z_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$$

em que: a_0 é o coeficiente aleatório estimado o grupo “classificacaoWeibull”, w_0 = coeficiente aleatório estimado o grupo “classeForma” e z_0 = coeficiente aleatório estimado o grupo “classeEscala”

2.6. Avaliação dos resultados

A aderência dos ajustes realizados por Weibull aos dados foi avaliada pelo teste de Kolmogorov-Smirnov (MASSEY JR, 1951), que compara a frequência acumulada estimada ($F_e(x)$) com a observada ($F_o(x)$), sendo a classe de maior divergência absoluta a estatística do teste (D):

$$D = \max |F_o(x) - F_e(x)|$$

Sendo assim, quanto maior o valor de D , maior é considerada a divergência entre os valores observados e os estimados pela função.

A avaliação das alternativas foi realizada pelas seguintes estatísticas:

a) Coeficiente de correlação (PIETILÄ et al., 2010; MÔRA, 2015):

$$r_{\hat{y}} = \frac{n^{-1}(\sum(\hat{y}_i - \hat{y}_m)(y_i - \bar{y}))}{\sqrt{(n^{-1}\sum(\hat{y}_i - \hat{y}_m)^2)(n^{-1}\sum(y_i - \bar{y})^2)}} \quad (9)$$

b) BIAS(%) (UZOH; OLIVER, 2008):

$$BIAS = \frac{\sum_{i=1}^n (\hat{Y} - Y)}{n} \quad (10)$$

$$BIAS(\%) = \frac{BIAS}{\sum_{i=1}^n Y/n} \cdot 100 \quad (11)$$

c) Raiz quadrada do erro quadrático médio (RMSE %) (*Root Mean Square Error*) (ISLAM et al., 2009; ZHAO et al., 2013; ARIAS-RODIL et al., 2015):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n}} \quad (12)$$

$$RMSE(\%) = \frac{RMSE}{\sum_{i=1}^n Y/n} \cdot 100 \quad (13)$$

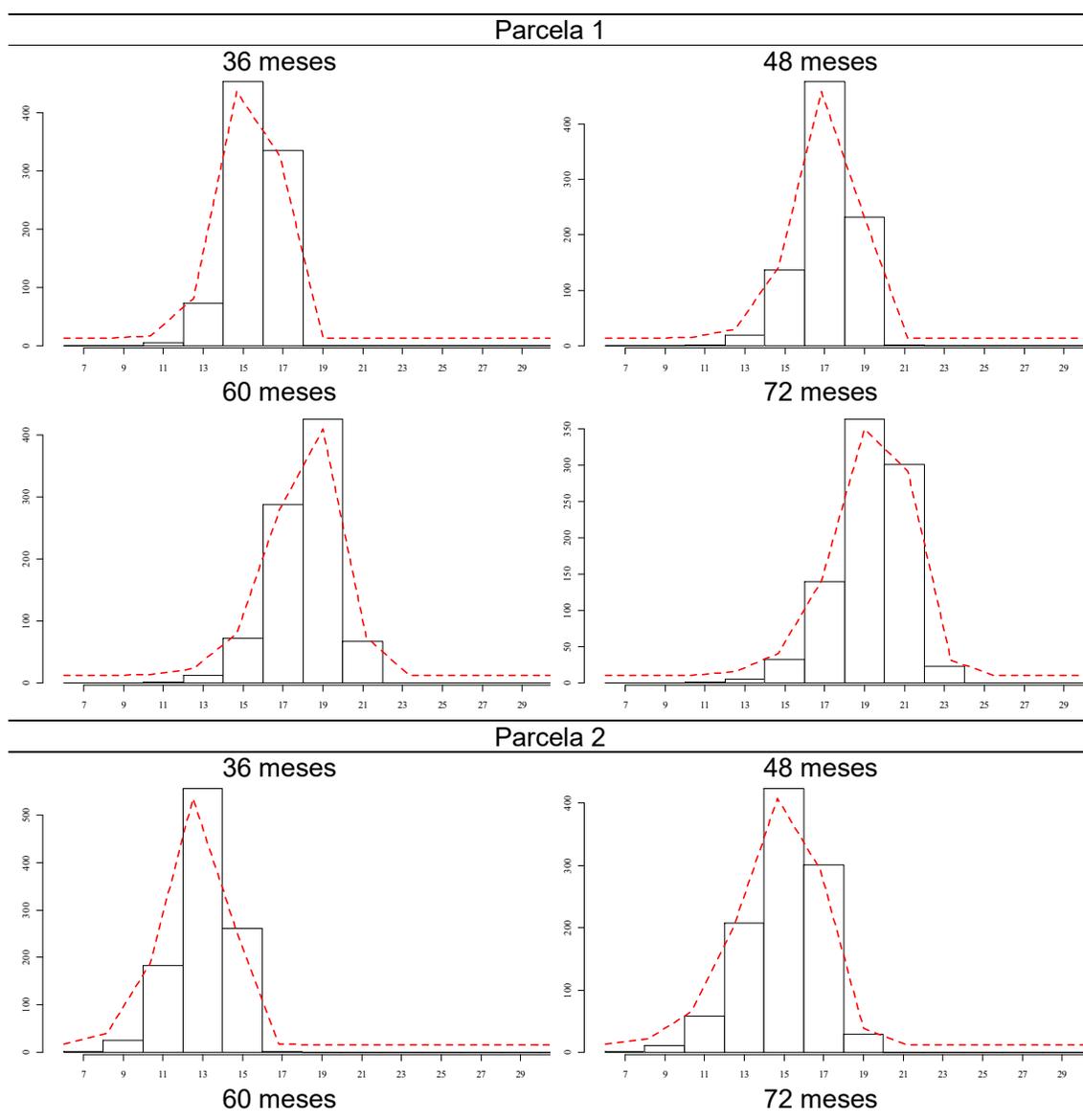
De forma a complementar a análise foram elaborados gráficos relacionando os volumes observados com os volumes estimados e gráficos de resíduos percentuais, obtidos pela expressão (MENDONÇA, 2010; MARTINS et al., 2014):

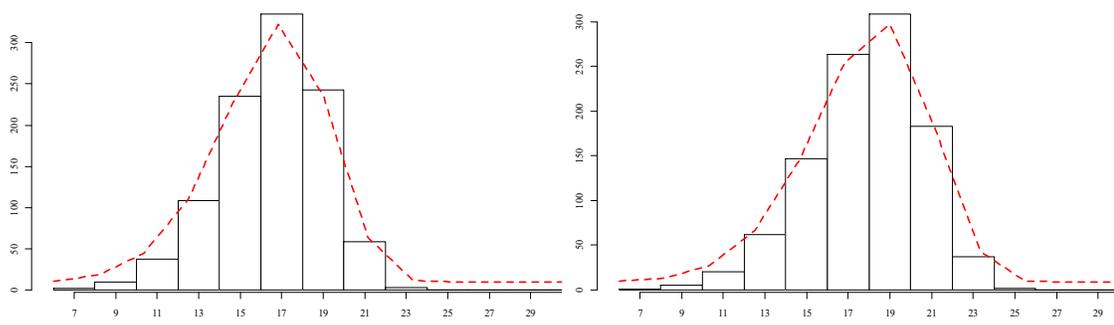
$$res(\%) = \left(\frac{y_i - \hat{y}_i}{y_i} \right) \cdot 100 \quad (14)$$

3. RESULTADOS E DISCUSSÃO

Ajustou-se a função Weibull de dois parâmetros para todas as parcelas em cada situação, sendo cada estimativa comparada com a distribuição observada. Todos os ajustes resultaram em aderência aos dados pelo teste de Kolmogorov-Smirnov a 1% de probabilidade, sendo o valor médio da estatística D igual a 0,0989. Na Figura 1 são apresentadas a distribuição de frequência de duas parcelas selecionadas aleatoriamente, com seus respectivos ajustes.

Figura 1 – Frequência observada e estimada pela função Weibull (---) em duas parcelas de eucalipto com idades de 36, 48, 60 e 72 meses, com amplitude de classe de diâmetro de 2 cm





Para o caso das estimativas de sobrevivência, isto é, o número de árvores por hectare, foi testada a abordagem mista para o ajuste da expressão 5, sem, contudo, obter-se resultados favoráveis. Sendo assim, este modelo foi ajustado considerando apenas os efeitos fixos. A equação de sobrevivência ajustada foi: $N_2 = N_1 \cdot e^{-(0.0000073)(I_2^{1.45271} - I_1^{1.45271})}$. As estatísticas de ajuste e de validação são apresentadas na Tabela 4.

Tabela 4 – Estatísticas utilizadas para avaliar o ajuste da equação para projeção de número de árvores por hectare ($N \text{ ha}^{-1}$), por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}\hat{y}}$	RMSE (%)	BIAS (%)
Ajuste				
1	S = 21	0,9991	0,42	-0,0780
	S = 27	0,9855	1,42	-0,0207
	S = 33	0,9968	0,66	0,0422
Validação				
1	S = 21	0,9670	1,56	-0,1236
	S = 27	0,9976	0,56	-0,0638
	S = 33	0,9999	0,08	-0,0792

Verifica-se que, tanto para o ajuste quanto para a validação realizada, o modelo se ajustou bem aos dados, e que sua capacidade de predição é adequada ao seu propósito.

Em seguida são apresentados os resultados do ajuste e validação dos grupos de alternativas propostas para a modelagem de distribuição de diâmetros utilizando dados do conjunto denominado critério 1, e a outra, do conjunto definido como critério 2.

3.1. Critério 1 – Alternativas usando a variável parcela como grupo de efeito aleatório

Nesta seção são apresentados os resultados do ajuste e aplicação da alternativa de modelagem de distribuição de diâmetros usando o modelo de efeitos

fixos e alternativas de modelos mistos, e parcela como grupo de efeitos aleatórios (Tabela 1), ajustadas e validadas com o conjunto de dados do critério 1. Testes foram realizados utilizando a classe diamétrica como grupo de efeitos aleatórios, inclusive adotando a abordagem multinível, na qual foram definidos 2 grupos aninhados de efeitos aleatórios, mas os resultados foram os mesmos, mantendo-se a parcela como grupo de efeitos aleatórios.

Na seção 3.1.1 são apresentados os resultados de cada submodelo do modelo de distribuição de diâmetros ajustado, e na seção seguinte, 3.1.2, os resultados das estatísticas comparando os volumes totais observados e estimados pela alternativa 1_1.

3.1.1. Resultado das Alternativas de modelagem de distribuição de diâmetros

O parâmetro escala na idade futura foi ajustado considerando o diâmetro máximo na idade futura como variável independente. Desta forma, foi realizado o ajuste de alternativas para obtenção desta variável e o resultado das estatísticas obtidas é apresentado na Tabela 5.

Observa-se que os resultados para a classe de produtividade Média ($S = 27$) foi mais acurado e menos viesado, com correlação de 0,93. Cabe destacar que é na classe de produtividade Média ($S = 27$) onde está a maior parte das parcelas utilizadas no ajuste (81% das parcelas), seguido pela classe de produtividade baixa ($S = 21$) com 14% das parcelas e a alta ($S = 33$) com 5% do total de parcelas. Entretanto, considerando a validação realizada (Tabela 6), a classe apresenta-se com viesada, com tendência de subestimação, assim como as demais classes.

No ajuste, observa-se a tendência de superestimação na classe de produtividade Baixa ($S = 21$), com BIAS relativo de 3,19, e tendência contrária na classe de produtividade Alta ($S=33$), com tendência de subestimação, sendo o BIAS (%) de -1,54. Quando avaliada a acurácia, a classe de produtividade Baixa é mais acurada do que a classe Alta, com valores de RMSE relativo de 5.251 e 5.51, respectivamente.

Tabela 5 – Estatísticas utilizadas para avaliar o ajuste das alternativas para estimar o diâmetro máximo (d_{max}) na idade futura, por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9437	4,43	1,9686
	S = 27	0,9344	4,45	-0,7964
	S = 33	0,9243	5,76	-2,6053
1_1	S = 21	0,9382	5,25	3,1973
	S = 27	0,9309	4,52	0,3182
	S = 33	0,9211	5,51	-1,5445

Tabela 6 – Estatísticas utilizadas para avaliar a aplicação da alternativa para obtenção do diâmetro máximo na idade futura em uma base de dados de validação, por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9710	2,75	-0,8576
	S = 27	0,9732	4,18	-2,8608
	S = 33	0,9848	5,71	-5,0992
1_1	S = 21	0,9563	2,72	-1,3108
	S = 27	0,9702	2,81	-1,6644
	S = 33	0,9798	3,26	-2,8063

Foi ajustada uma equação para projeção do parâmetro escala, e as estatísticas calculadas são apresentadas na Tabela 7. A maior acurácia ocorre na classe de produtividade Alta (S=33), embora seja mais viesada, com BIAS relativo de 1.27, com tendência de superestimação. A acurácia obtida na classe de produtividade Média, apresenta RMSE relativo de 3.9 e BIAS (%) de -0.32, com leve tendência de subestimação, o que também ocorre ao verificar-se o resultado obtido para a classe de produtividade Baixa, com BIAS relativo de -0.683, sendo, mais viesada e menos acurada, com RMSE (%) de 4.78.

Tabela 7 – Estatísticas utilizadas para avaliar o ajuste das alternativas para estimar o parâmetro β_2 da função Weibull de 2 parâmetros, por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9356	4,69	0,2802
	S = 27	0,9415	4,15	-0,3319
	S = 33	0,9289	3,87	0,7141
1_1	S = 21	0,9453	4,79	-0,6832
	S = 27	0,9592	3,90	-0,3248
	S = 33	0,9531	3,50	1,2791

Verifica-se que a alternativa 1_1 é mais acurada e menos tendenciosa que a alternativa 1_0 na classe de produtividade Média (S = 27). Nas demais classes, a

alternativa 1_1 apresenta-se menos acurada e mais viesada, contudo com valores próximos aos apresentados na alternativa 1_0.

Ao realizar a validação das equações em uma base de dados independente, percebe-se que a alternativa 1_0 apresenta-se como mais acurada e menos viesada na classe de produtividade Média e Baixa, entretanto, não apresenta o mesmo comportamento classe de produtividade Alta (Tabela 8). Observa-se um viés acentuado desta alternativa na classe de produtividade Média, com tendência de superestimação, contrária à tendência obtida no ajuste, com BIAS relativo de 1.833.

Tabela 8 – Estatísticas utilizadas para avaliar a aplicação da alternativa para obtenção do parâmetro β_2 na idade futura em uma base de dados de validação, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9768	2,65	2,0405
	S = 27	0,9762	2,21	1,6476
	S = 33	0,9831	2,70	1,9203
1_1	S = 21	0,9695	2,67	1,9465
	S = 27	0,9730	2,42	1,8329
	S = 33	0,9896	2,17	1,8470

Os resultados da estimativa do parâmetro de forma na idade futura são apresentados na Tabela 9. Percebe-se uma acurácia moderada em todas as classes de produtividade para as alternativas apresentadas, fato que pode ter relação com a baixa correlação apresentada pelo parâmetro de forma com características do povoamento (GUIMARÃES, 1994).

A alternativa 1_0 é mais acurada, e menos viesada na classe de produtividade Média (S = 27), contudo apresenta maior viés na classe de produtividade Alta (S = 33)

Embora menos acuradas no ajuste, o comportamento é inverso na validação, contudo, ainda apresentando tendência, porém de superestimação. Apesar disso, as classes de produtividade Média e Alta apresentaram resultados mais acurados, com 9.95% na classe de produtividade Média (S = 27) e 9.3 de RMSE (%) na classe de produtividade Alta.

Tabela 9 – Estatísticas utilizadas para avaliar o ajuste da alternativa 1_1 para estimar o parâmetro γ_2 da função Weibull de 2 parâmetros, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9577	13,94	-2,498
	S = 27	0,9479	13,97	3,040
	S = 33	0,9445	18,38	14,116
1_1	S = 21	0,9281	15,95	-1,3316
	S = 27	0,9447	17,01	-5,6027
	S = 33	0,9213	15,70	-0,9368

Embora apresente RMSE relativo inferior, portanto mais acurado na classe de produtividade Baixa da alternativa 1_0 e que esta seja menos viesada, nas classes Média (S = 27) e Alta (S = 33), a alternativa 1_0 apresenta-se como menos acurada e mais viesada.

Tabela 10 – Estatísticas utilizadas para avaliar a aplicação da alternativa para obtenção do parâmetro γ_2 na idade futura em uma base de dados de validação, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9858	7,78	1,134
	S = 27	0,9816	10,20	6,272
	S = 33	0,9850	13,33	12,114
1_1	S = 21	0,9802	8,52	2,4363
	S = 27	0,9784	9,95	3,2326
	S = 33	0,9599	9,29	1,7927

Com os parâmetros forma e escala disponíveis na idade futura, foi possível calcular a função Weibull de dois parâmetros acumulada com o propósito de obter a estatística D do teste não paramétrico Kolmogorov-Smirnov, no qual obteve-se todos os resultados significativos a 5%, tanto para o ajuste, quando na validação. Foram selecionados resultados de parcelas usadas no ajuste, e na validação, apresentando a estatística D e o gráfico de probabilidade acumulada para as alternativas 1_0 e 1_1, apresentados na Figura 2 e na Figura 3, respectivamente, nas quais foram selecionadas aleatoriamente 4 parcelas para ajuste e outras quatro para apresentação dos resultados de validação.

Figura 2 – Frequência observada acumulada e frequências estimadas acumuladas, calculadas a partir do ajuste dos parâmetros forma e escala da função Weibull obtidas pelas equações ajustadas para as alternativas 1_0 (A_1_0) e 1_1 (A_1_1).

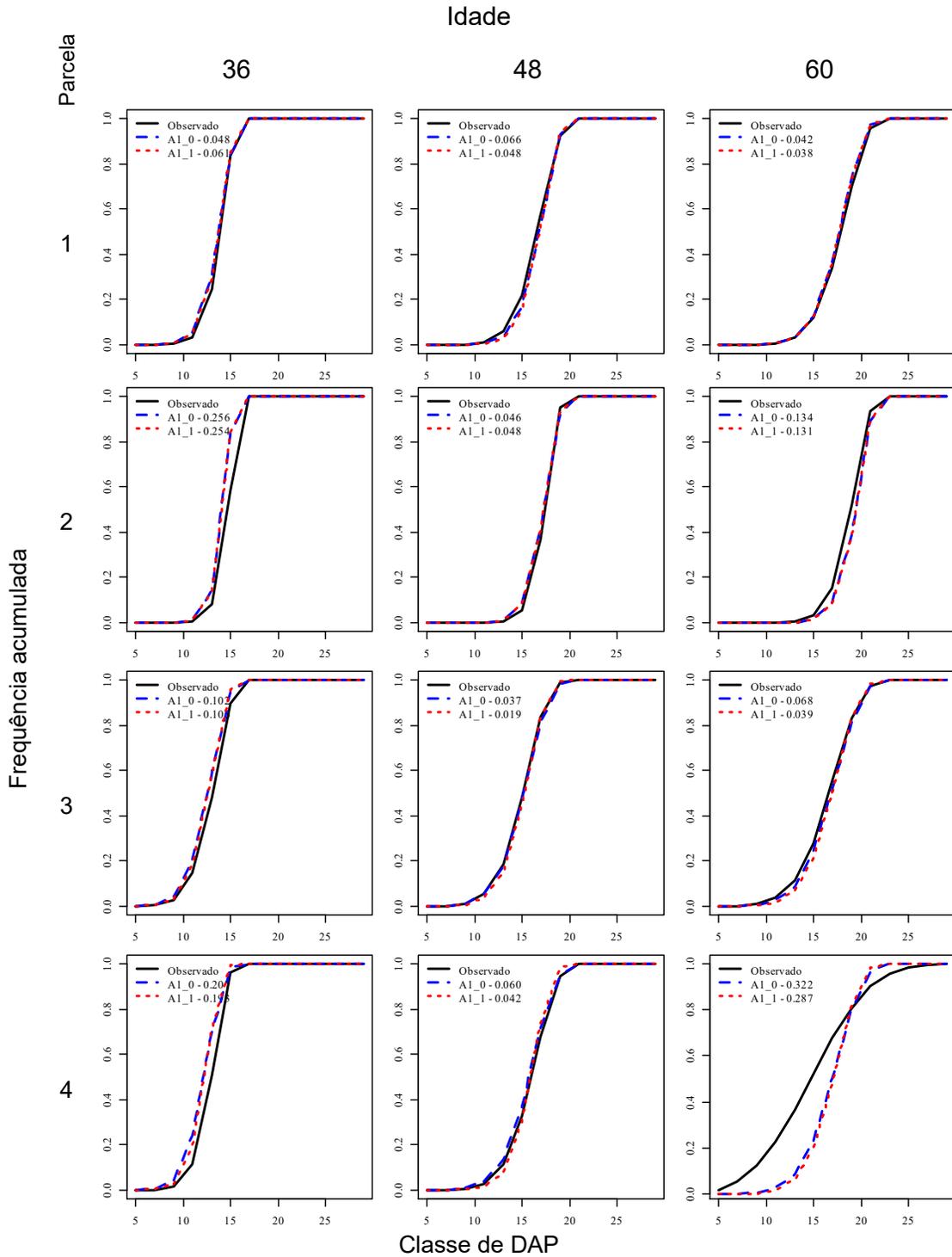
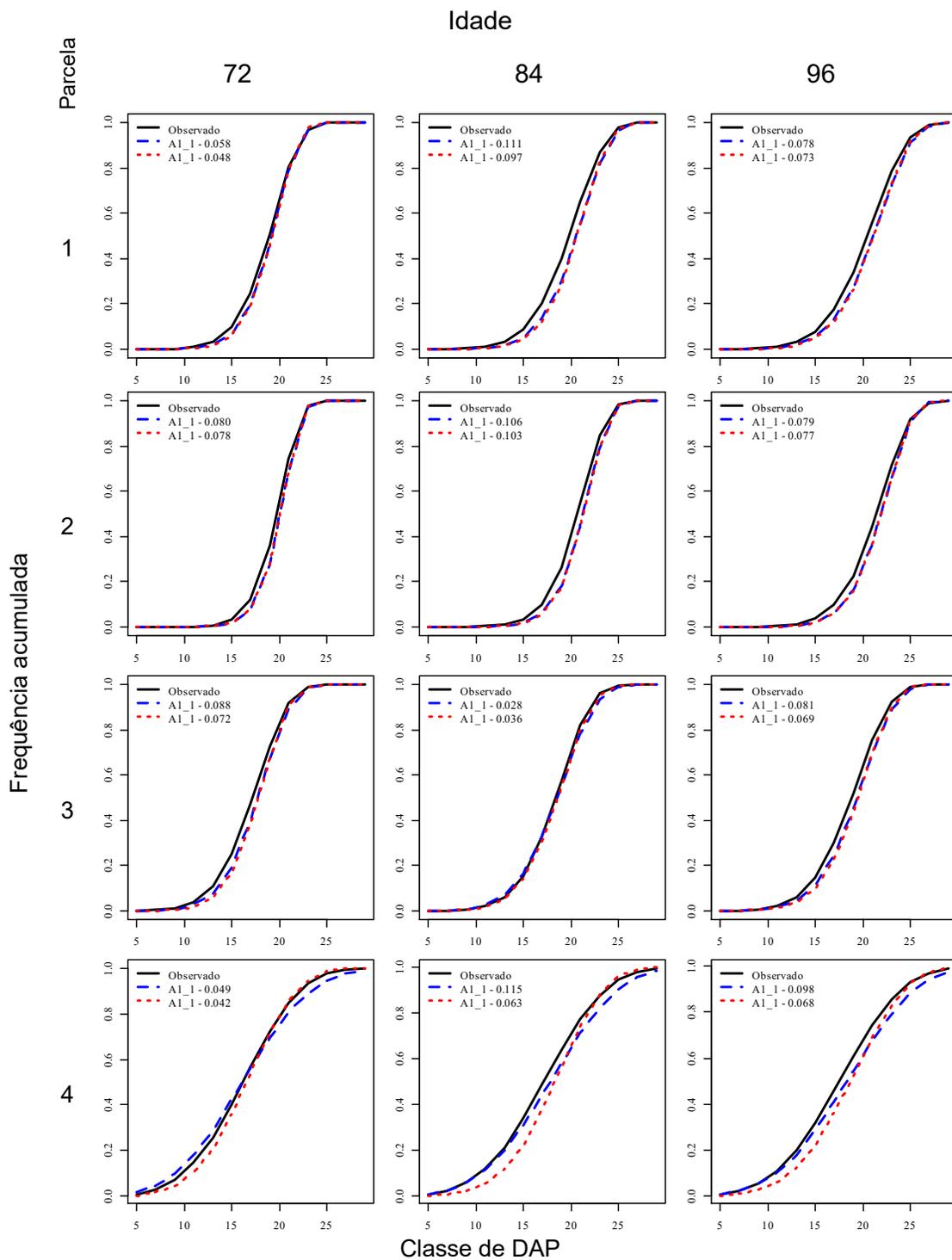


Figura 3 – Frequência observada acumulada e frequências estimadas acumuladas, calculadas a partir do das equações ajustas dos parâmetros forma e escala da função Weibull para a base de validação, para as alternativas 1_0 (A_1_0) e 1_1 (A_1_1).



3.1.2. Resultados da aplicação das alternativas de modelagem de distribuição de diâmetros

Com o ajuste e aplicação do modelo de distribuição de diâmetros apresentado na seção 3.1.1 realizado, obteve-se os volumes estimados por classe de diâmetro, que para em seguida foram totalizados por parcela (em $\text{m}^3 \text{ha}^{-1}$), para proceder a análise apresentada nesta seção. Uma parcela foi selecionada aleatoriamente para apresentação dos volumes observados e estimados por classes de diâmetros obtidos no ajuste (Figura 4) e na validação (Figura 5).

Figura 4 – Volumes por classe de diâmetro observados e estimados aos 36, 48 e 60 meses para uma parcela selecionada aleatoriamente com os dados de ajuste (critério 1).

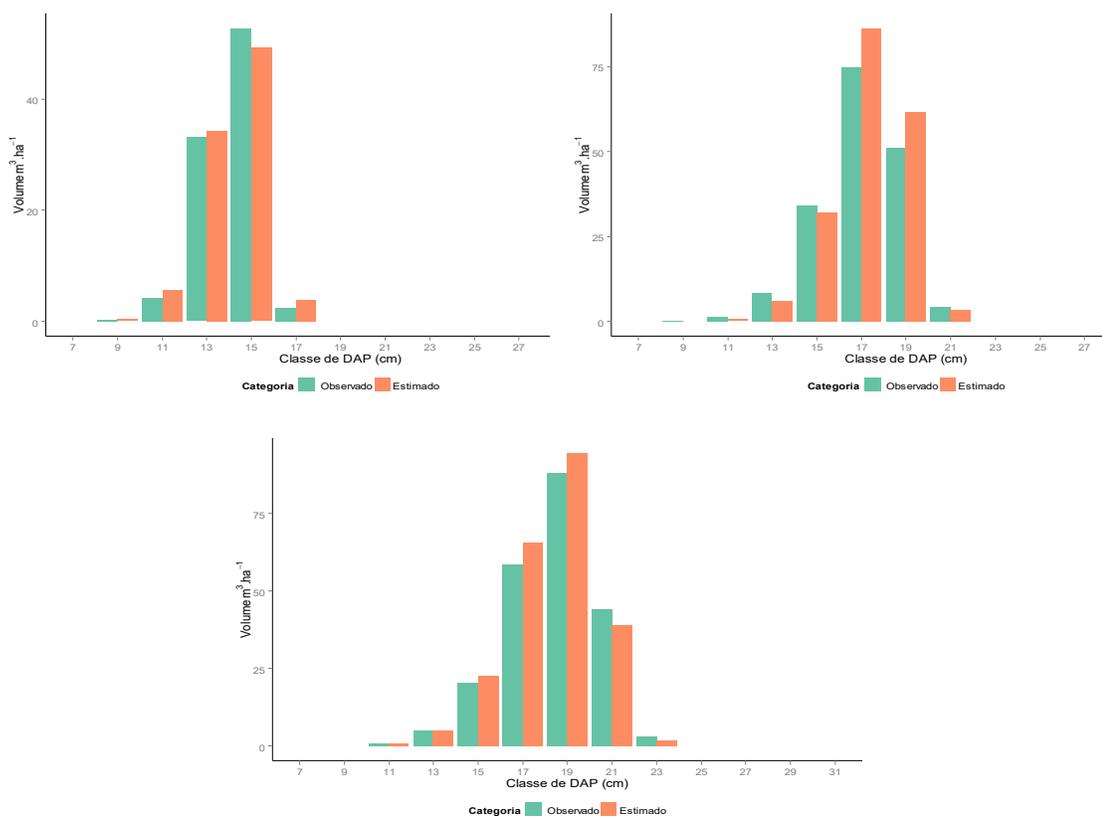
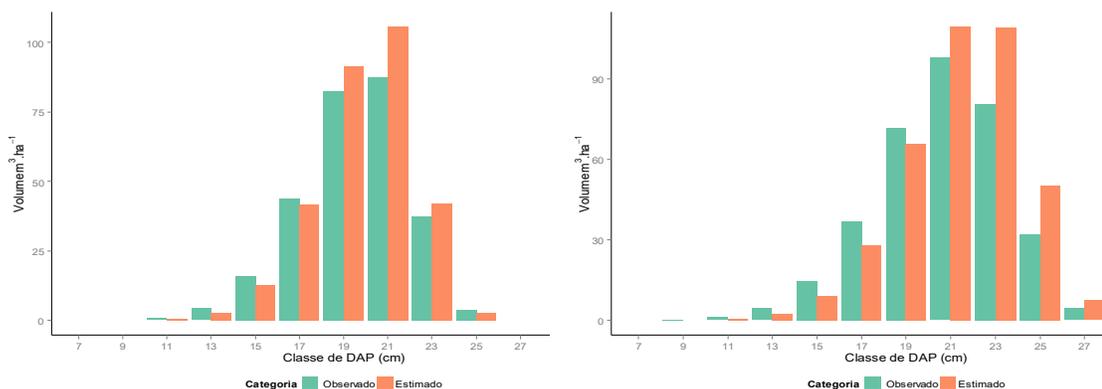


Figura 5 – Volumes por classe de diâmetro observados e estimados aos 72, 84 meses para a mesma parcela selecionada na Figura 4.



Avaliando-se as estatísticas apresentadas na Tabela 11, é possível concluir que a alternativa 1_1 apresenta uma acurácia moderada em todas as classes de produtividade, com RMSE relativo de 10.58 na classe de produtividade Média, com tendência média de superestimação. Nesta classe, para menores valores de volume a tendência pode ser confirmada na Figura 6 e na Figura 7. Em povoamentos com volume abaixo de 200 m³ ha⁻¹ a tendência é contrária, apresentando uma leve de subestimação em todas as classes de produtividade.

Tabela 11 – Estatísticas utilizadas para avaliar o volume totalizado por parcela (m³ ha⁻¹) obtido a partir das alternativas de modelagem de distribuição de diâmetros, detalhados por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9397	13,79	-1,3080
	S = 27	0,9475	11,75	2,2890
	S = 33	0,9466	13,30	8,0690
1_1	S = 21	0,9632	12,01	-2,257
	S = 27	0,9634	10,58	2,676
	S = 33	0,9799	10,13	5,944

Apesar das estatísticas para a validação apresentarem maior acurácia em relação aos dados de ajuste, os resultados tornaram-se mais viesados, com aumento expressivo do BIAS (%) em todas as classes de produtividade, tendendo à superestimação em todas as classes de produtividade, conforme apresentado na Figura 8 e na Figura 9.

Tabela 12 – Estatísticas utilizadas para avaliar o volume do povoamento ($\text{m}^3 \text{ha}^{-1}$) obtido por meio da aplicação das alternativas de modelagem de distribuição de diâmetros em uma base de dados independente, totalizados por parcela e detalhados por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1_0	S = 21	0,9670	6,56	4,3190
	S = 27	0,9662	8,31	6,8490
	S = 33	0,9885	11,88	11,2280
1_1	S = 21	0,9660	6,62	4,239
	S = 27	0,9610	8,95	7,377
	S = 33	0,9921	11,45	11,098

Nota-se pelos volumes observados *versus* estimados para as alternativas 1_0 e 1_1 (Figura 6), uma maior dispersão dos pontos na alternativa 1_0 em todas as classes de produtividade quando se avalia o ajuste do modelo.

Avaliando-se os gráficos de resíduos relativos, a tendência de subestimação é verificada nas parcelas com menores volumes e superestimação em parcelas com volumes maiores (Figura 7).

Figura 6 – Volumes totais por parcela estimados *versus* observados, obtidos pelas equações ajustadas referentes às alternativas de modelagem de distribuição de diâmetros, em cada classe de produtividade.

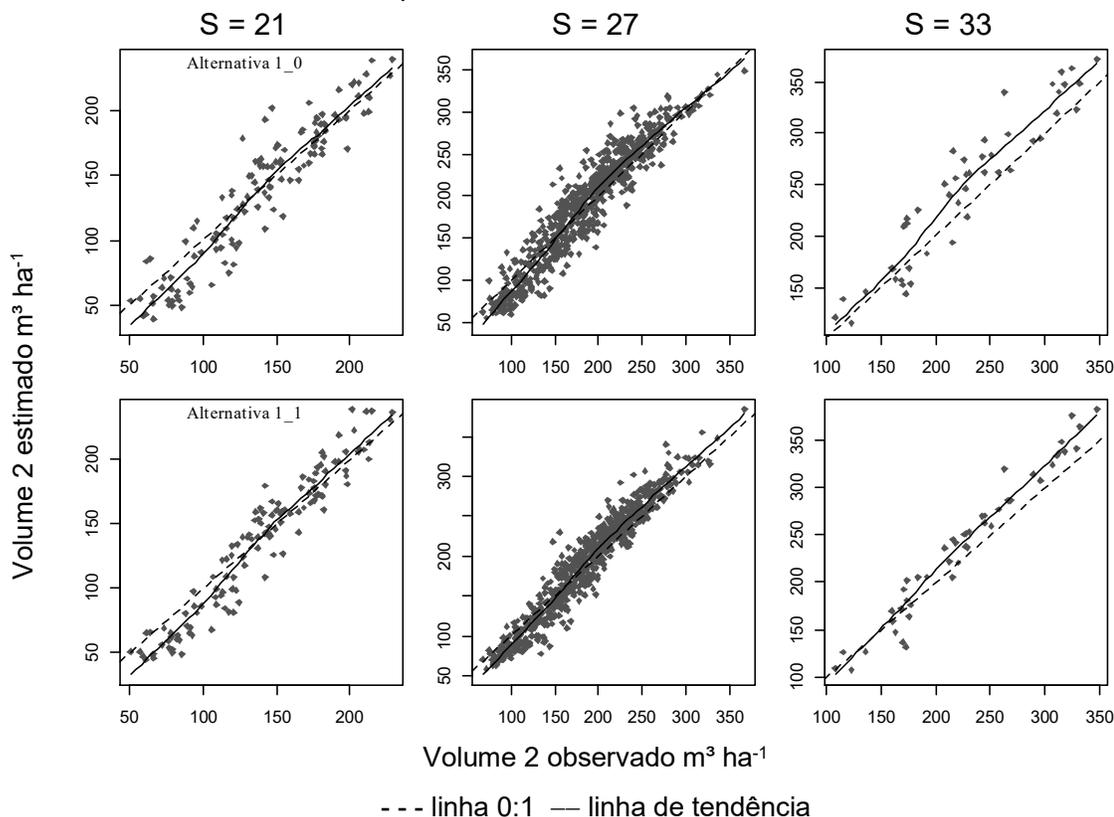
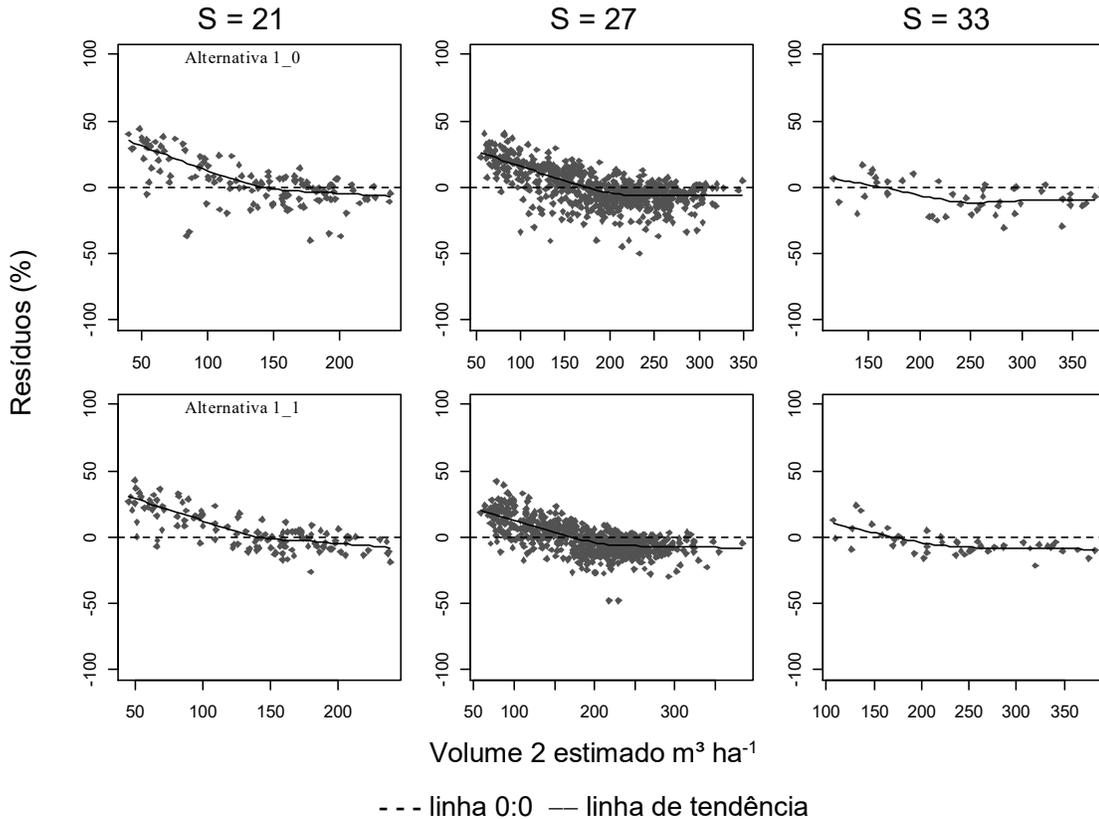
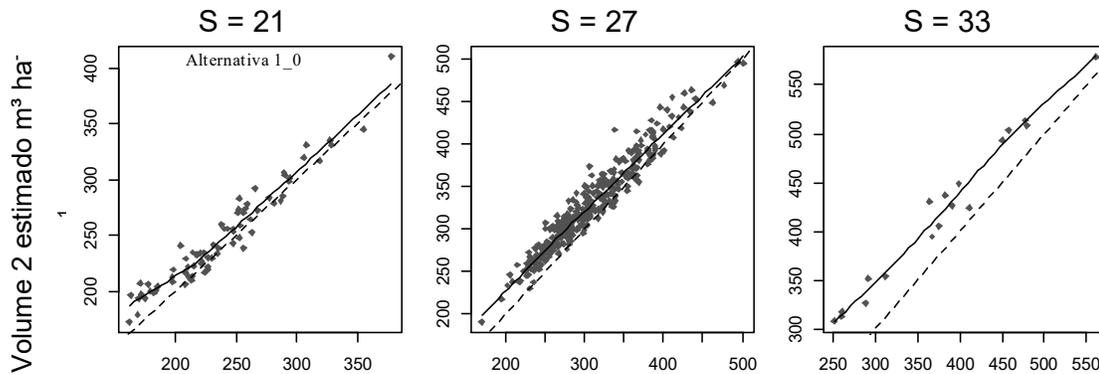


Figura 7 – Distribuição dos resíduos percentuais em função dos volumes totais estimados, em cada classe de produtividade para o modelo de distribuição de diâmetros proposto nas alternativas 1_0 e 1_1.



Considerando a validação, a tendência de superestimação é confirmada nas Figura 8 e na Figura 9 em todas as classes de produtividade.

Figura 8 – Volumes totais por parcela estimados *versus* observados, obtidos pela aplicação das alternativas 1_0 e 1_1 de modelagem de distribuição de diâmetros em uma base de dados independente, em cada classe de produtividade.



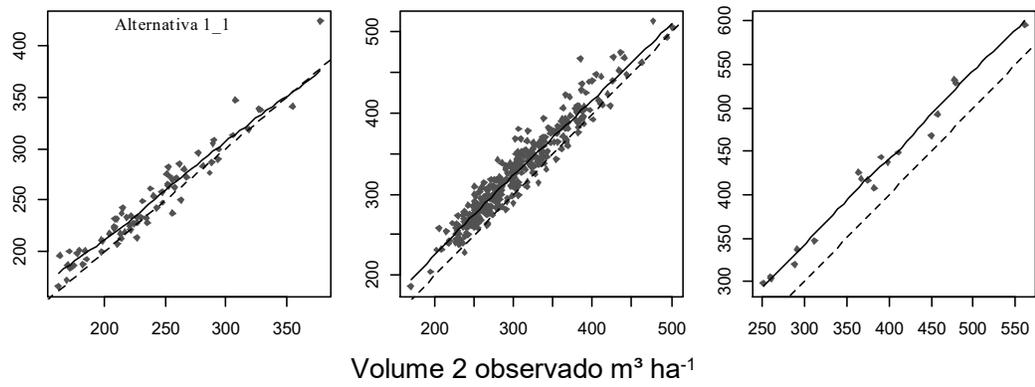
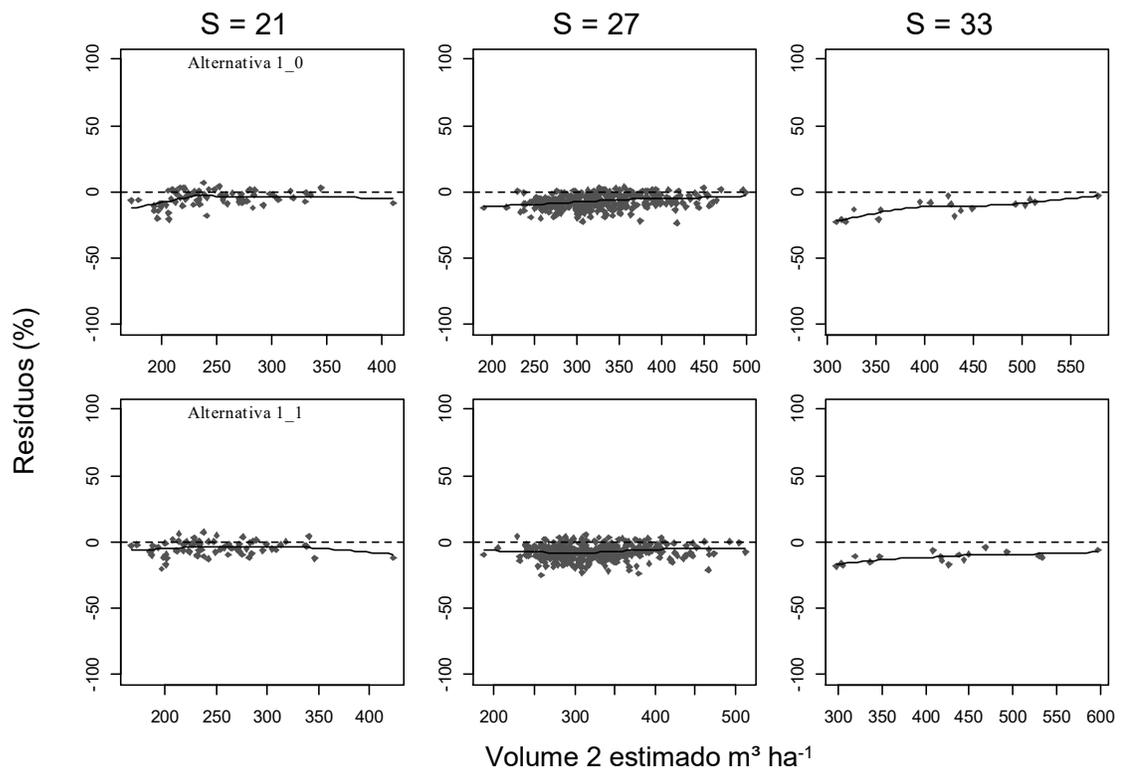


Figura 9 – Distribuição dos resíduos percentuais em função dos volumes totais estimados, em cada classe de produtividade para o modelo de distribuição de diâmetros proposto nas alternativas 1_0 e 1_1 para a base de dados de validação (aplicação).



3.2. Critério 2 – Alternativas usando outras variáveis como grupos de efeitos aleatórios

Nesta seção são apresentados os resultados do ajuste das alternativas de modelagem de distribuição de diâmetros que fazem parte do grupo 2, que utilizaram outras variáveis como grupos de efeitos aleatórios, e foram ajustadas e validadas com o conjunto de dados do critério 2.

3.2.1. Alternativas de modelagem para se estimar a distribuição de diâmetros

O ajuste de cada modelo que compõem o grupo 2 de alternativas de modelagem de distribuição de diâmetros foi realizado com o conjunto de dados do critério 2, e seus respectivos resultados são apresentados nesta seção.

Utilizada como variável independente no modelo representado pela equação (3), o diâmetro máximo na idade futura representado pela equação (5) foi ajustado considerando as quatro alternativas determinadas neste capítulo. Os resultados do ajuste são apresentados na Tabela 13.

Tabela 13 – Estatísticas utilizadas para avaliar o ajuste das quatro alternativas para estimar o diâmetro máximo (d_{max}) na idade futura, por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}\hat{y}}$	RMSE (%)	BIAS (%)
1	S = 21	0,9775	4,16	1,7611
	S = 27	0,9688	4,34	-1,0261
	S = 33	0,9483	6,38	-2,3991
2	S = 21	0,9761	3,23	1,1394
	S = 27	0,9771	2,99	-0,3304
	S = 33	0,9767	3,33	-0,2676
3	S = 21	0,9665	3,89	1,3226
	S = 27	0,9733	3,47	-0,7625
	S = 33	0,9737	4,14	-0,5828
4	S = 21	0,9727	3,65	1,5634
	S = 27	0,9633	3,86	-0,4846
	S = 33	0,9492	5,38	-1,8014

Verifica-se a superioridade da alternativa 2 (não linear mista), que utiliza a variável categórica “classificacaoWeibull” para agrupamento aleatório, com valores de RMSE (%) inferiores em todas as classes de produtividade para todas as alternativas definidas. A alternativa 2 apresenta ainda, valores de BIAS (%) próximos de zero, com leve tendência de superestimação na classe de produtividade baixa (S = 21), e subestimação nas classes média (S = 27) e alta (S = 33).

No caso da validação, as estatísticas apresentadas na Tabela 14 indicam que as alternativas 2 e 4 apresentaram resultados similares, com vantagem para a alternativa 2 para as estatísticas $r_{y\hat{y}}$ e $RMSE$ (%), mas no caso do $BIAS$ (%), a alternativa 4 mostrou-se, de forma geral, menos viesada.

Tabela 14 – Estatísticas utilizadas para avaliar a validação das alternativas para obtenção do diâmetro máximo na idade futura em uma base de dados de validação, com parcelas independentes daquelas usadas no ajuste.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1	S = 21	0,9532	4,45	1,7859
	S = 27	0,9615	4,00	0,1055
	S = 33	0,9759	4,31	-2,5763
2	S = 21	0,9610	3,84	1,7882
	S = 27	0,9632	3,56	0,7179
	S = 33	0,9597	3,55	-0,7130
3	S = 21	0,9469	4,41	1,9408
	S = 27	0,9574	3,81	0,4918
	S = 33	0,9517	4,26	-1,3200
4	S = 21	0,9572	3,94	1,7112
	S = 27	0,9570	3,77	0,3929
	S = 33	0,9663	3,44	-1,3531

As estimativas para dos coeficientes fixos das alternativas avaliadas são apresentadas na Tabela 15.

Tabela 15 – Estimativas dos coeficientes fixos das alternativas avaliadas para $d \max_2$

	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
$\hat{\alpha}_0$	25.78**	25.59**	24.94**	26.68**

** - significativo a 1% pelo Test T para modelos fixos, e Teste de Wald para modelos mistos

Em seguida foram ajustadas as alternativas para a modelagem do parâmetro escala na idade futura (β_2) da função Weibull de dois parâmetros para as alternativas definidas. Em destaque na Tabela 16, a alternativa 2 para estimativa do parâmetro escala na idade futura apresenta superioridade sob as demais alternativas avaliadas, cabendo salientar que o diâmetro máximo na idade futura obtido pela alternativa 2 (Tabela 14) foi utilizado como *input* no modelo representado pela equação 3 para obtenção do parâmetro escala.

Tabela 16 – Estatísticas utilizadas para avaliar o ajuste das quatro alternativas para estimar o parâmetro escala na idade futura da função Weibull de 2 parâmetros, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1	S = 21	0,9684	3,68	0,3026
	S = 27	0,9709	3,32	-0,0827
	S = 33	0,9475	3,67	1,0446
2	S = 21	0,9803	2,66	0,7404
	S = 27	0,9739	2,70	-0,2877
	S = 33	0,9512	3,33	0,0706
3	S = 21	0,9717	3,50	0,5584
	S = 27	0,9705	3,19	-0,1287
	S = 33	0,9455	3,55	0,5326
4	S = 21	0,9693	3,24	0,6309
	S = 27	0,9706	2,91	-0,3178
	S = 33	0,9437	3,60	0,5240

As estimativas para o coeficiente das equações ajustadas para as alternativas para o parâmetro escala da função Weibull são apresentados na Tabela 17.

Tabela 17 – Estimativas dos coeficientes fixos das alternativas avaliadas para β_2

	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
$\hat{\alpha}_0$	1.185**	1.241**	1.237**	1.235**

** - significativo a 1% pelo Test T para modelos fixos, e Teste de Wald para modelos mistos

As estatísticas resultantes das estimativas realizadas para o parâmetro de forma na idade futura são apresentadas na Tabela 18. Verifica-se a superioridade das alternativas não lineares mistas (2 a 4) sob a alternativa não linear ao observar-se os valores de RMSE (%) e BIAS (%). O valor de RMSE (%) para a classe de produtividade baixa (S = 21) da alternativa 1 apresenta valor de 14.24, sendo reduzido a 6.2920 na mesma classe de produtividade na alternativa 2, 8.42 na alternativa 3 e com uma redução menos expressiva na alternativa 4, com valor de RMSE (%) de 13.01.

Considerando o BIAS (%), a alternativa 2 apresenta valores de -1.5075 na classe de produtividade baixa (S = 21) e 0.31 para a classe de produtividade média (S = 27), indicando uma pequena tendência de subestimação e superestimação, respectivamente, porém inferior à alternativa 1, usada como referência para comparação. Apesar da classe de produtividade alta (S = 33) da alternativa 2 ter apresentado valor de RMSE (%) de 11.772, superior ao valor da mesma estatística na mesma classe na alternativa 3, o BIAS (%) é inferior, reiterando a superioridade da alternativa 2 frente às demais.

Tabela 18 – Estatísticas utilizadas para avaliar o ajuste das quatro alternativas para estimar o parâmetro γ_2 da função Weibull de 2 parâmetros, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1	S = 21	0,9569	14,24	-3,4100
	S = 27	0,9471	13,88	1,9830
	S = 33	0,9450	17,20	12,6370
2	S = 21	0,9903	6,29	-1,5075
	S = 27	0,9783	8,18	0,3145
	S = 33	0,9608	11,77	3,4077
3	S = 21	0,9834	8,43	-0,8622
	S = 27	0,9730	9,18	0,7864
	S = 33	0,9653	11,58	4,1370
4	S = 21	0,9656	13,01	-2,7910
	S = 27	0,9511	13,19	0,8186
	S = 33	0,9368	15,98	10,0797

As estimativas parâmetro de forma são apresentadas na Tabela 19.

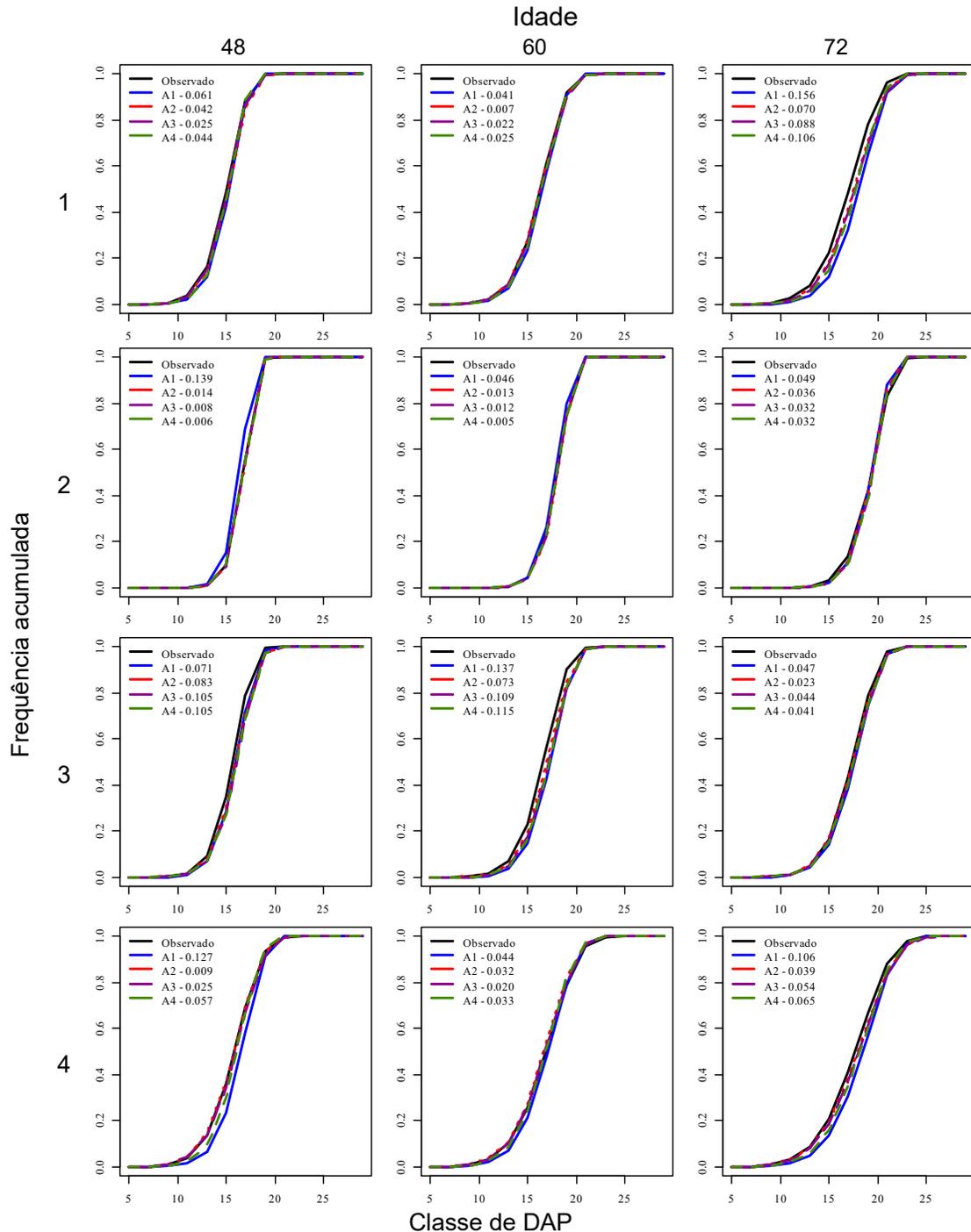
Tabela 19 – Estimativas dos coeficientes fixos das alternativas avaliadas para γ_2

	Alternativa 1	Alternativa 2	Alternativa 3	Alternativa 4
α_0	9.507**	12.9**	16.46**	9.794**

** - significativo a 1% pelo Test T para modelos fixos, e Teste de Wald para modelos mistos

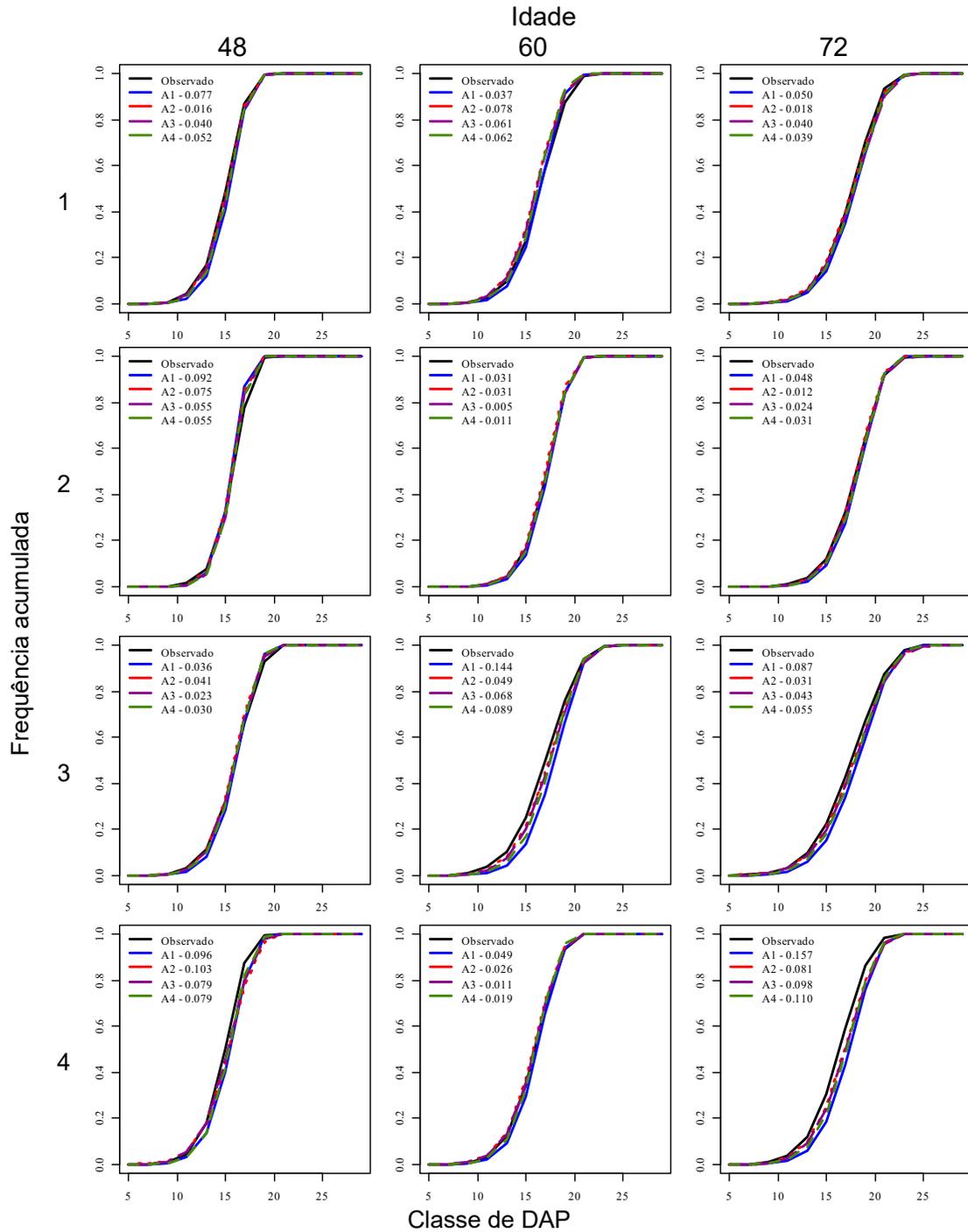
Com as estimativas dos parâmetros de escala e forma da função Weibull para as quatro alternativas, procedeu-se a avaliação da frequência observada acumulada *versus* as frequências estimadas para cada uma das alternativas, sendo a estatística D não significativa para todas as parcelas utilizadas neste capítulo. O gráfico (Figura 10) com quatro parcelas selecionadas aleatoriamente ajuda a ilustrar esta afirmativa.

Figura 10 – Frequência observada acumulada e frequências estimadas acumuladas, calculadas pelos valores ajustados para os parâmetros forma e escala da função Weibull obtidas pelas equações ajustadas para as quatro alternativas.



Verifica-se que o valor da estatística D para as alternativas que utilizam a modelagem mista é menor na maioria das parcelas seleccionadas aleatoriamente nas diferentes idades consideradas, indicando um melhor ajuste, o que pode ser confirmado nos gráficos apresentados na Figura 11.

Figura 11 – Frequência observada acumulada e frequências estimadas acumuladas, calculadas usando os parâmetros ajustados forma e escala da função Weibull em uma base de dados independente, para as quatro alternativas.



Os resultados confirmam a superioridade das alternativas mistas na modelagem de distribuição de diâmetros.

3.2.2. Alternativas de modelagem de distribuição de diâmetros para se estimar o volume por classe de diâmetro

Como resultado final do ajuste das alternativas de modelagem de distribuição de diâmetros do grupo 2, obteve-se os volumes por classe de diâmetros (Apêndice A – Volumes estimados e observados obtidos no ajuste e na validação das alternativas 1(não mista) e 2 (mista) para o conjunto de dados do critério 2.), os quais foram totalizados por parcela a cada período de medição, cujas estatísticas de ajuste são apresentadas na Tabela 20.

Tabela 20 – Estatísticas utilizadas para avaliar o volume total do povoamento ($m^3 \text{ ha}^{-1}$) obtido pelas quatro alternativas de modelagem de distribuição de diâmetros, totalizados por parcela e detalhados por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}\hat{y}}$	RMSE (%)	BIAS (%)
1	S = 21	0,9747	9,76	-0,0783
	S = 27	0,9771	9,30	3,0921
	S = 33	0,9746	12,59	8,8899
2	S = 21	0,9791	7,83	-0,0861
	S = 27	0,9778	7,82	1,7009
	S = 33	0,9754	9,18	5,4058
3	S = 21	0,9774	9,26	0,5565
	S = 27	0,9773	8,78	2,7735
	S = 33	0,9734	10,55	6,8876
4	S = 21	0,9714	9,17	-0,2988
	S = 27	0,9748	8,40	1,7321
	S = 33	0,9725	10,70	6,9380

Com os valores de RMSE (%) e BIAS (%) inferiores em todas as classes de produtividade, a alternativa 2 configura-se como a mais adequada. Apesar da alternativa 1 apresentar um valor inferior de BIAS (%) na classe de produtividade baixa (S = 21), sendo menos viesada, verifica-se que a magnitude da diferença é insignificante. O melhor desempenho, e conseqüente melhoria ocorreu na classe de produtividade alta (S = 33), com uma redução de BIAS (%) de 8.8899 para 5.4058.

Após a aplicação do modelo em uma base de dados independente, pode-se observar (Tabela 21) que a tendência de melhoria da acurácia das alternativas mistas é mantida, embora os resultados de validação não sejam tão acurados como os de ajuste, o que já era esperado. Nota-se também que a alternativa 4 apresentou resultados melhores que a alternativa 2 no caso da validação.

Tabela 21 – Estatísticas utilizadas para avaliar o volume ($\text{m}^3 \text{ha}^{-1}$) obtido por meio da aplicação das alternativas de modelagem de distribuição de diâmetros em uma base de dados independente, totalizados por parcela e detalhados por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
1	S = 21	0,9730	10,05	1,3674
	S = 27	0,9710	10,34	3,6127
	S = 33	0,9793	11,09	7,3315
2	S = 21	0,9628	10,08	1,0525
	S = 27	0,9701	8,79	2,4770
	S = 33	0,9750	10,33	6,0666
3	S = 21	0,9755	9,75	1,5028
	S = 27	0,9694	10,08	3,4849
	S = 33	0,9844	9,72	6,4210
4	S = 21	0,9691	9,15	0,9870
	S = 27	0,9697	8,90	2,2548
	S = 33	0,9799	9,38	5,6704

Na Figura 12, corrobora-se a acurácia da alternativa 2 na classe de produtividade Média (S = 27), semelhante à alternativa 4 na mesma classe de produtividade. As tendências de superestimação são confirmadas em todas as classes de produtividade em todas as alternativas, contudo mais acentuadas na classe de produtividade Alta (S = 33).

Analisando-se a dispersão residual percentual (Figura 13), verifica-se uma tendência de superestimar o volume em povoamentos menos produtivos, em todas as classes de produtividade. Apesar da presença de pontos dispersos, é perceptível a melhoria apresentada na alternativa 2 comparada às demais alternativas, em todas as classes de produtividade.

Figura 12 – Volumes totais por parcela estimados *versus* observados, obtidos pelas equações ajustadas referentes às quatro alternativas de modelagem de distribuição de diâmetros avaliadas, em cada classe de produtividade.

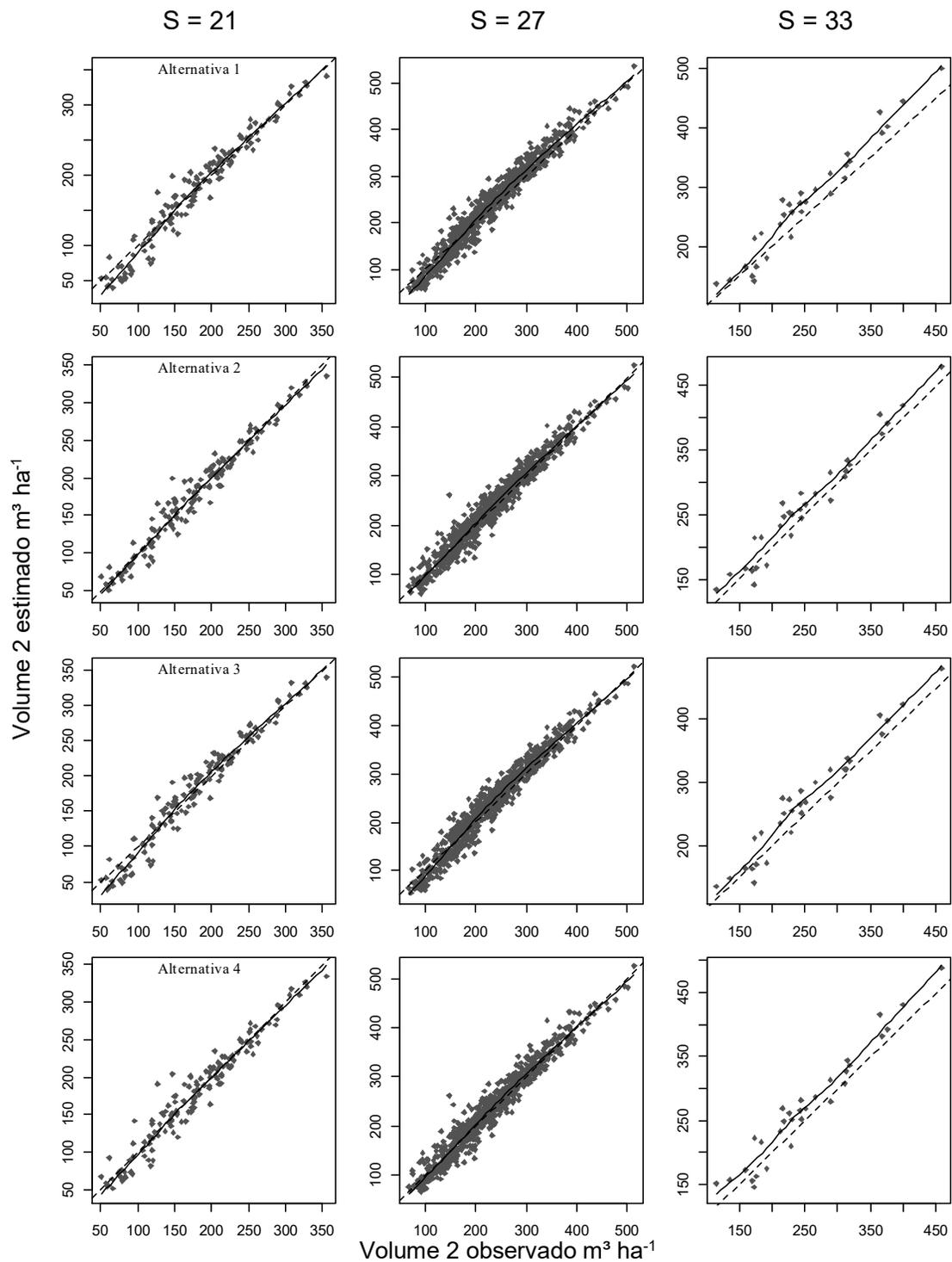


Figura 13 – Distribuição dos resíduos percentuais em função dos volumes totais estimados, em cada classe de produtividade para os modelos de distribuição de diâmetros.

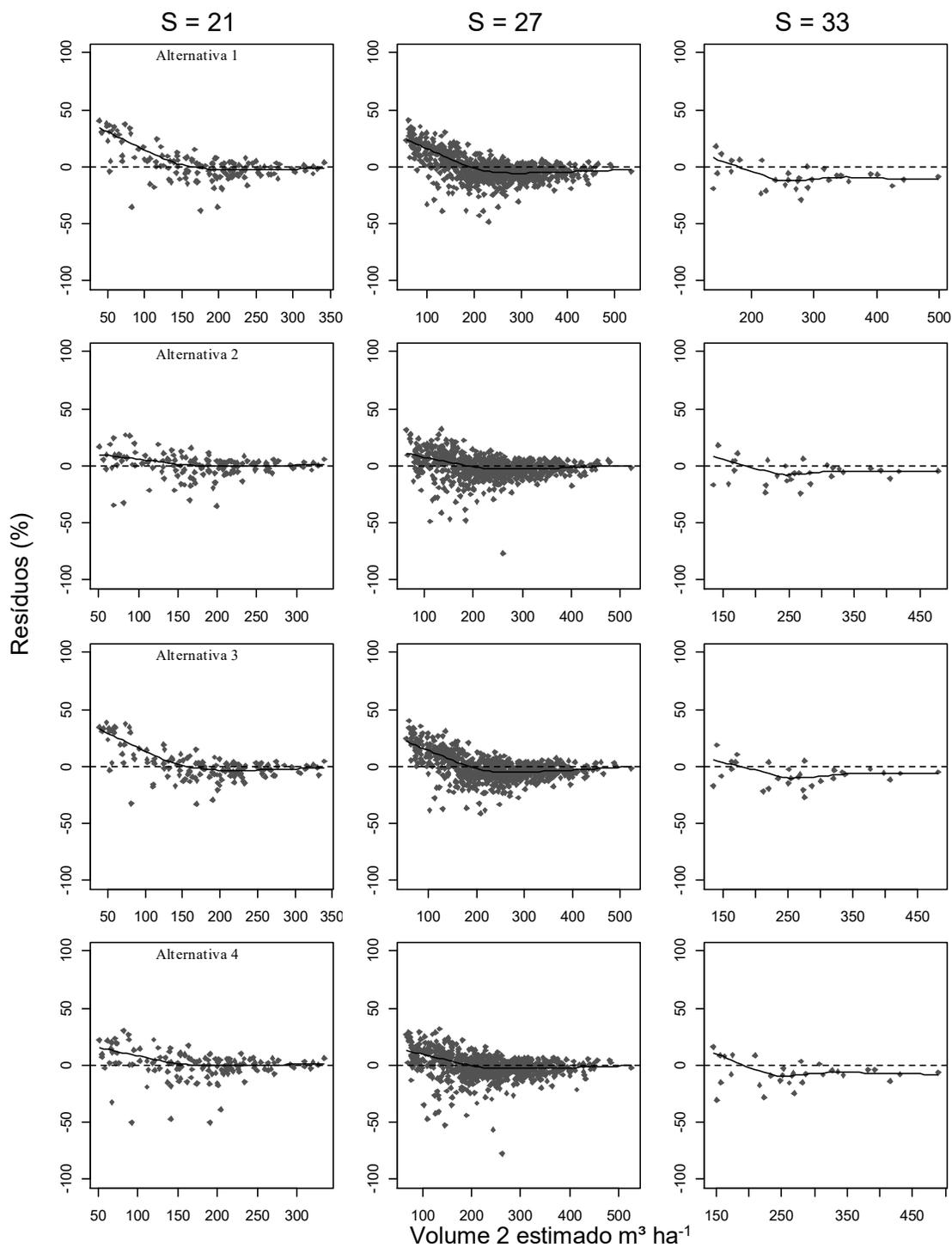


Figura 14 – Volumes totais por parcela estimados *versus* observados, obtidos pela aplicação das equações ajustadas em uma base de dados independente referentes às quatro alternativas de modelagem de distribuição de diâmetros avaliadas, em cada classe de produtividade.

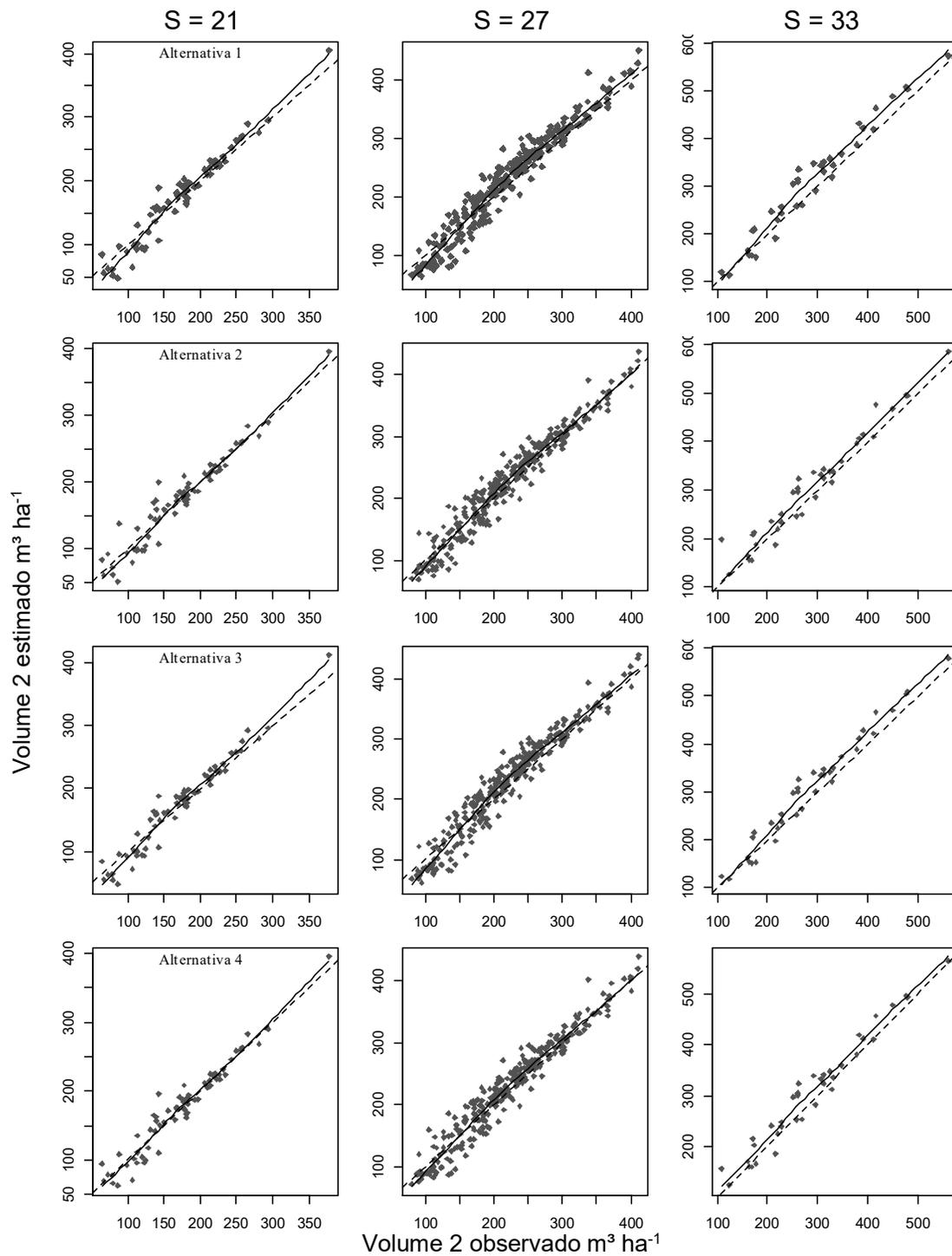
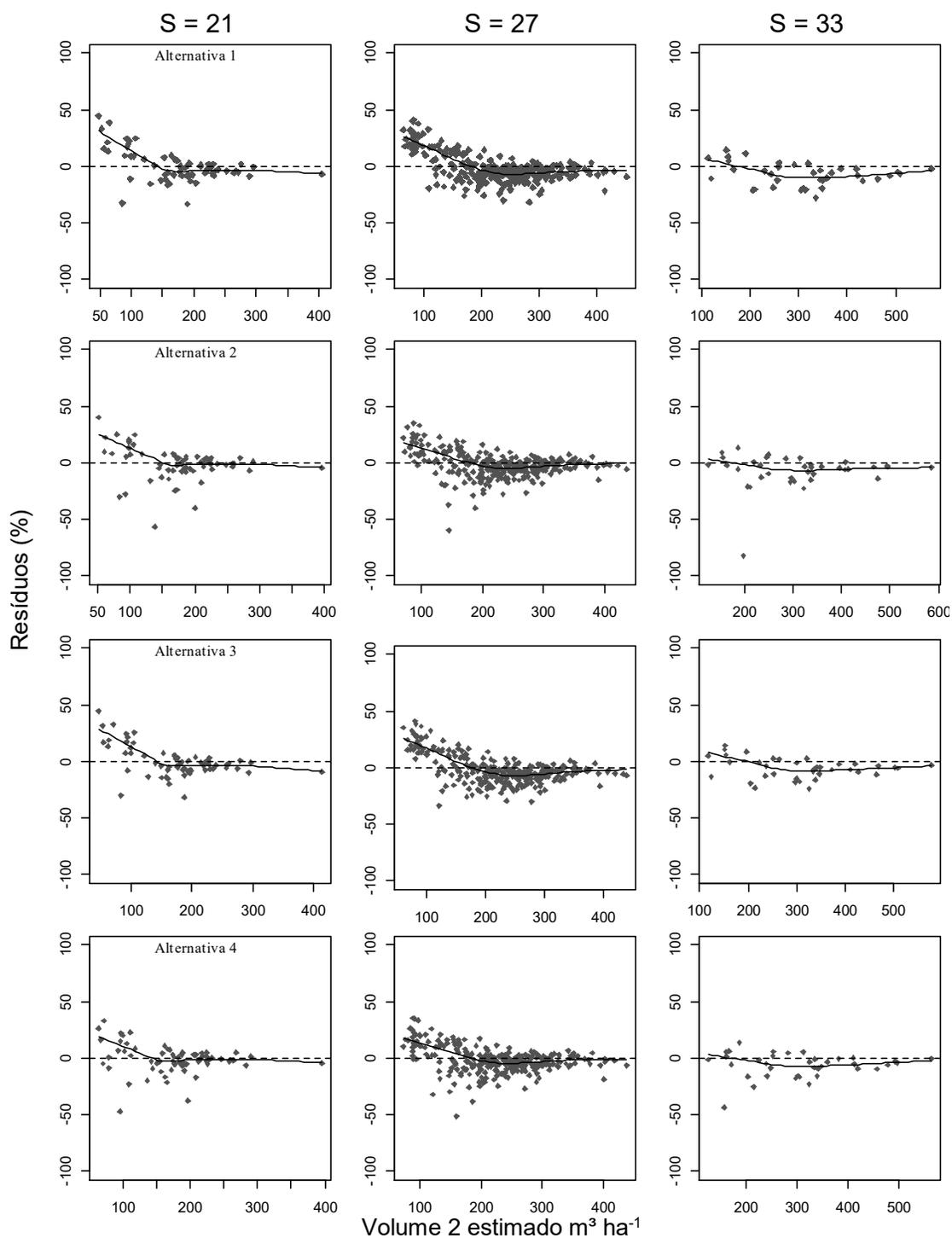


Figura 15 – Distribuição dos resíduos percentuais em função dos volumes totais estimados, pela aplicação das equações ajustadas em uma base de dados independente para cada classe de produtividade para os modelos de distribuição de diâmetros.



3.3. Comparação das alternativas ajustadas e validadas para os critérios definidos

Comparando-se as estatísticas para as alternativas usando modelagem de efeitos fixos (alternativa 1_0 e alternativa 1), verifica-se no ajuste que a alternativa 1_0 é mais acurada em todas as classes de produtividade. Apesar disso, apresenta-se mais viesada nas classes de produtividade Média (S = 27) e Alta (S = 33). Contudo, ao realizar a avaliação gráfica dos resíduos, verifica-se maior dispersão de pontos em torno da linha 0:1 na alternativa 1_0, principalmente na classe de produtividade Média (S = 27), que concentra 81% das parcelas, no ajuste e na validação. A dispersão residual é semelhante no ajuste, mas apresenta tendência de subestimativa de parcelas com maior volume na classe de produtividade Baixa (S = 21) para a alternativa 1_0. Tal tendência é verificada, apesar de ligeira, na mesma alternativa, para a classe de produtividade Média (S = 27). Na validação, a tendência de superestimar da alternativa 1_0 é mais pronunciada em todas as classes de produtividade. Diante dessa análise, é possível afirmar que a alternativa 1 é mais acurada e menos viesada, quando comparada com a alternativa 1_0.

Avaliando-se as alternativas mistas, tem-se a para o critério 1, a alternativa 1_1 e para o critério 2 as alternativas 2 a 4. A alternativa 2 apresenta maior acurácia e menor viés em todas as classes de produtividade no ajuste, contudo, na validação a alternativa 4 a supera nas classes de produtividade Baixa (S = 21) e Alta (S = 33), tanto em acurácia, quanto em viés, sendo este ligeiramente menor. Pela análise gráfica dos volumes observados *versus* estimados no ajuste, a superioridade da alternativa 2 é confirmada, e na validação, pode se considerá-la equivalente à alternativa 4.

Comparando-se as alternativas fixas e mistas, a alternativa 2 é mais acurada e menos viesada em todas as classes de produtividade, frente à alternativa 1, indicando então que os efeitos aleatórios produziram um melhor desempenho da modelagem para se estimar o volume por classe de diâmetro.

4. CONCLUSÕES

Para modelos fixos e mistos, as alternativas ajustadas e validadas com dados do critério 2 foram mais acuradas e menos viesadas que as alternativas ajustadas tendo a parcela com grupos de efeitos aleatórios (critério 1).

Com os resultados obtidos para as condições em que foi desenvolvido este estudo, é possível afirmar que a modelagem mista para modelos de distribuição de diâmetros apresenta-se mais acurada e menos viesada quando comparada com abordagens com modelos de efeitos fixos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ARAÚJO JÚNIOR, C. A.; NOGUEIRA, G. S.; OLIVEIRA, M. L. R. DE; et al. Projeção da distribuição diamétrica de povoamentos de eucalipto em diferentes amplitudes de classe. **Pesq. agropec. bras.**, Brasília, v. 45, n. 11, p. 1275–1281, 2010. SciELO Brasil.

ARIAS-RODIL, M.; CASTEDO-DORADO, F.; CÁMARA-OBREGÓN, A.; DIÉGUEZ-ARANDA, U. Fitting and calibrating a multilevel mixed-effects stem taper model for *Maritime Pine* in NW Spain. **PloS one**, v. 10, n. 12, p. e0143521, 2015. Public Library of Science. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143521>>. Acesso em: 31/12/2015.

BAILEY, R. L.; DELL, T. R. Quantifying diameter distributions with the Weibull function. **Forest Science**, v. 19, n. 2, p. 97–104, 1973. Society of American Foresters.

BINOTI, D. H. B.; BINOTI, M. L. M. DA S.; LEITE, H. G. Modelagem da distribuição diamétrica de povoamentos equiâneos de eucalipto utilizando a função logística generalizada. **Revista Árvore**, v. 39, n. 4, p. 707–711, 2015. Sociedade de Investigações Florestais. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622015000400707&lng=pt&nrm=iso&tlng=en>. Acesso em: 7/1/2016.

BINOTI, D. H. B.; BINOTI, M. L. M. DA S.; LEITE, H. G.; SILVA, A. A. L. DA; ALBUQUERQUE, A. C. Modelagem da distribuição de diâmetros utilizando automatos celulares e redes neurais artificiais. **CERNE**, v. 19, n. 4, p. 677–685, 2013.

BINOTI, D. H. B.; LEITE, H. G.; NOGUEIRA, G. S.; et al. Uso da função Weibull de três parâmetros em um modelo de distribuição diamétrica para plantios de eucalipto submetidos a desbaste. **Revista Árvore**, v. 34, n. 1, p. 147–156, 2010. SciELO Brasil.

BINOTI, M. L. M. DA S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, v. 37, n. 4, p. 639–645, 2013. Sociedade de Investigações Florestais. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000400007&lng=pt&nrm=iso&tlng=en>. Acesso em: 31/12/2015.

BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **CERNE**, v. 21, n. 1, p. 97–105, 2015. UFLA - Universidade Federal de Lavras. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-77602015000100097&lng=en&nrm=iso&tlng=pt>. Acesso em: 31/12/2015.

BURKHART, H.; TOMÉ, M. Diameter-Distribution Models for Even-Aged Stands. **Modeling Forest Trees and Stands SE - 12**. p.261–297, 2012. Springer Netherlands.

CALAMA, R.; MONTERO, G. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. **Silva Fennica**, v. 39, n. 1, p. 37–54, 2005.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração Florestal: Perguntas e Respostas**. 4th ed. Viçosa, MG, 2013.

CARVALHO, S. DE P. C. E; CALEGARIO, N.; SILVA, F. F. E; et al. Modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* clonal. **CERNE**, v. 17, n. 4, p. 541–548, 2011. Universidade Federal de Lavras.

CASTRO, R. V. O.; SOARES, C. P. B.; MARTINS, F. B.; LEITE, H. G. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos. **Pesquisa Agropecuária Brasileira**, v. 48, n. 3, p. 287–295, 2013. SciELO Brasil.

CRECENTE-CAMPO, F.; SOARES, P.; TOMÉ, M.; DIÉGUEZ-ARANDA, U. Modelling annual individual-tree growth and mortality of Scots pine with data obtained at irregular measurement intervals and containing missing observations. **Forest Ecology and Management**, v. 260, n. 11, p. 1965–1974, 2010. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112710005153>>. Acesso em: 31/12/2015.

DIAMANTOPOULOU, M. J.; ÖZÇELİK, R.; CRECENTE-CAMPO, F.; ELER, Ü. Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. **Biosystems Engineering**, v. 133, p. 33–45, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1537511015000367>>. Acesso em: 7/1/2016.

ELZHOV, T. V.; MULLEN, K. M.; SPIESS, A.-N.; BOLKER, B. minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm. Disponível em: <<http://cran.r-project.org/package=minpack.lm>>. Acesso em: 31/12/2015.

FRAGA FILHO, C. V.; SCHMIDT, L. N. Weibull 2 e 3p Maximum Likelihood. Disponível em: <<https://sourceforge.net/projects/weibull23p/>>. Acesso em: 20/11/2014.

FRAGA FILHO, C. V.; SIMIQUELI, A. P. .; ALTOE, W. . Fgmutils: Forest Growth Model Utilities. Disponível em: <<https://cran.r-project.org/web/packages/Fgmutils/index.html>>. Acesso em: 31/12/2015.

GÖRGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. **Computers and Electronics in Agriculture**, v. 116, p. 221–227, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169915001994>>. Acesso em: 9/12/2015.

GUIMARÃES, D. P. **Desenvolvimento de um modelo de distribuição diamétrica de passo invariante para prognose e projeção da estrutura de povoamentos de eucalipto**, 1994. 160 f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 1994.

ISLAM, M. N.; KURTTILA, M.; MEHTÄTALO, L.; HAARA, A. Analyzing the effects of inventory errors on holding-level forest plans: the case of measurement error in the basal area of the dominated tree species. **Silva Fennica**, v. 43, n. 1, p. 71–85, 2009. Finnish Society of Forest Science.

LEITE, H. G. **Ajuste de um modelo de estimação de frequência e produção por classe de diâmetro, para povoamentos de Eucalyptus saligna Smith**, 1990. 97f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 1990.

MALTAMO, M.; GOBAKKEN, T. Predicting Tree Diameter Distributions. In: M. Maltamo; E. Næsset; J. Vauhkonen (Eds.); **Forestry Applications of Airborne Laser Scanning**, Managing Forest Ecosystems. v. 27, p.177–191, 2014. Dordrecht: Springer Netherlands. Disponível em: <<http://link.springer.com/10.1007/978-94-017-8663-8>>. Acesso em: 8/1/2016.

MARTINS, F. B.; SOARES, C. P. B.; SILVA, G. F. DA. Individual tree growth models for eucalyptus in northern Brazil. **Scientia Agricola**, v. 71, n. 3, p. 212–225, 2014. SciELO Brasil.

MASSEY JR, F. J. The Kolmogorov-Smirnov test for goodness of fit. **Journal of the American statistical Association**, v. 46, n. 253, p. 68–78, 1951. Taylor & Francis Group.

MENDONÇA, A. R. DE E. **Modelagem não linear do crescimento e da produção de plantações florestais**, 2010. 84f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras, 2010.

MÔRA, R. **Funções de afilamento de forma variável e modelagem de efeitos mistos em fustes de *Pinus taeda* e *Eucalyptus saligna***, 2015. 277f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2015.

NORD-LARSEN, T.; CAO, Q. V. A diameter distribution model for even-aged beech in Denmark. **Forest Ecology and Management**, v. 231, n. 1-3, p. 218–225, 2006. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112706003616>>. Acesso em: 7/1/2016.

PIETILÄ, I.; KANGAS, A.; MÄKINEN, A.; MEHTÄTALO, L. Influence of growth prediction errors on the expected losses from forest decisions. **Silva Fennica**, v. 44, n. 5, p. 829–843, 2010.

PINHEIRO, J.; BATES, D.; DEBROY, S.; SARKAR, D. nlme: Linear and Nonlinear Mixed Effects Models. Disponível em: <<http://cran.r-project.org/package=nlme>>. Acesso em: 31/12/2015.

UZOH, F. C. C. C.; OLIVER, W. W. Individual tree diameter increment model for managed even-aged stands of ponderosa pine throughout the western United States using a multilevel linear mixed effects model. **Forest ecology and management**, v. 256, n. 3, p. 438–445, 2008. Elsevier. Disponível em: <<http://www.treesearch.fs.fed.us/pubs/40984#.VoV-xYIcjD4.mendeley>>. Acesso em: 31/12/2015.

VISMARA, E. DE S.; MEHTÄTALO, L.; BATISTA, J. L. F. Linear mixed-effects models and calibration applied to volume models in two rotations of *Eucalyptus grandis* plantations. **Canadian Journal of Forest Research**, p. 132–141, 2015. NRC Research Press. Disponível em: <<http://www.nrcresearchpress.com/doi/abs/10.1139/cjfr-2014-0435#.VoWA8-JYWug>>. Acesso em: 31/12/2015.

XU, Q.; HOU, Z.; MALTAMO, M.; TOKOLA, T. Calibration of area based diameter distribution with individual tree based diameter estimates using airborne laser scanning. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 93, p. 65–75, 2014. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0924271614000689>>. Acesso em: 7/1/2016.

ZHAO, L.; LI, C.; TANG, S. Individual-tree diameter growth model for fir plantations based on multi-level linear mixed effects models across southeast China. **Journal of Forest Research**, v. 18, n. 4, p. 305–315, 2013.

6. APÊNDICES

Apêndice A – Volumes estimados e observados obtidos no ajuste e na validação das alternativas 1 (não mista) e 2 (mista) para o conjunto de dados do critério 2.

Figura 16 – Volumes por classe de diâmetro observados e estimados aos 36, 48 e 60, 72, 84 e 96 meses para uma parcela selecionada aleatoriamente com os dados de ajuste da alternativa 1 (critério 2).

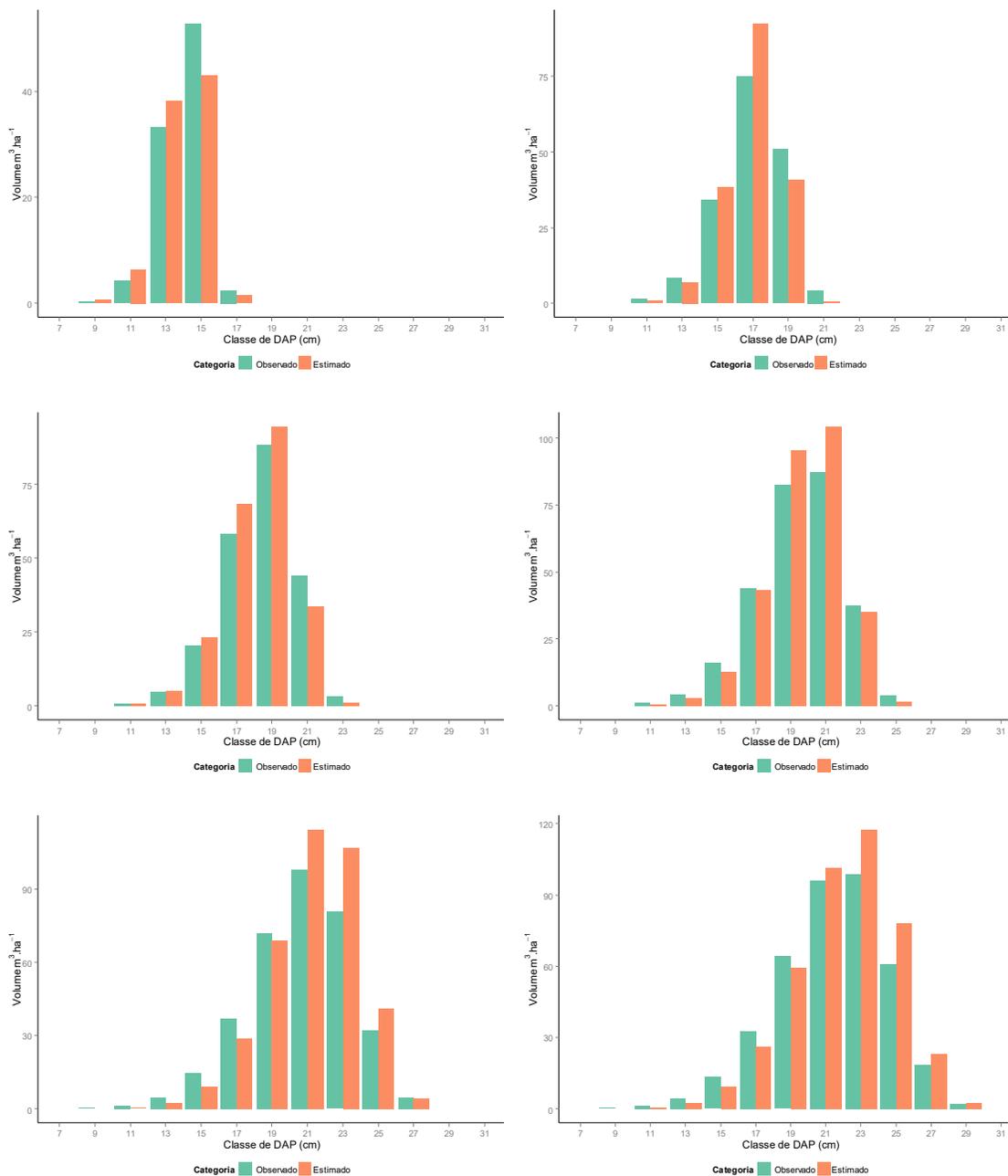


Figura 17 – Volumes por classe de diâmetro observados e estimados aos 36, 48 e 60, 72, 84 e 96 meses para outra parcela, selecionada aleatoriamente do grupo de validação (critério 2).

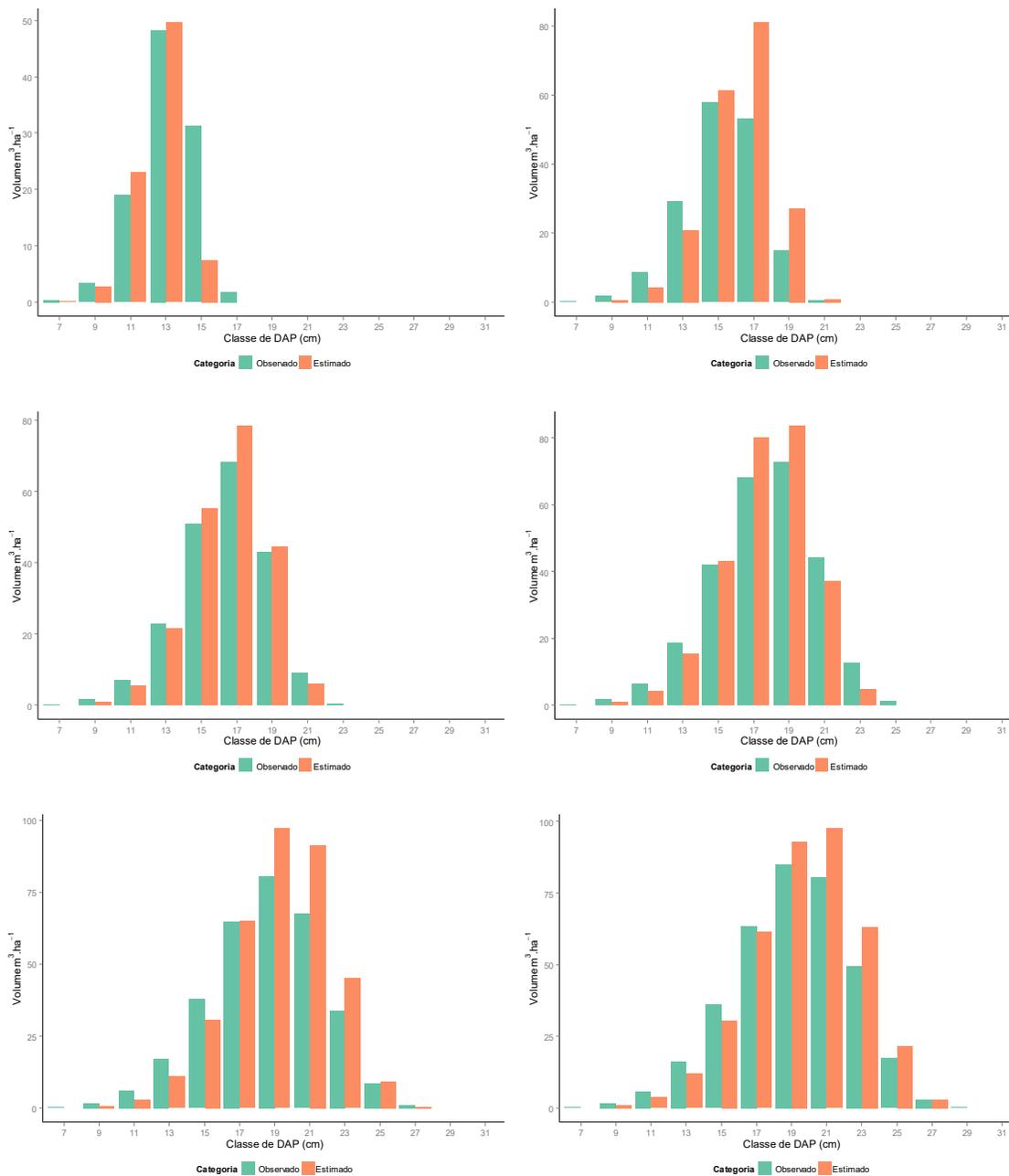


Figura 18 – Volumes por classe de diâmetro observados e estimados aos 36, 48 e 60, 72, 84 e 96 meses para uma parcela selecionada aleatoriamente com os dados de ajuste da alternativa 2 (critério 2).

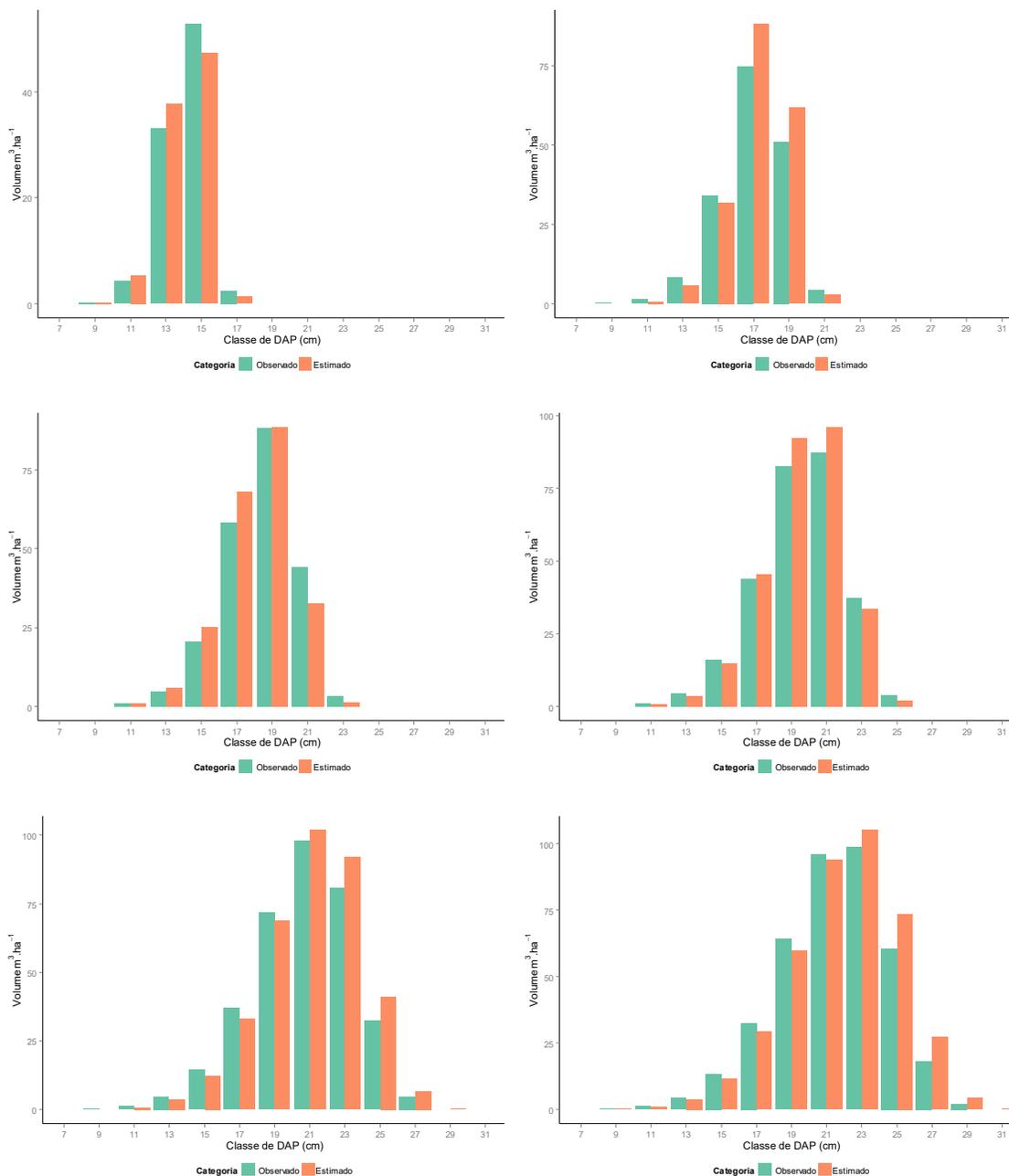
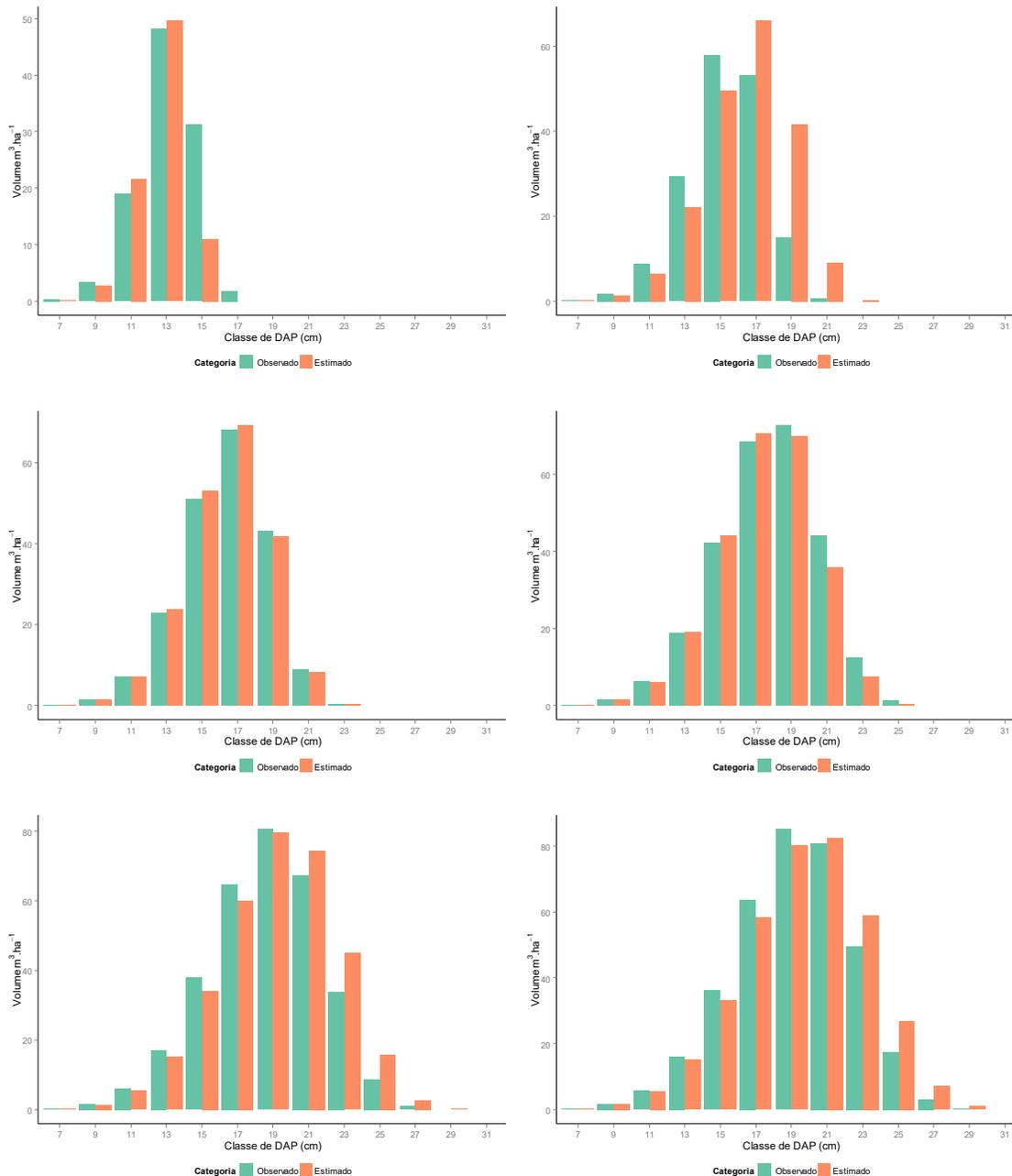


Figura 19 – Volumes por classe de diâmetro observados e estimados aos 36, 48 e 60, 72, 84 e 96 meses para outra parcela, selecionada aleatoriamente do grupo de validação (critério 2).



CAPÍTULO 3 – MODELAGEM DO CRESCIMENTO E PRODUÇÃO DE POVOAMENTOS DE EUCALIPTO, EM NÍVEL DE ÁRVORES INDIVIDUAIS, COM E SEM EFEITO ALEATÓRIO

1. INTRODUÇÃO

Modelos de crescimento e produção são ferramentas essenciais para o planejamento da gestão florestal, estimativas acuradas dos estoques atuais e suas alterações são imprescindíveis para implementação de alternativas de gestão para a tomada de decisões (UZOH; OLIVER, 2008). Castro et al. (2013) afirmam que embora modelos no nível de povoamento atendam de maneira eficiente aos usuários, sendo a categoria mais utilizada no Brasil desde meados da década de 80, muitas vezes o detalhamento das informações acerca do crescimento e produção dos povoamentos é necessário, principalmente para estimar a produção de madeira de alta qualidade e multiprodutos.

Os modelos de crescimento e produção em nível de árvore individual simulam o crescimento de cada árvore em diâmetro, altura e copa, decidindo se a árvore vive ou morre, calculando o crescimento e o volume, de forma individual (MARTINS et al., 2014). A partir dos volumes individuais calculados, obtêm-se os volumes totais por parcela, por meio da soma dos volumes das árvores contidas nestas parcelas. São três os submodelos componentes deste nível de modelo: probabilidade de mortalidade, ingresso e crescimento dimensional (altura e diâmetro).

A mortalidade pode ser classificada em dois tipos: regular e irregular, sendo a primeira ocasionada por fatores não catastróficos como competição, fatores genéticos e senescência, a qual pode ser mensurada. Já a mortalidade irregular ocorre por causas não intrínsecas à árvore, como incêndios, ataques de pragas, danos mecânicos, portanto, difíceis de serem previstos e mensurados. Sendo assim, a mortalidade é um aspecto essencial, no entanto, de difícil avaliação e modelagem (CASTRO et al., 2015). O ingresso representa outro submodelo, e diz respeito ao número de indivíduos que medidos em uma idade i não foram medidos em idades anteriores, por não apresentarem tamanho mínimo de acordo com o determinado.

O uso de modelos de árvores individuais no Brasil ainda é restrito, seja pela indisponibilidade de dados apropriados para seu ajuste, sua complexidade, o maior número de funções de o compõem e a falta de estudos que demonstrem seus

resultados contribuem a baixa utilização (CASTRO et al., 2013, 2014; MARTINS et al., 2014).

O avanço dos recursos de computação, com maior poder de processamento resultado no uso de técnicas estatísticas alternativas, dentre as quais destaca-se os modelos de regressão não lineares generalizados, generalizados de efeito aleatório e de efeito aleatório (CARVALHO et al., 2011). Os mesmos autores citam trabalhos que envolvem a aplicação de tais modelos, como Budhathoki et al. (2008), Calegario et al. (2005a, b), Fang e Bailey (2001), Gregoire et al. (1995) e Pilar-Díaz e Couto (1999). Outras abordagens como o uso de redes neurais artificiais (CASTRO et al., 2013; BINOTI et al., 2013, 2015), aprendizado de máquina e técnicas de sensoriamento remoto tem ganhado espaço (GÖRGENS et al., 2015), dentre outros.

O uso de modelos mistos permite identificar as diferentes fontes de variabilidade estocásticas que não são explicadas pela parte fixa do modelo dividindo-se a variância residual em diferentes componentes (CALAMA; MONTERO, 2005). Neste tipo de modelo, as medidas dos indivíduos não precisam ser igualmente espaçadas e balanceadas e as análises podem ser conduzidas com os dados de indivíduos que tiveram medidas perdidas (árvores que não foram medidas dentro de parcelas, por exemplo) ou que apresentam ausência de informação em algum momento do estudo. Portanto, na ausência de dados, os modelos mistos oferecem mais uma vantagem de utilização frente aos outros tipos de modelos para dados longitudinais e de medidas repetidas (UZOH; OLIVER, 2008).

Na modelagem de efeitos aleatórios, os parâmetros de um modelo ajustado podem variar entre os grupos, sendo assim, o modelo pode ser calibrado para um determinado grupo afim de melhorar a qualidade das predições dentro do referido grupo (VISMARA et al., 2015).

O objetivo deste capítulo foi avaliar e comparar duas estratégias para a estimativa de volume pela projeção do diâmetro a 1,3m de altura (DAP) e altura total (HT) de modelos em nível de árvore individual, com e sem efeitos aleatórios.

2. MATERIAL E MÉTODOS

2.1. Descrição dos dados

Os dados utilizados neste capítulo são descritos na seção 3.4 da tese. A base de dados com as variáveis das árvores individuais, pareadas na idade 1 e idade 2 foi criada com a função *createPairedData* do pacote *Fgmutils* (FRAGA FILHO et al., 2015), criado durante o desenvolvimento deste trabalho e disponibilizado nos repositórios da linguagem R.

2.2. Divisão da base de dados para ajuste e validação dos modelos

Com o objetivo de avaliar diferentes grupos na calibração de modelos mistos, os dados foram divididos utilizando-se dois critérios: no primeiro (critério 1), o objetivo foi avaliar o comportamento dos modelos para idades futuras, definindo a variável parcela como grupo. No segundo critério (critério 2), foram definidos grupos formados por variáveis relacionadas ao povoamento, disponíveis na base de dados do inventário.

De acordo com o critério 1, os dados foram divididos em dois conjuntos: o primeiro para ajuste, e o segundo para validação dos modelos. Do conjunto de ajuste, participaram todas as parcelas com idade futura (idade 2) até 60 meses, e no segundo, de projeção (ou validação), foram selecionadas as mesmas parcelas do grupo de ajuste, porém com idade futura igual ou superior a 72 meses.

De acordo com o critério 2, o conjunto de parcelas foi dividido aleatoriamente também em dois grupos, sendo 70% das parcelas utilizadas para ajuste dos modelos e 30% para validação, ou seja, de acordo com este critério, as parcelas do conjunto de ajuste não estavam presentes no conjunto de validação.

2.3. Alternativas de modelagem em nível de árvores individuais

Para realizar a projeção do crescimento do diâmetro a 1,3 m de altura (DAP) e da altura total (HT) das árvores, foram avaliadas duas estratégias, sendo a primeira formada pelo modelo que apresentou melhor desempenho dentre cinco modelos para estimativa de crescimento em diâmetro (cm) e altura (m) apresentados na Tabela 1 (MARTINS et al., 2014), denominada a partir de agora de estratégia 1.

Tabela 1 – Submodelos utilizados para estimar o crescimento em diâmetro e altura.

Tipo/Autor	Modelo
Pienaar e Schiver (1981)	$Y_2 = Y_1 \cdot e^{(-\beta_0(l_2^{\beta_1} - l_1^{\beta_1}))} + \varepsilon \quad (1)$
Logístico / Amaro et al. (1998)	$Y_2 = Y_1 + \left(\frac{\beta_0}{1 + e^{(\beta_1 - \beta_2 l_2)}} - \frac{\beta_0}{1 + e^{(\beta_1 - \beta_2 l_1)}} \right) + \varepsilon \quad (2)$
Richards, Zeide (1993)	$Y_2 = Y_1 + \left(\frac{\beta_0}{1 + e^{(\beta_1 - \beta_2 l_2) / \beta_3}} - \frac{\beta_0}{1 + e^{(\beta_1 - \beta_2 l_1) / \beta_3}} \right) + \varepsilon \quad (3)$
Schumacher / Tomé (2003)	$Y_2 = Y_1 + e^{(\beta_0 - \beta_1 l_2)} - e^{(\beta_0 - \beta_1 l_1)} + \varepsilon \quad (4)$
Linear adaptado de Bella (1971) e Campos e Leite (2009)	$Y_2 = Y_1 + \left(\beta_0 + \beta_1 \left(\frac{1}{l_2} - \frac{1}{l_1} \right) + \beta_2 BAI + \beta_3 S \right) + \varepsilon \quad (5)$

em que: Y_2 é o diâmetro (em cm) ou a altura total (m) em idade futura; Y_1 é o diâmetro (cm) ou altura (m) em idade atual; l_2 é a idade futura (meses); l_1 é a idade atual (meses); S é o índice de local; BAI é o índice de área basal, com equação $BAI = d_i^2 / q^2$; q é o diâmetro quadrático (cm); β_i são os coeficientes do modelo; e é o erro aleatório;

A segunda estratégia avaliada, denominada aqui de estratégia 2, é composta pelos modelos lineares propostos neste trabalho, ajustados posteriormente com efeitos aleatórios. O modelo de crescimento em diâmetro foi assim definido:

$$\widehat{DAP}_2 = \beta_0 + \beta_1 DAP_1 + \beta_2 l_1 + \beta_3 l_2 + \beta_4 \bar{D}_1 + \beta_5 DMIN_1 + \beta_6 DMAX_1 + \varepsilon \quad (6)$$

em que: \widehat{DAP}_2 = diâmetro (em cm) estimado na uma idade futura; DAP_1 = diâmetro (cm) em idade atual; \bar{D}_1 = média dos diâmetros (cm) da parcela na idade atual; $DMAX$ = diâmetro máximo (cm) da parcela na idade atual; $DMIN$ = diâmetro mínimo (cm) da parcela na idade atual; l_1 = idade atual (meses); l_2 = idade futura (meses); β_i = coeficientes do modelo; e ε = erro aleatório;

Como parte da estratégia 2, o modelo de crescimento em altura total foi assim definido:

$$\widehat{HT}_2 = \beta_0 + \beta_1 HT_1 + \beta_2 I_1 + \beta_3 IID3_1 + \beta_4 I_2 + \varepsilon \quad (7)$$

em que: \widehat{HT}_2 = altura total (em m) estimada na uma idade futura; HT_1 = altura total (m) na idade atual; $IID3 = \frac{DAP^2}{D^2} \cdot \frac{HT}{H}$ = índice de competição independente de distância (MARTINS, 2011); I_1 = idade atual (meses); I_2 = idade futura (meses); β_i = coeficientes do modelo; e ε = erro aleatório;

Com base no modelo escolhido na estratégia 1, foram definidas alternativas de modelagem mistas, avaliadas neste capítulo.

Depois de concluído o ajuste dos modelos apresentados (Tabela 1), procedeu-se com a seleção de um modelo de crescimento em diâmetro a 1,3 m de altura (cm) e um para altura total (m), que fez parte dos conjuntos de alternativas, apresentadas adiante.

Tendo como referência as estratégias apresentadas, foram propostos dois conjuntos de alternativas de modelagem com e sem efeitos aleatórios, com o propósito de avaliar seu desempenho quando à acurácia e tendenciosidade, quando comparadas aos resultados das estatísticas do ajuste e análise gráfica dos resíduos dos modelos que compõem cada estratégia.

O primeiro conjunto é formado por alternativas que não utilizaram o identificador de cada árvore ou parcela como grupo de efeitos aleatórios na calibração, tendo definidas outras variáveis com este propósito. Como opções ao uso do identificador da árvore, apresentam-se a classe diamétrica da árvore, a classe de altura, e o material genético, sendo estas variáveis candidatas a formarem grupos de efeitos aleatórios. A variável material genético foi considerada como variável de agrupamento de efeitos aleatórios (Tabela 2), e outras três variáveis foram utilizadas como grupos para calibração das novas alternativas de modelagem mista.

Tabela 2 – Variáveis criadas para representar grupos de efeitos aleatórios no ajuste das alternativas propostas utilizando com modelos mistos.

Grupo	Variável (is) agrupada (s)	Amplitude
Classe de Altura	Altura (m)	2 m
Classe de Altura (2)	Altura (m)	3 m
Classe diamétrica	Diâmetro (cm)	2 cm

Classe Dimensional	Diâmetro (cm) e Altura (m)	2 cm e 2 m
Material Genético	Material Genético	-

A variável *Classe Dimensional* é formada pela combinação (concatenação) das variáveis classe de altura e classe diamétrica, e fornece uma classificação simultânea (de dupla entrada) para cada árvore quanto à classe de altura total e diâmetro nas quais a árvore foi classificada.

Considerando os novos grupos de efeitos aleatórios apresentados na Tabela 2, alternativas de modelagem de efeitos aleatórios (conjunto de alternativas 1) foram propostas, tendo como referência o modelo de melhor desempenho selecionado com base na Tabela 1. Este conjunto de alternativas foi ajustado e validado com o conjunto de dados do critério 1.

O segundo conjunto refere-se a alternativas de modelagem apresentadas na Tabela 3, que foram propostas para realização da calibração em nível de árvore individual, utilizando a variável identificadora da árvore como grupo de efeitos aleatórios. Embora a modelagem de árvore individual esteja sendo avaliada, a variável parcela também foi avaliada como grupo de efeitos aleatórios, já que representa um conjunto de árvores individuais com características similares, provenientes do mesmo local. Tais alternativas foram ajustadas e validadas com o conjunto de dados estabelecido de acordo com o critério 2.

Tabela 3 – Alternativas de modelagem avaliadas para projeção do volume em nível de povoamento utilizando a parcela como grupo de efeitos aleatórios (critério 2).

Alternativa	Modelo
2_0	<p>Modelos lineares fixos, para projeção de DAP e HT</p> $\widehat{DAP}_2 = \beta_0 + \beta_1 DAP_1 + \beta_2 I_1 + \beta_3 I_2 + \beta_4 \bar{D}_1 + \beta_5 DMIN_1 + \beta_6 DMAX_1 + \varepsilon$ $\widehat{HT}_2 = \beta_0 + \beta_1 HT_1 + \beta_2 I_1 + \beta_3 IID3_1 + \beta_4 I_2 + \varepsilon$
2_1	$\widehat{DAP}_2 = (\beta_0 + b_0) + \beta_1 DAP_1 + \beta_2 I_1 + \beta_3 I_2 + (\beta_4 + b_3) \bar{D}_1 + \beta_5 DMIN_1 + \beta_6 DMAX_1 + \varepsilon$ <p>, com o identificador da árvore como variável de agrupamento de efeitos aleatórios, e diâmetro médio atual como inclinação aleatória.</p> $\widehat{HT}_2 = (\beta_0 + b_0) + \beta_1 HT_1 + \beta_2 I_1 + \beta_3 IID3_1 + (\beta_4 + b_4) I_2 + \varepsilon$ <p>, com o identificador da árvore como variável de agrupamento de efeitos aleatórios, e idade futura como inclinação aleatória.</p>

2_2_0	Alternativas para projeção de DAP e HT fixas do modelo selecionado da Tabela 1, cuja equação apresentou melhor desempenho nas estatísticas avaliadas e na análise gráfica.
2_2	$\widehat{DAP}_2 = (\beta_0 + b_0) + \beta_1 DAP_1 + \beta_2 I_1 + \beta_3 I_2 + (\beta_4 + b_3) \bar{D}_1 + \beta_5 DMIN_1 + \beta_6 DMAX_1 + \varepsilon$ <p>, com a parcela como variável de agrupamento de efeitos aleatórios, e diâmetro médio atual como inclinação aleatória.</p>
2_3	$\widehat{HT}_2 = (\beta_0 + b_0) + \beta_1 HT_1 + \beta_2 I_1 + (\beta_3 + b_3) IID3_1 + \beta_4 I_2 + \varepsilon$ <p>, com a parcela como variável de agrupamento de efeitos aleatórios, e IID3 na idade atual como inclinação aleatória.</p>
	Alternativas para projeção de DAP e HT mistas do modelo selecionado da Tabela 1, cuja equação apresentou melhor desempenho nas estatísticas avaliadas e na análise gráfica, com o identificador da árvore como variável de agrupamento de efeitos aleatórios.

em que: S = índice de local; B_2 área basal projetada para a idade futura; β - parâmetros fixos do modelo, b parâmetros aleatórios do modelo misto.

De posse do diâmetro a 1,3 m de altura (DAP) e altura total estimados pelas alternativas de modelagem, calculou-se os volumes individuais das árvores dentro de cada parcela com a equação de volume apresentada na seção 3.4 da tese. Então, procedeu-se com a soma dos volumes individuais para obtenção dos volumes da parcela e conseqüentemente o volume total dos povoamentos ($m^3 \text{ ha}^{-1}$), em cada idade.

2.4. Avaliação dos resultados

A avaliação das alternativas foi realizada pelas seguintes estatísticas:

a) Coeficiente de correlação (PIETILĂ et al., 2010; MÔRA, 2015):

$$r_{\hat{y}y} = \frac{n^{-1}(\sum(\hat{y}_i - \hat{y}_m)(y_i - \bar{y}))}{\sqrt{(n^{-1}\sum(\hat{y}_i - \hat{y}_m)^2)(n^{-1}\sum(y_i - \bar{y})^2)}} \quad (8)$$

b) BIAS(%) (UZOH; OLIVER, 2008):

$$BIAS = \frac{\sum_{i=1}^n (\hat{Y} - Y)}{n} \quad (9)$$

$$BIAS(\%) = \frac{BIAS}{\sum_{i=1}^n Y / n} \cdot 100 \quad (10)$$

c) Raiz quadrada do erro quadrático médio (RMSE %) (*Root Mean Square Error*), (ISLAM et al., 2009; ZHAO et al., 2013; ARIAS-RODIL et al., 2015):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n}} \quad (11)$$

$$RMSE(\%) = \frac{RMSE}{\sum_{i=1}^n Y / n} \cdot 100 \quad (12)$$

De forma a complementar a análise foram elaborados gráficos relacionando os volumes observados com os volumes estimados e gráficos de resíduos percentuais, obtidos pela expressão (MARTINS et al., 2014):

$$res(\%) = \left(\frac{y_i - \hat{y}_i}{y_i} \right) \cdot 100 \quad (13)$$

3. RESULTADOS E DISCUSSÃO

Nesta seção os resultados são apresentados de acordo com os critérios empregados na divisão dos dados em grupos. Antes, porém, foi selecionado o modelo da estratégia 1, cujo resultado é apresentado na seção 3.1, a seguir.

3.1. Avaliação dos modelos de crescimento em diâmetro e altura total

Observa-se na Tabela 4 que as equações dos modelos ajustados para diâmetro, apresentam valores próximos entre si para as estatísticas avaliadas ($r_{y\hat{y}}$, RMSE% e BIAS%), exceto a equação referente ao modelo de Schumacher e Tomé (2003) com $r_{y\hat{y}}$ inferior aos demais, apresentando um valor de 0,9430, RMSE (%) e 6,8060, superior aos modelos apresentados, portanto menos acurado, embora com viés semelhante. Comparando com os demais modelos, destaca-se a equação referente ao modelo Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014) com valor de 0,9807 de $r_{y\hat{y}}$, maior dentre os demais, mais acurado com valores inferiores de RMSE (%), e com viés significativamente menor, com BIAS (%) de 3,6060 e 0,0035, respectivamente.

Tabela 4 – Estatísticas utilizadas para avaliar o ajuste dos modelos não lineares de crescimento em diâmetro na altura do peito (DAP).

Modelo	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
Linear adaptado de Bella (1971) e Campos e Leite (2009), (MARTINS et al., 2014)	0,9807	3,61	0,0035
Pienaar e Schiver (1981)	0,9748	4,12	-0,2911
Logístico / Amaro et al. (1998)	0,9750	4,30	-0,4033
Richards, Zeide (1993)	0,9749	4,32	-0,4260
Schumacher / Tomé (2003)	0,9430	6,81	-0,2839

Para o crescimento em diâmetro, nota-se uma leve tendência de superestimar árvores com menores diâmetros nos modelos 1, 2, 3, e mais acentuada no modelo 4 (Figura 1). Martins et al. (2014) justificam que são tendências comuns de ocorrerem, contudo difíceis de serem explicadas. O modelo Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014) apresentou melhor distribuição residual, com a nuvem de pontos menos dispersa em relação a linha 0:1, corroborando

os resultados das estatísticas apresentadas para as equações dos modelos de crescimento em diâmetro.

Avaliando-se os gráficos de resíduos percentuais na Figura 2, nota-se que os modelos apresentaram distribuição dos resíduos similar, com desvantagem para o modelo 4 e uma ligeira vantagem para o modelo 5.

Figura 1 – Diâmetros na altura do peito estimados *versus* observados pelas equações ajustadas para os cinco modelos de crescimento em diâmetro.

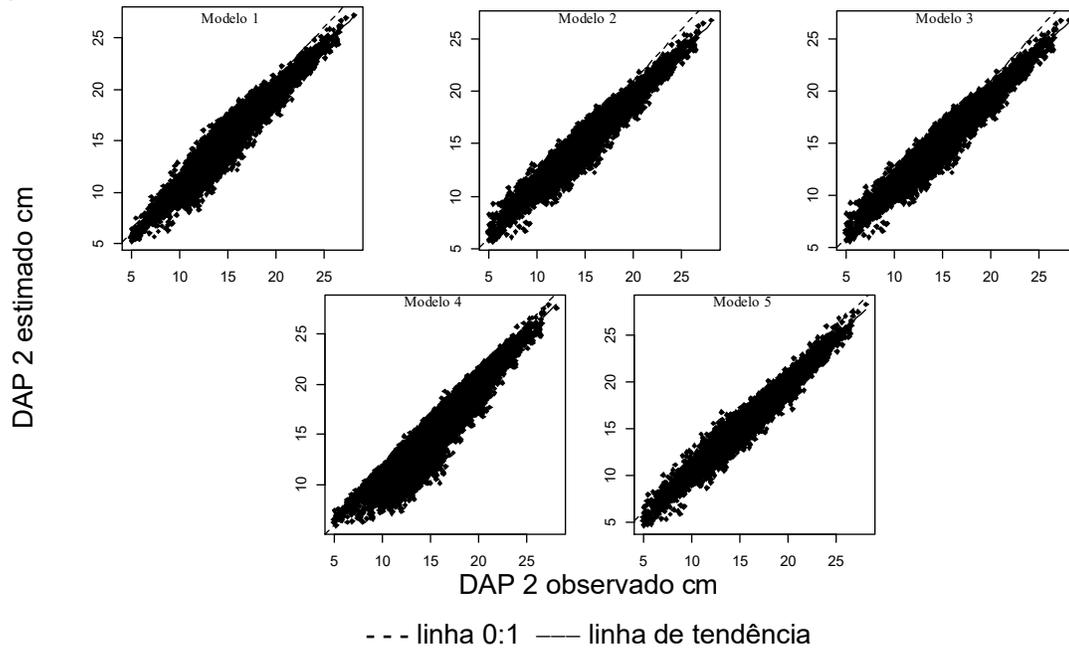
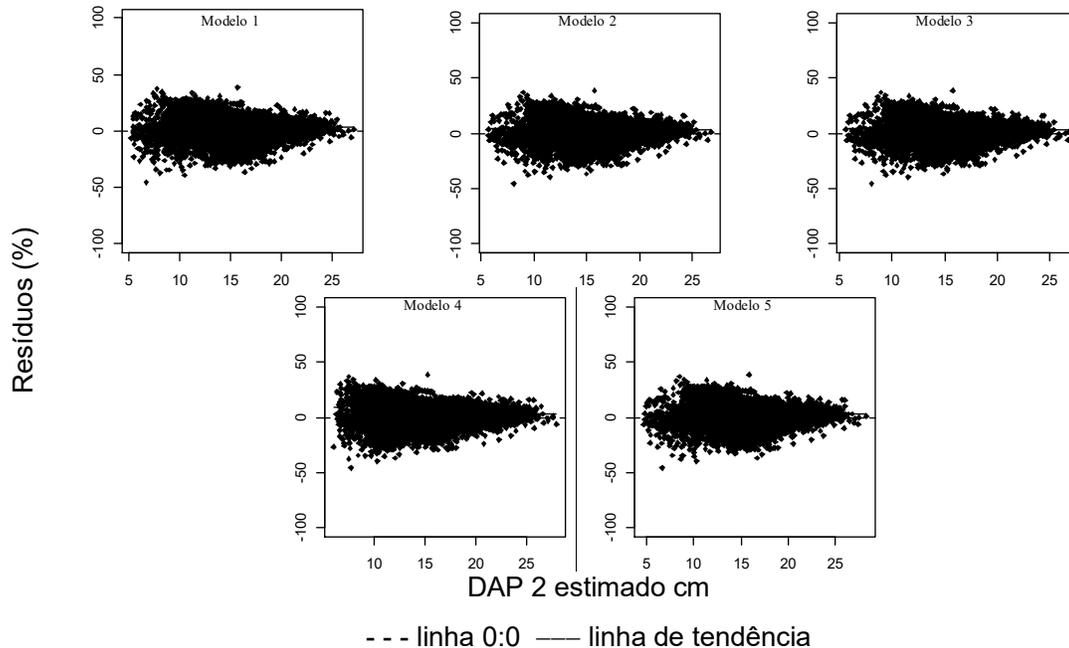


Figura 2 – Distribuição dos resíduos percentuais em função dos diâmetros estimados para os cinco modelos de crescimento em diâmetro.



Avaliando as estatísticas das equações referentes aos modelos de crescimento em altura, apresentadas na Tabela 5, a equação ajustada para o modelo Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014) também obteve melhor resultado dentre as estatísticas avaliadas, apresentando 0,9538, mais acurado com RMSE (%) de 5,2265 quando comparado com as demais equações e também menos viesado BIAS (%), com valor insignificante perante os demais.

Tabela 5 – Estatísticas utilizadas para avaliar o ajuste dos modelos não lineares de crescimento em altura total (HT).

Modelo	$r_{\hat{y}y}$	RMSE (%)	BIAS (%)
Linear, adaptado de Bella (1971) e Campos e Leite (2009), (MARTINS et al., 2014)	0,9538	5,23	$1,30 \cdot 10^{-6}$
Logístico, Amaro et al. (1998)	0,9431	5,77	-0,5061
Richards, Zeide (1993)	0,9431	5,77	-0,5026
Pienaar e Schiver (1981)	0,9372	6,32	-0,8204
Schumacher, Tomé (2003)	0,9339	8,37	-0,4215

Apresentando desempenho também inferior para crescimento em altura para os dados utilizados, figura a equação ajustada para o modelo de Schumacher e Tomé (2003), seguida pelo resultado da equação referente ao modelo de Pienaar e Schiver (1981).

Como observado no crescimento em diâmetro, percebeu-se uma tendência de superestimação com árvores de menor tamanho, acentuada no caso do modelo 4. O modelo 5 apresentou a mesma tendência, diferente do que ocorreu para o crescimento em diâmetro, que pode ser explicado pelo fato da altura ser uma variável com maior grau de dificuldade em se mensurar quando comparada ao diâmetro, sendo estimada por equações hipsométricas, carregando consigo os erros provenientes das mesmas.

Com base na avaliação das estatísticas apresentadas, pode-se afirmar que as equações de crescimento em diâmetro a 1,30 m de altura (DAP) e altura total (HT) do modelo Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014) (expressão 5) apresentam-se como as melhores opções para explicar o comportamento dessas variáveis para os dados analisados, segundo as estatísticas avaliadas e a análise gráfica dos resíduos.

Figura 3 – Altura total estimadas *versus* observadas pelas equações ajustadas para os cinco modelos de crescimento em altura.

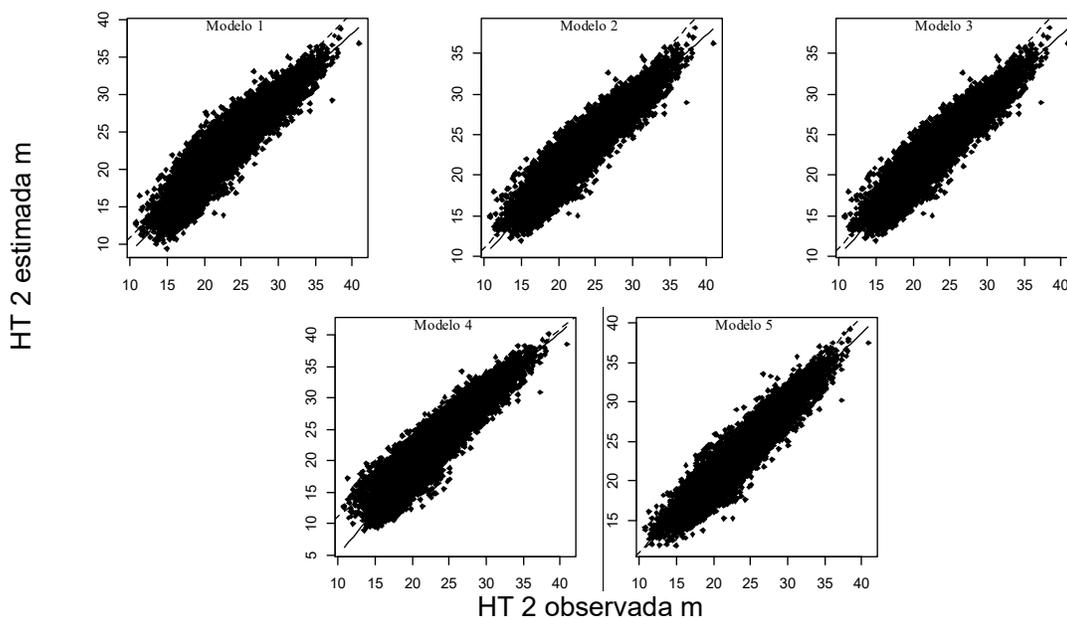
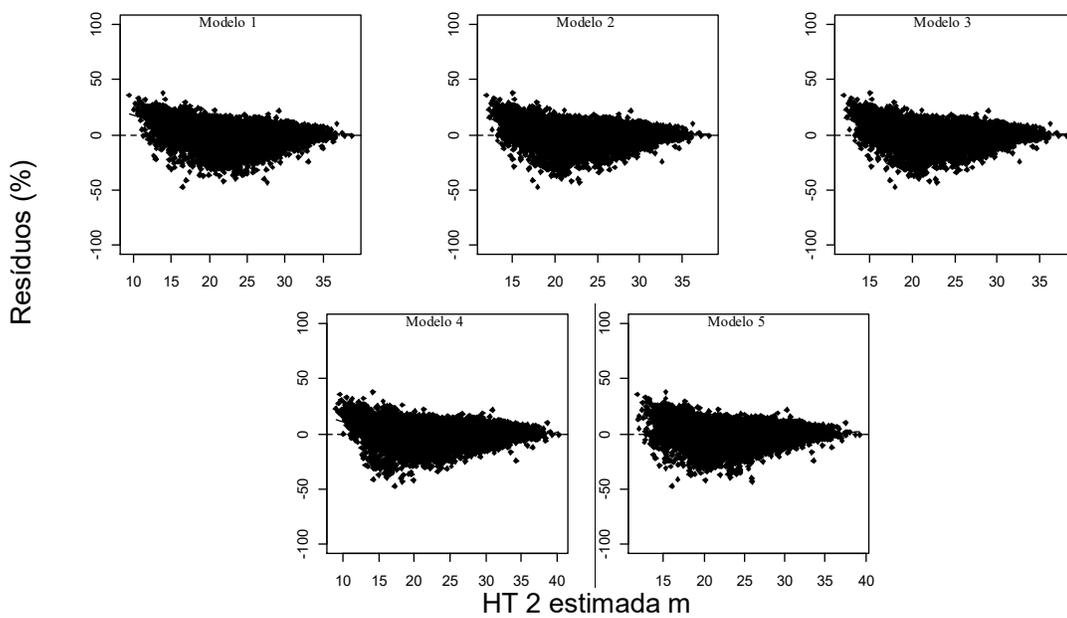


Figura 4 – Distribuição dos resíduos percentuais em função da altura (HT) estimada para os cinco modelos de crescimento em altura.



3.2. Critério 1 – Alternativas usando outras variáveis como grupos de efeitos aleatórios

Baseando-se nas equações de crescimento em diâmetro e altura referentes ao modelo Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014), foram definidas seis alternativas de modelagem mista, somadas à alternativa fixa (1) totalizam sete alternativas apresentadas na Tabela 6, avaliadas por classe de produtividade neste capítulo utilizando o conjunto de dados do critério 1. Salienta-se que a primeira alternativa foi avaliada na seção anterior sem considerar as classes de produtividade.

Tabela 6 – Alternativas de modelagem definidas para o crescimento em diâmetro na altura do peito e altura total de árvores individuais.

Alternativa	Modelo
1	$Y_2 = Y_1 + \left(\beta_0 + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S \right) + \varepsilon$, Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014)
2	$Y_2 = Y_1 + \left((\beta_0 + b_0) + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S \right) + \varepsilon$, modelo misto Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014), tendo a variável 'ClasseDimensional' como variável de agrupamento de efeitos aleatórios.
3	$Y_2 = Y_1 + \left((\beta_0 + b_0) + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S \right) + \varepsilon$, Modelo misto Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014), tendo a variável <i>classe</i> Y_1 como variável de agrupamento de efeitos aleatórios.
4	$Y_2 = Y_1 + \left((\beta_0 + b_0) + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S + b_4 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) \bar{H}_1 \right) + \varepsilon$, modelo misto Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014), tendo a variável 'ClasseDimensional' como variável de agrupamento de efeitos aleatórios, e a altura média na idade 1 como inclinação aleatória
5	$Y_2 = Y_1 + \left((\beta_0 + b_0) + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S + b_4 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) S \right) + \varepsilon$, modelo misto Linear adaptado de Bella (1971) e Campos e Leite

(2009) (MARTINS et al., 2014), tendo a variável 'ClasseDimensional' como variável de agrupamento de efeitos aleatórios, e o índice de local como inclinação aleatória

$$Y_2 = Y_1 + \left((\beta_0 + b_0) + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S + b_4 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) Hd_1 \right) + \varepsilon,$$

- 6 modelo misto Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014), tendo a variável 'ClasseDimensional' como variável de agrupamento de efeitos aleatórios, e a altura dominante na idade 1 como inclinação aleatória

$$Y_2 = Y_1 + \left((\beta_0 + b_0) + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S \right) + \varepsilon, \text{ modelo misto}$$

- 7 Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014), tendo o material genético como variável de agrupamento de efeitos aleatórios.

em que: Y_2 = diâmetro (em cm) ou altura (m) em idade futura; Y_1 = diâmetro (cm) ou altura (m) em idade atual; I_2 = idade futura (meses); I_1 = idade atual (meses); S = índice de local; $BAI = d_i^2 / q^2$; q = diâmetro quadrático (cm); β_i = coeficientes do modelo; e ε = erro aleatório;

3.2.1. Ajuste das alternativas de modelagem para crescimento em diâmetro

Na Tabela 7 são apresentadas as estatísticas das equações ajustadas para as alternativas da estratégia 1 por classe de produtividade. Verifica-se que todas as alternativas apresentaram valores semelhantes nas estatísticas das equações ajustadas para crescimento em diâmetro.

Na classe de produtividade baixa ($S = 21$), o RMSE (%) da alternativa 5, apesar de próximo das demais alternativas na mesma classe, é mais acurado, acompanhado de um valor de $r_{\hat{y}\hat{y}} = 0,9793$, inferior ao valor da mesma estatística da alternativa 6.

O valor de RMSE (%) de 3,2110 na classe de produtividade média ($S = 27$) para a alternativa 4 é mais acurado comparado à mesma classe nas demais alternativas. Também nesta classe, observa-se o valor da $r_{\hat{y}\hat{y}}$ de 0,9843, superior aos valores desta estatística na classe de produtividade em outras alternativas e BIAS (%) de -0,1600. A alternativa 5 apresentou um valor de RMSE (%) de 3,5270, inferior, contudo, próximo aos valores desta estatística na mesma classe de produtividade. A correlação calculada para esta alternativa nesta classe é de 0,9846, superior às demais alternativas, e o BIAS (%) de 0,0868, insignificante quando é comparado aos valores apresentados nas alternativas 4 e 6.

Tabela 7 – Estatísticas utilizadas para avaliar o ajuste das sete alternativas para estimar o crescimento de diâmetro na altura do peito (DAP), por classe de produtividade.

Alternativa	Classe	$r_{\hat{y}\hat{y}}$	RMSE (%)	BIAS (%)
1	S = 21	0,9786	3,94	1,2686
	S = 27	0,9811	3,53	-0,2055
	S = 33	0,9809	4,09	-0,2516
2	S = 21	0,9788	3,85	1,0970
	S = 27	0,9835	3,30	-0,1913
	S = 33	0,9824	3,83	0,0330
3	S = 21	0,9783	3,92	1,1270
	S = 27	0,9816	3,48	-0,2046
	S = 33	0,9812	4,02	0,2361
4	S = 21	0,9797	3,73	0,9267
	S = 27	0,9843	3,21	-0,1600
	S = 33	0,9833	3,69	-0,0132
5	S = 21	0,9793	3,72	0,6555
	S = 27	0,9838	3,26	-0,1170
	S = 33	0,9846	3,53	0,0868
6	S = 21	0,9797	3,74	0,9721
	S = 27	0,9840	3,24	-0,1665
	S = 33	0,9833	3,68	-0,0470
7	S = 21	0,9785	3,87	0,9745
	S = 27	0,9822	3,43	-0,1736
	S = 33	0,9821	3,85	0,1213

Os coeficientes estimados para as alternativas avaliadas são apresentados na Tabela 8, sendo todos foram significativos a 1%.

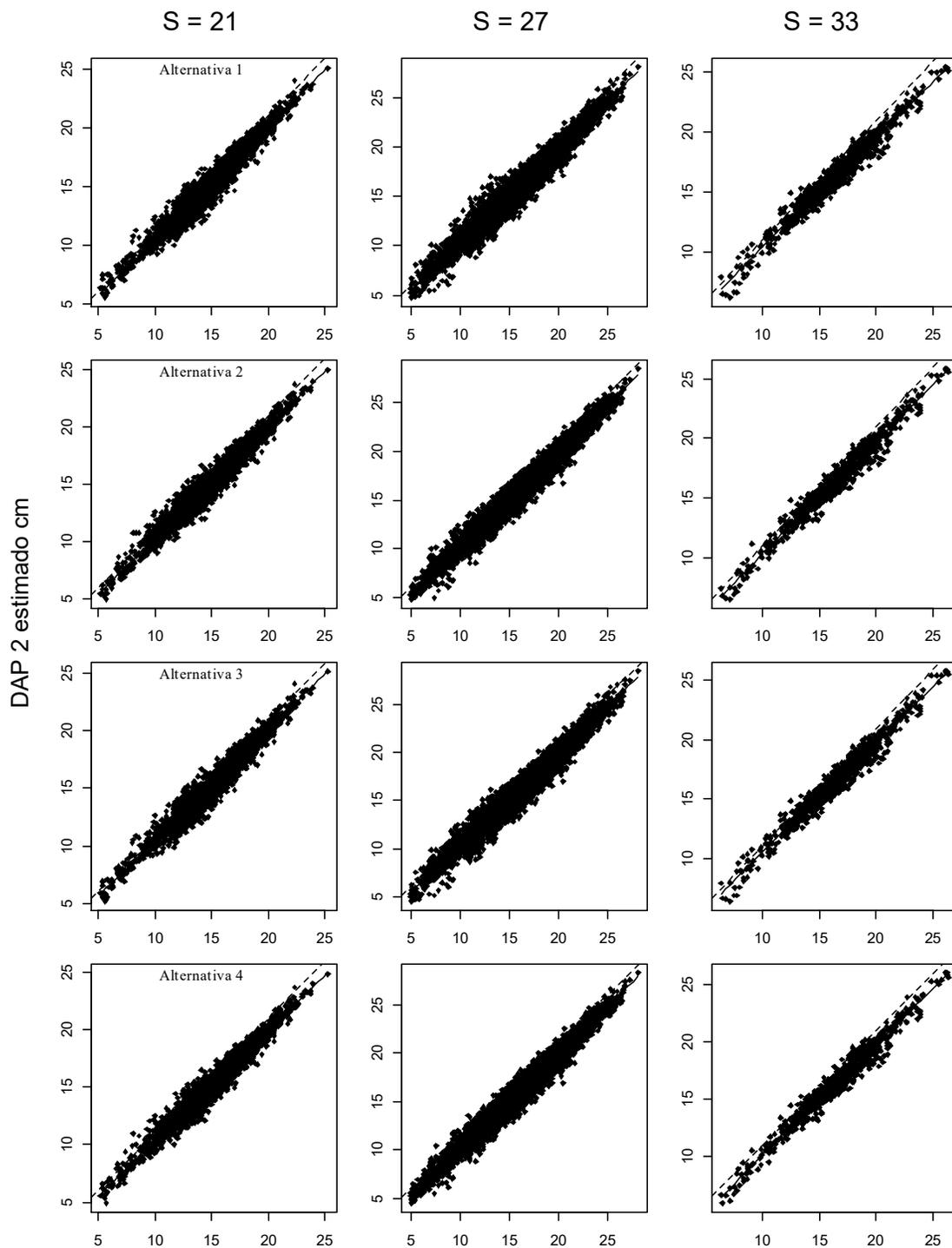
Tabela 8 – Estimativas dos coeficientes das equações referentes às alternativas propostas para crescimento em diâmetro (DAP).

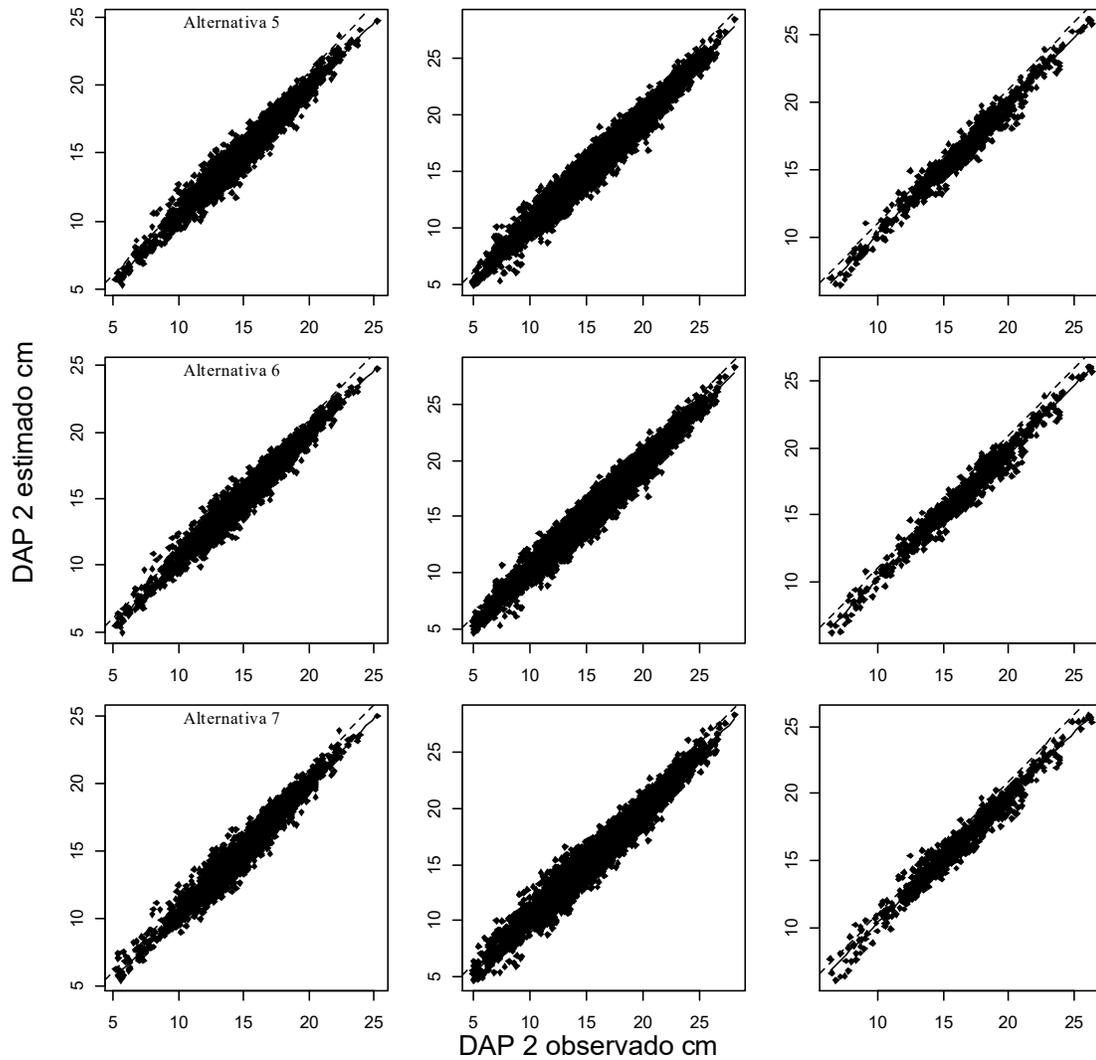
Alternativa	Coeficientes			
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
1	0.3373**	-213.5**	1.1360**	-0.0437**
2	-1.0860**	-116.5**	1.4080**	0.0192**
3	-0.6715**	-178.4**	1.7220**	-0.0216**
4	-1.0450**	-198.4**	1.1130**	0.0135**
5	-0.8220**	-81.83**	1.4250**	0.0116**
6	-1.0030**	-174.9**	1.1980**	0.0121**
7	-0.4152**	-214.5**	1.3100**	-0.0217**

** - significativo a 1% pelo Test T para modelos fixos, e Teste de Wald para modelos mistos

Avaliando os gráficos da Figura 5, verifica-se comportamento semelhante ao diâmetro estimado *versus* diâmetro observado em todas as classes de produtividade, com uma leve tendência de superestimação em todas as classes de produtividade.

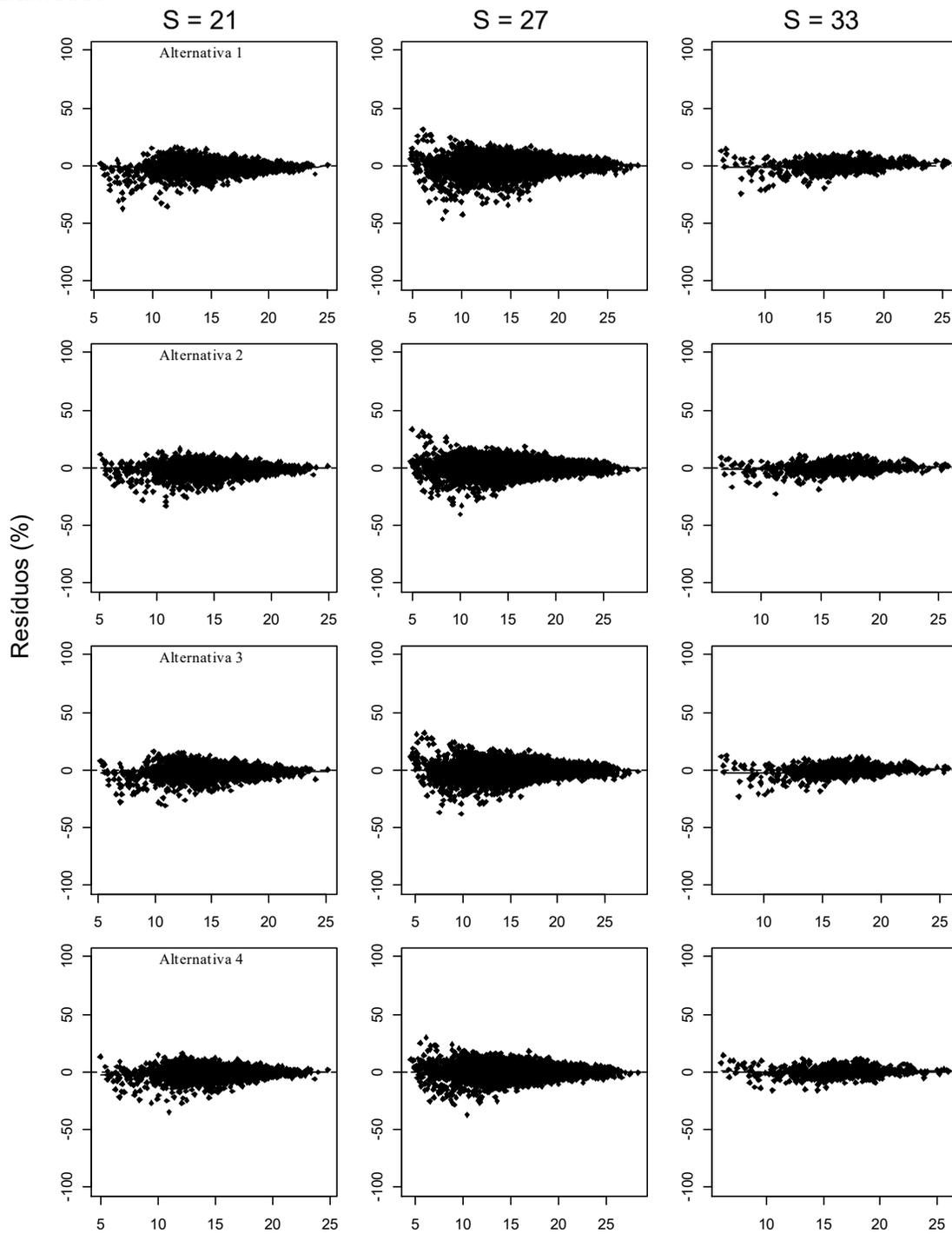
Figura 5 – Diâmetros na altura do peito estimados *versus* observados pelas equações ajustadas referentes aos sete modelos de crescimento em diâmetro, em cada classe de produtividade.

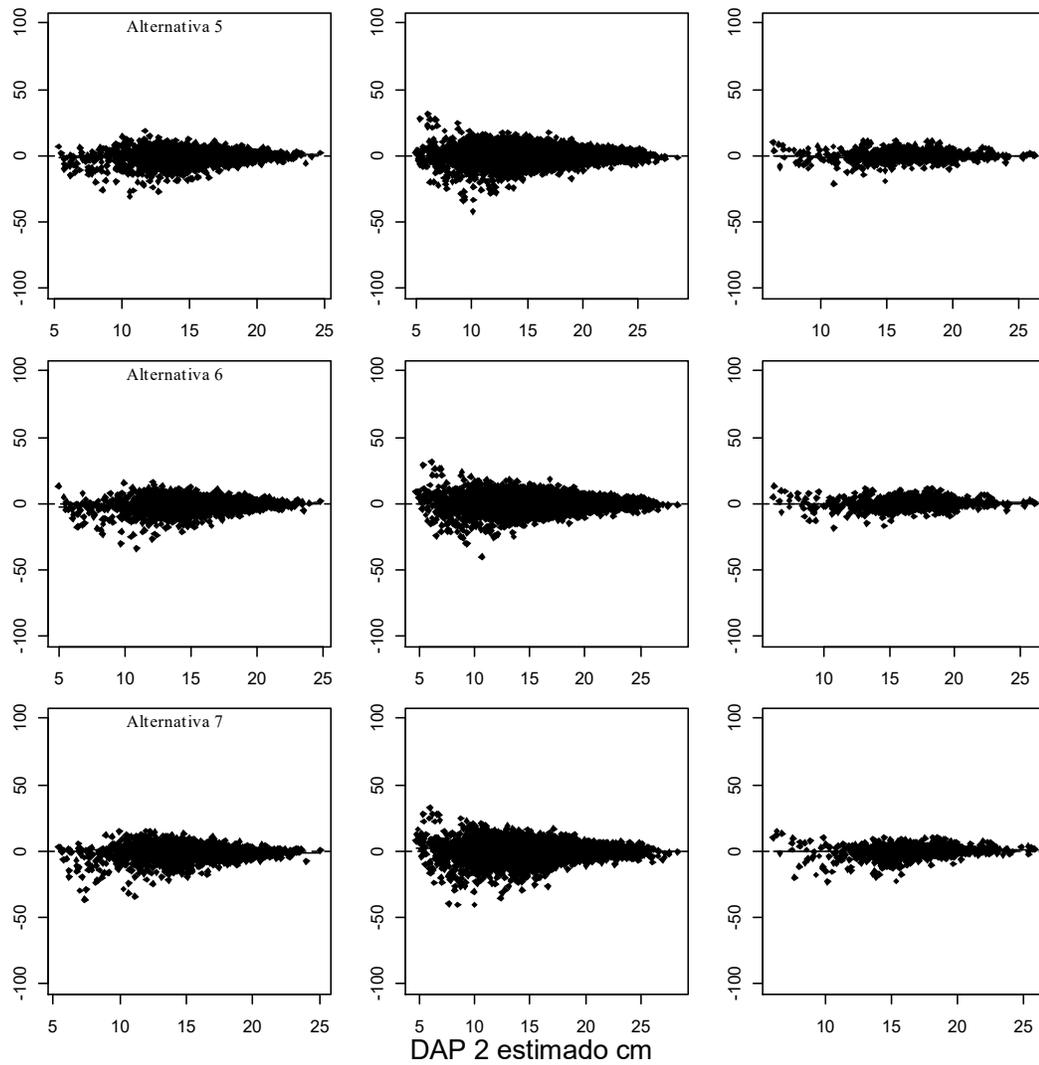




Na avaliação dos resíduos em porcentagem, apresentados na Figura 6, embora semelhantes, destaca-se a distribuição dos resíduos mais concentrada na alternativa 4, em todas as classes de produtividade. As alternativas 1, 3 e 7 apresentam maior dispersão também em todas as classes de produtividade.

Figura 6 – Distribuição dos resíduos percentuais em função dos diâmetros (DAP) estimados, em cada classe de produtividade para os modelos de crescimento em diâmetro.





As alternativas mistas (2 a 7) para modelagem de crescimento em diâmetro apresentaram comportamentos semelhantes quando comparadas à alternativa 1, que utiliza modelagem fixa. Contudo, pela avaliação das estatísticas de ajuste em todas as classes de produtividade e pela análise gráfica dos resíduos percentuais, pode-se afirmar que a alternativa 4 foi ligeiramente mais acurada na estimação do crescimento de diâmetro frente às demais alternativas propostas.

3.2.2. Ajuste das alternativas de modelagem para crescimento em altura total

As estatísticas obtidas a partir das equações para estimativa de crescimento de altura, apresentadas na Tabela 9, semelhantes em todas as classes de produtividade, assim como ocorreu na avaliação das alternativas para estimativa de crescimento de diâmetro. Contudo, destaca-se a alternativa 5, apresentando maior acurácia em todas as classes de produtividade, com RMSE (%) apresentando valores de 5.3700, 4.6810 e 4.1040 nas classes baixa, média e alta, respectivamente, sendo inferiores aos mesmos valores nestas classes para todas as alternativas.

Tabela 9 – Estatísticas utilizadas para avaliar o ajuste das sete alternativas para estimar o crescimento de altura total (HT), por classe de produtividade.

Alternativa	Classe	r_{yy}	RMSE (%)	BIAS (%)
1	S = 21	0,9419	6,40	2,7840
	S = 27	0,9532	5,06	-0,3839
	S = 33	0,9571	4,63	-1,5410
2	S = 21	0,9479	6,01	2,8930
	S = 27	0,9582	4,74	-0,3988
	S = 33	0,9613	4,46	-1,6010
3	S = 21	0,9436	6,25	3,0200
	S = 27	0,9564	4,89	-0,4316
	S = 33	0,9586	4,50	-1,3230
4	S = 21	0,9496	5,84	2,7130
	S = 27	0,9591	4,69	-0,3725
	S = 33	0,9633	4,35	-1,5350
5	S = 21	0,9499	5,37	1,4570
	S = 27	0,9589	4,68	-0,2111
	S = 33	0,9628	4,10	-0,5715
6	S = 21	0,9495	5,85	2,7210
	S = 27	0,9590	4,69	-0,3729
	S = 33	0,9637	4,34	-1,5590
7	S = 21	0,9447	6,17	2,5050
	S = 27	0,9544	5,00	-0,3319
	S = 33	0,9567	4,71	-1,6970

Verifica-se ainda, que a alternativa 5 é menos viesada quando se avalia a estatística BIAS (%) em todas as classes de produtividade, com tendência de subestimação para árvores mais altas (Figura 7 e Figura 8) com valores de 1.4570 para a classe de produtividade baixa (S = 21), -0.2111 na classe de produtividade média (S=27) e -0.5715 na classe de produtividade alta (S = 33). A correlação (r_{yy}) foi

superior apenas na classe de produtividade baixa ($S = 21$), com valor de 0.9499, e inferior nas demais classes para esta alternativa.

Todos os coeficientes fixos estimados apresentados na Tabela 10 foram significativos a 1% para as alternativas propostas para estimativa de crescimento de altura.

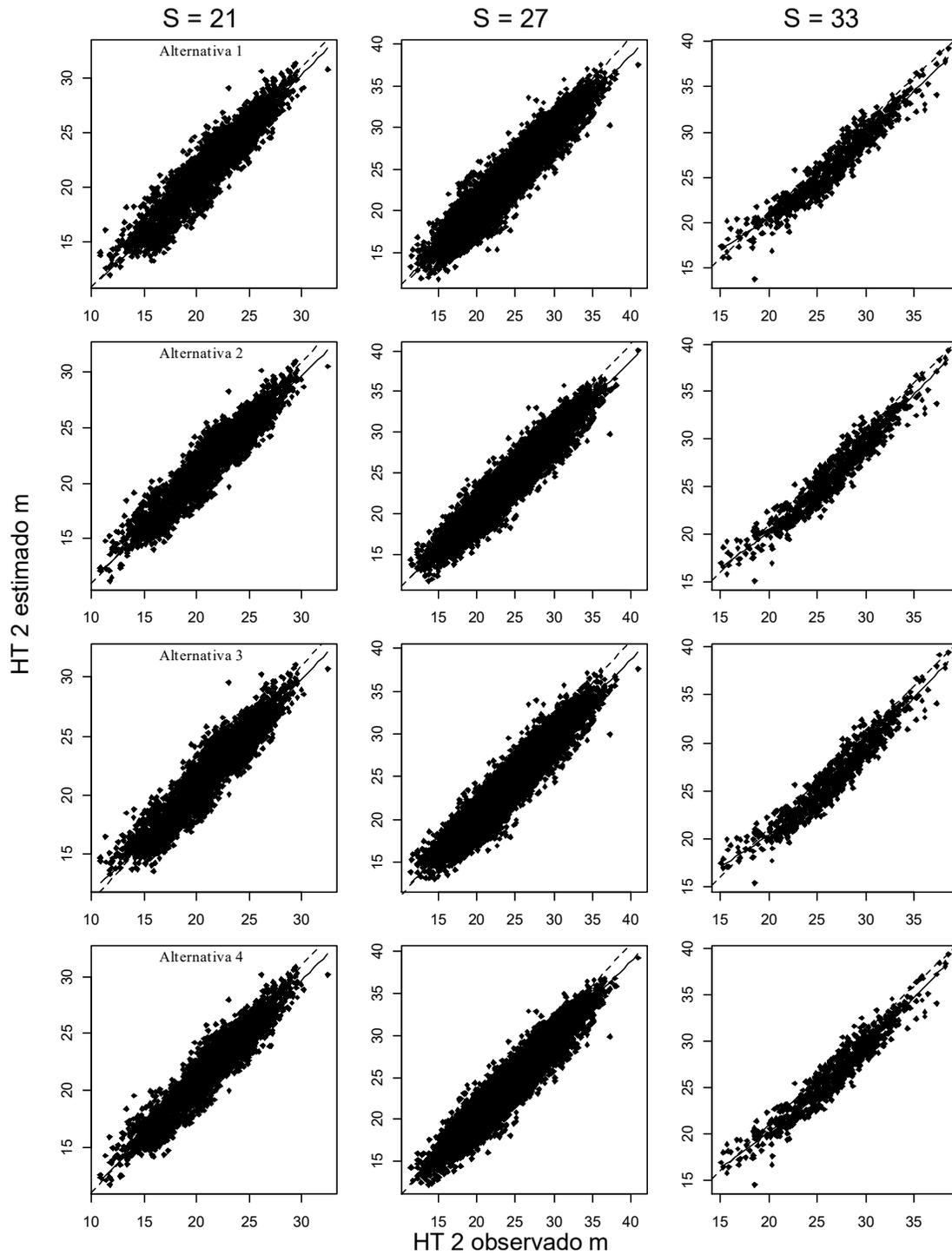
Tabela 10 – Estimativas dos coeficientes das equações referentes às alternativas propostas para crescimento em altura total (HT).

Alternativa	Coeficientes			
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
1	1,4670**	-355,10**	1,2590**	-0,0608**
2	-0,7362**	-147,30**	1,0730**	0,0712**
3	-1,0470**	-157,30**	2,1090**	0,0531**
4	-0,6686**	-224,00**	0,7733**	0,0580**
5	0,0140**	91,47**	1,2780**	0,0435**
6	-0,8567**	-228,10**	0,8051**	0,0613**
7	2,0390**	-352,70**	1,2590**	-0,0745**

** - significativo a 1% pelo Test T para modelos fixos, e Teste de Wald para modelos mistos

Observando-se as alturas estimadas em comparação com as alturas observadas na Figura 7, não é possível detectar diferenças entre as alternativas analisadas para crescimento em altura, exceto na classe de produtividade média da alternativa 6, onde os pontos apresentam uma leve aproximação da linha 0:1, nos menores valores de altura. Ao analisar graficamente os resíduos percentuais (Figura 8), observa-se que nas classes de produtividade média ($S = 27$) e alta ($S = 33$), a dispersão dos resíduos é menor nas alternativas 4 e 6. A alternativa 1 apresenta maior dispersão dos resíduos em percentual em todas as classes de produtividade avaliadas.

Figura 7 – Alturas totais estimadas *versus* observadas pelas equações ajustadas referentes aos sete modelos de crescimento em altura total, em cada classe de produtividade.



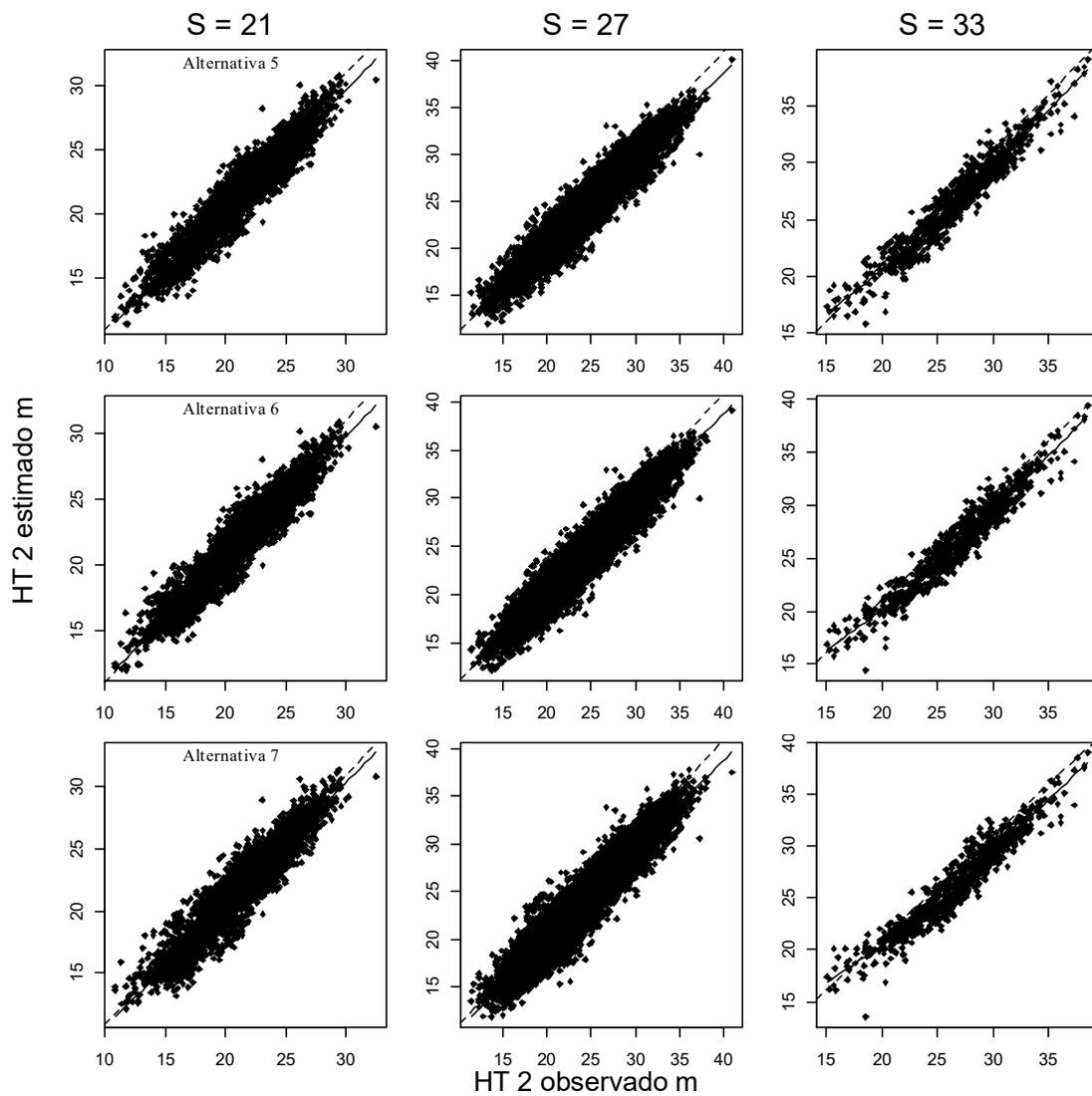
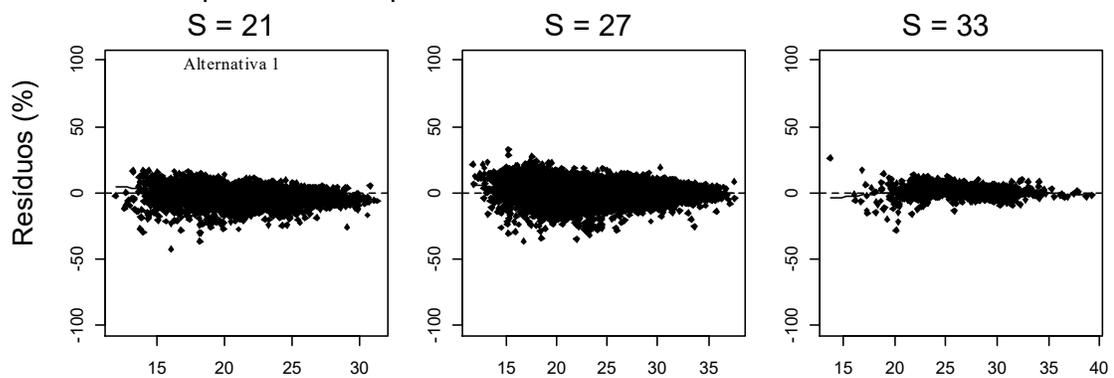
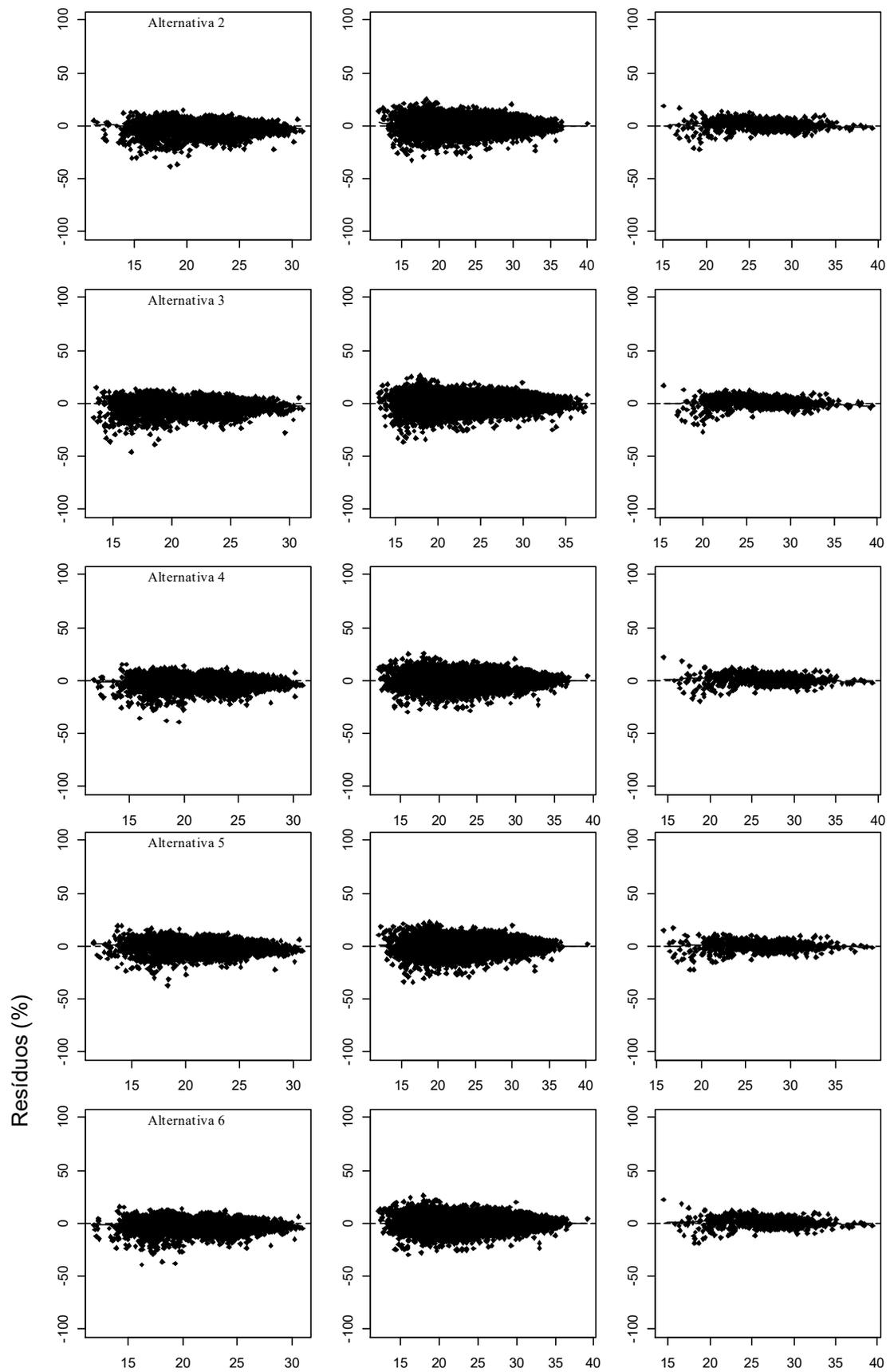
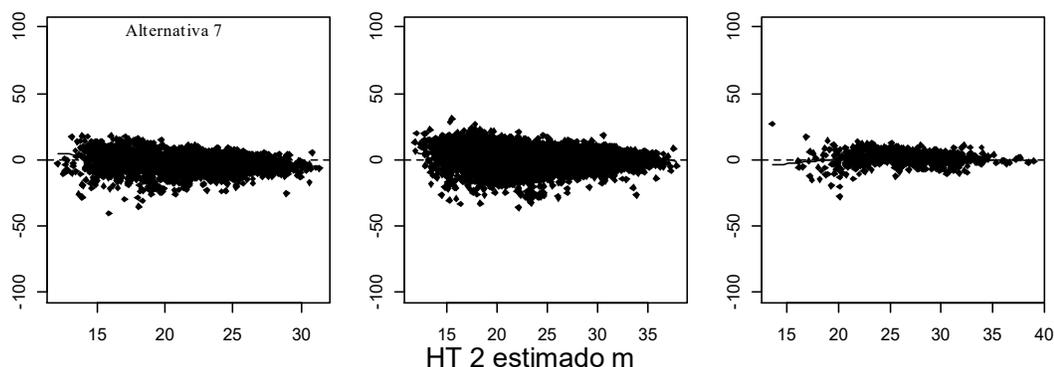


Figura 8 – Distribuição dos resíduos percentuais em função das alturas estimadas, em cada classe de produtividade para os modelos de crescimento em altura total.







3.2.3. Resultados de volume do modelo de árvore individual

As estatísticas obtidas pela avaliação do volume total obtido comparado ao volume observado são apresentadas na Tabela 11.

A alternativa 5 apresenta RMSE (%) com valor de 6,97, mais acurado quando comparada às demais alternativas na classe de produtividade baixa ($S = 21$), e menos viesado em todas as classes de produtividade (Figura 9), na mesma classe em outras alternativas. Este resultado é confirmado ao avaliar tais estatísticas, obtidas após aplicação da alternativa 5 em uma base de dados de validação com dados independentes do ajuste (Tabela 12). Embora a correlação nesta alternativa seja inferior aos valores das alternativas 1, 4 e 6, verifica-se maior correlação ao realizar a validação desta alternativa.

O mesmo comportamento é observado na classe de produtividade alta ($S = 33$) para esta alternativa. Contudo, apesar de próximos, o valor de RMSE (%) para a alternativa 5 nesta classe de produtividade nas estatísticas de validação foram ligeiramente superiores para RMSE (%) e BIAS (%) no ajuste. Na classe de produtividade média ($S = 27$), o RMSE (%) obtido na alternativa 4 também foi mais acurado, com diferença insignificante quando comparado ao valor obtido na mesma classe da alternativa 5.

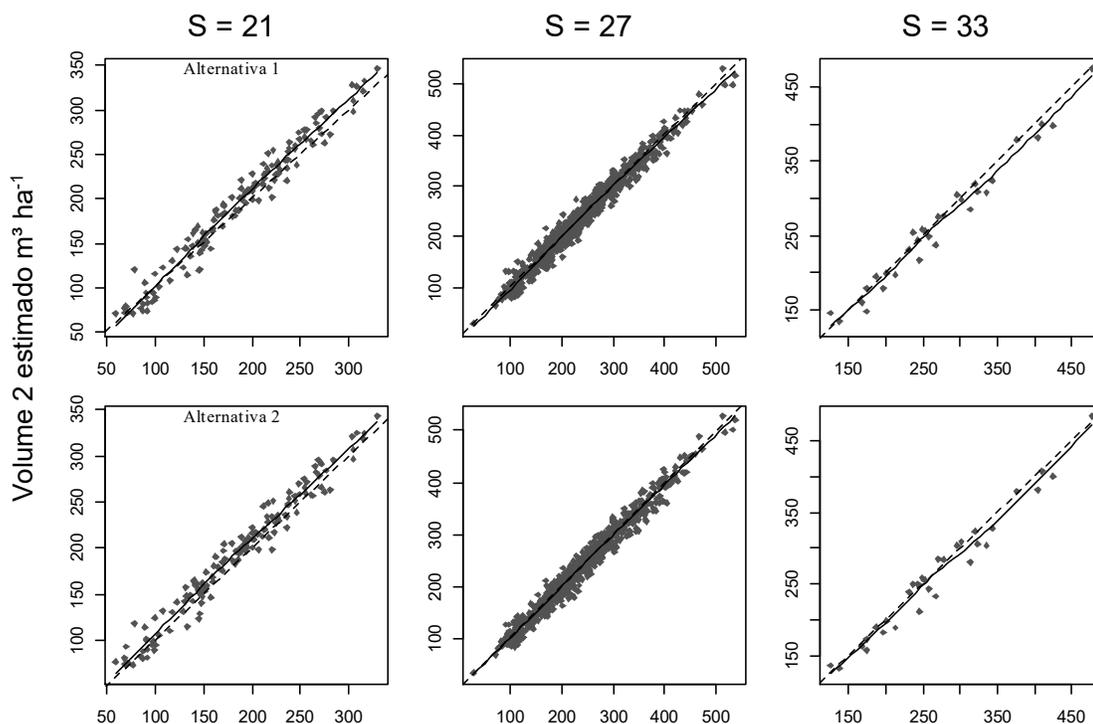
Tabela 11 – Estatísticas utilizadas para avaliar o volume obtido a partir do ajuste das sete alternativas de crescimento em diâmetro e altura, por classe de produtividade.

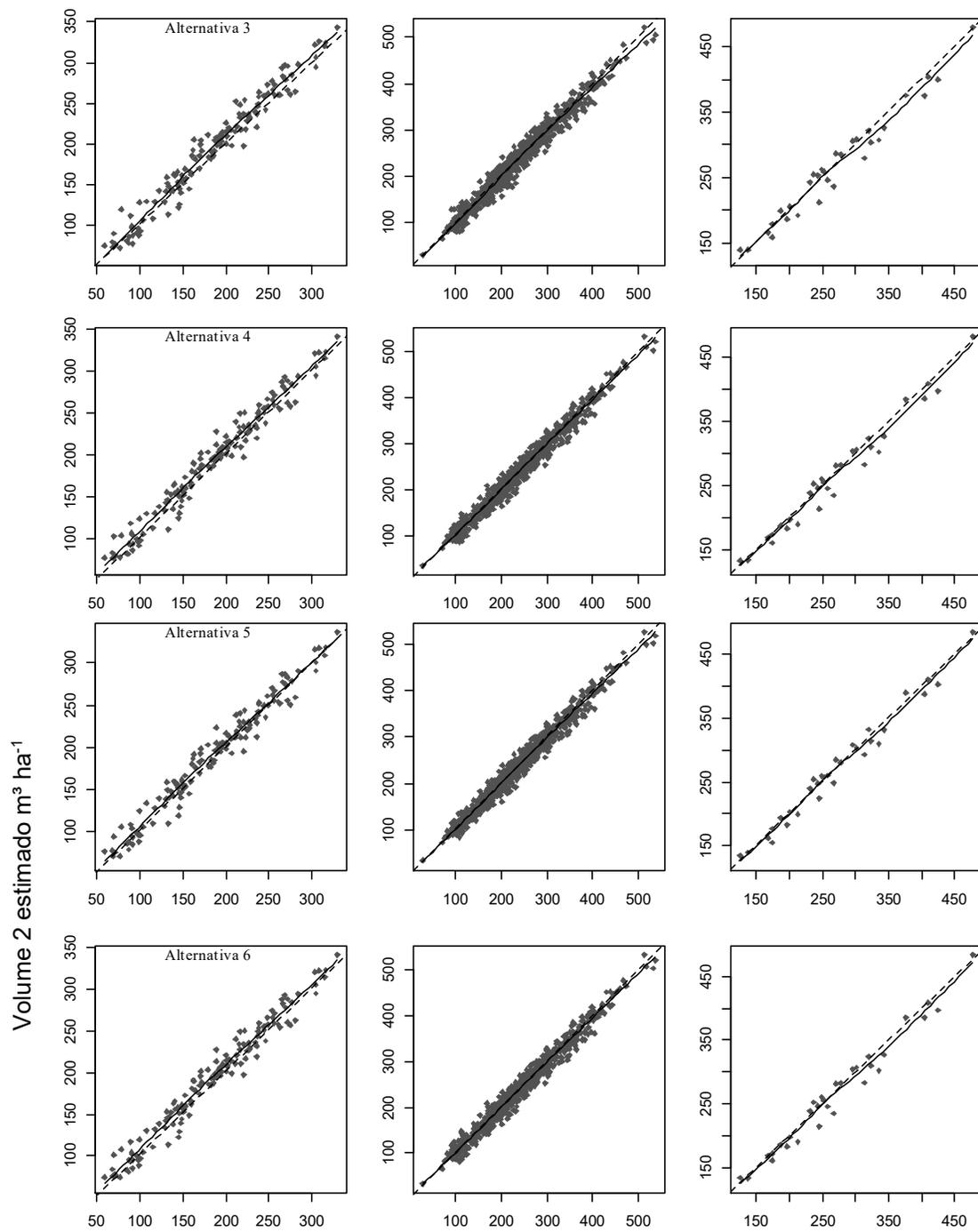
Alternativa	Classe	$r_{\hat{y}y}$	RMSE (%)	BIAS (%)
1	S = 21	0,9829	8,30	4,4746
	S = 27	0,9855	5,93	-0,4373
	S = 33	0,9864	5,73	-2,3953
2	S = 21	0,9815	8,26	4,9111
	S = 27	0,9858	5,84	-0,7751
	S = 33	0,9840	5,99	-2,1455
3	S = 21	0,9790	8,96	5,2579

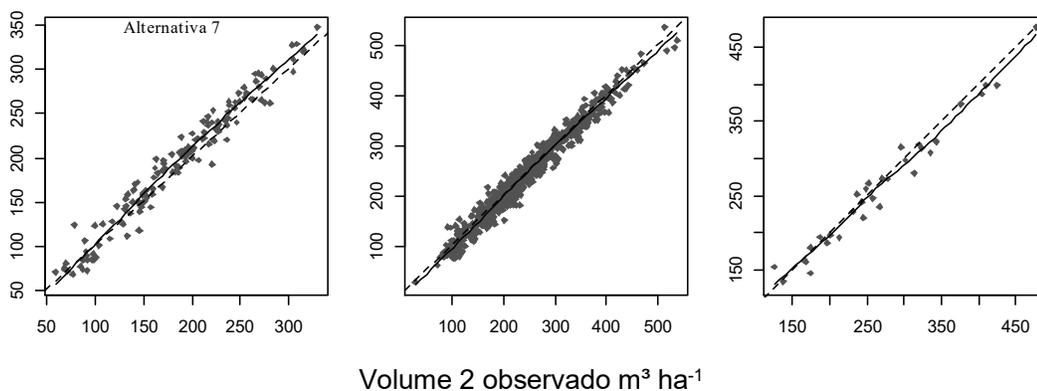
	S = 27	0,9838	6,24	-0,8314
	S = 33	0,9823	6,18	-1,721
4	S = 21	0,9823	7,71	4,2376
	S = 27	0,9867	5,65	-0,7018
	S = 33	0,9855	5,68	-1,9427
5	S = 21	0,9817	6,97	2,4434
	S = 27	0,9863	5,70	-0,4715
	S = 33	0,9891	4,69	-0,7678
6	S = 21	0,9822	7,78	4,3286
	S = 27	0,9865	5,70	-0,7039
	S = 33	0,9854	5,71	-1,9887
7	S = 21	0,9797	8,86	4,80562
	S = 27	0,9835	6,34	-0,5173
	S = 33	0,9841	6,04	-2,2511

Apesar da ligeira superioridade da alternativa 5, observa-se que a magnitude dos valores das estatísticas obtidas para os dados de ajuste em todas as alternativas para todas as classes de produtividade é semelhante, o que pode ser confirmado pelos gráficos de volume estimados *versus* observados (Figura 9).

Figura 9 – Volumes estimados *versus* observados pelas equações ajustadas referentes aos sete modelos de crescimento em diâmetro e altura total, em cada classe de produtividade.



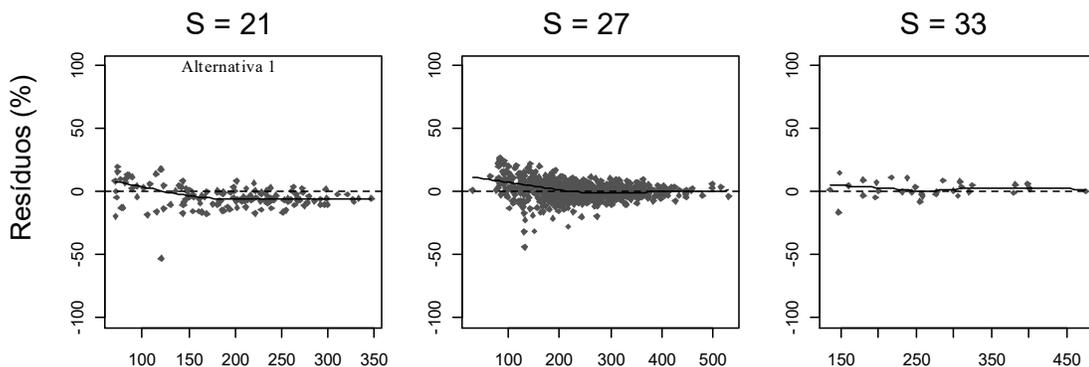


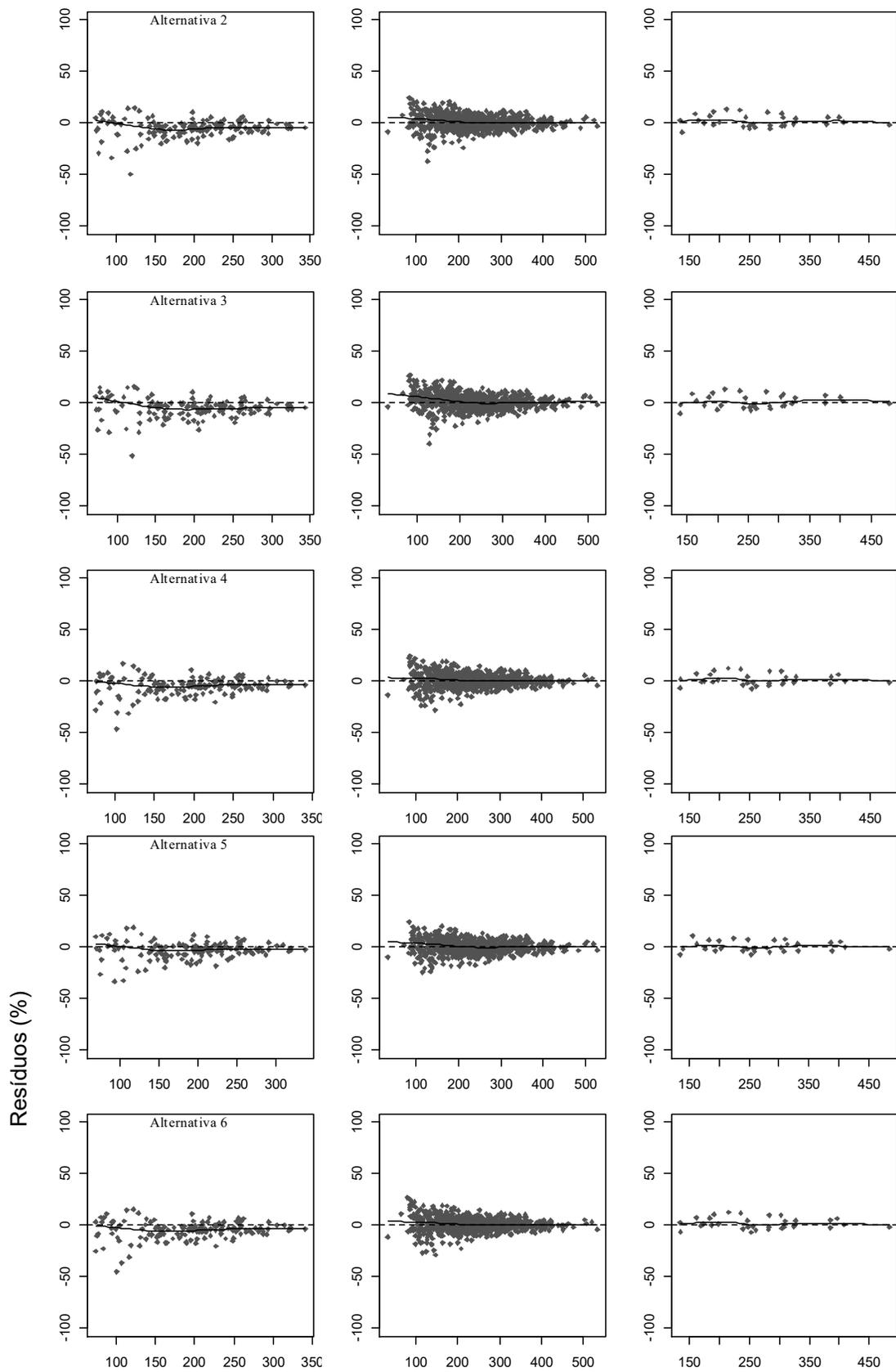


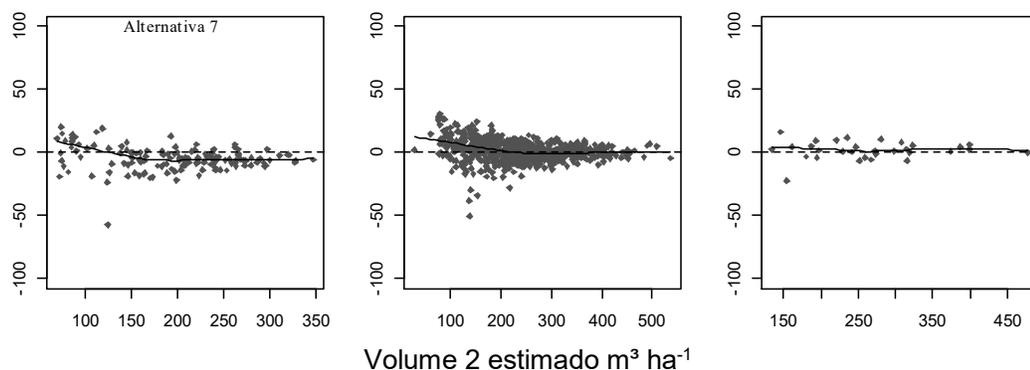
Avaliando a dispersão dos resíduos (Figura 10), observa-se uma tendência de superestimativa dos volumes nas classes de produtividade média ($S = 27$) nas alternativas 1, 3 e 7. O mesmo comportamento ocorre na classe de produtividade alta ($S = 33$) da alternativa 1, e menos dispersa na alternativa 7.

Observa-se na classe de produtividade baixa ($S = 21$) em todas as alternativas, tendência à subestimação a partir de volumes de amplitude média até as mais altas. Embora com um desvio, as alternativas 4 e 6, apresentando uma dispersão reduzida em todas as classes de produtividade quando comparada às demais alternativas ajustadas. No entanto, observando os resíduos percentuais (Figura 12) resultantes da aplicação das equações de crescimento a uma base de dados independente, verifica-se uma tendência de superestimação nas alternativas 4 e 6 em todas as classes de produtividade para valores menores de volume. Fato que não ocorre quando se avalia a aplicação da alternativa 2, que manteve comportamento semelhante no ajuste e na aplicação, corroborando com as estatísticas de validação apresentadas na Tabela 12.

Figura 10 – Distribuição dos resíduos percentuais em função dos volumes estimados, em cada classe de produtividade para os modelos de crescimento em diâmetro e altura.







Ainda considerando a aplicação realizada (Tabela 12) com o propósito de validar os modelos de crescimento de diâmetro e altura, cabe destacar a estabilidade mantida pela alternativa 2 caracterizada pela pouca variação, corroborada pelas análises gráficas realizadas.

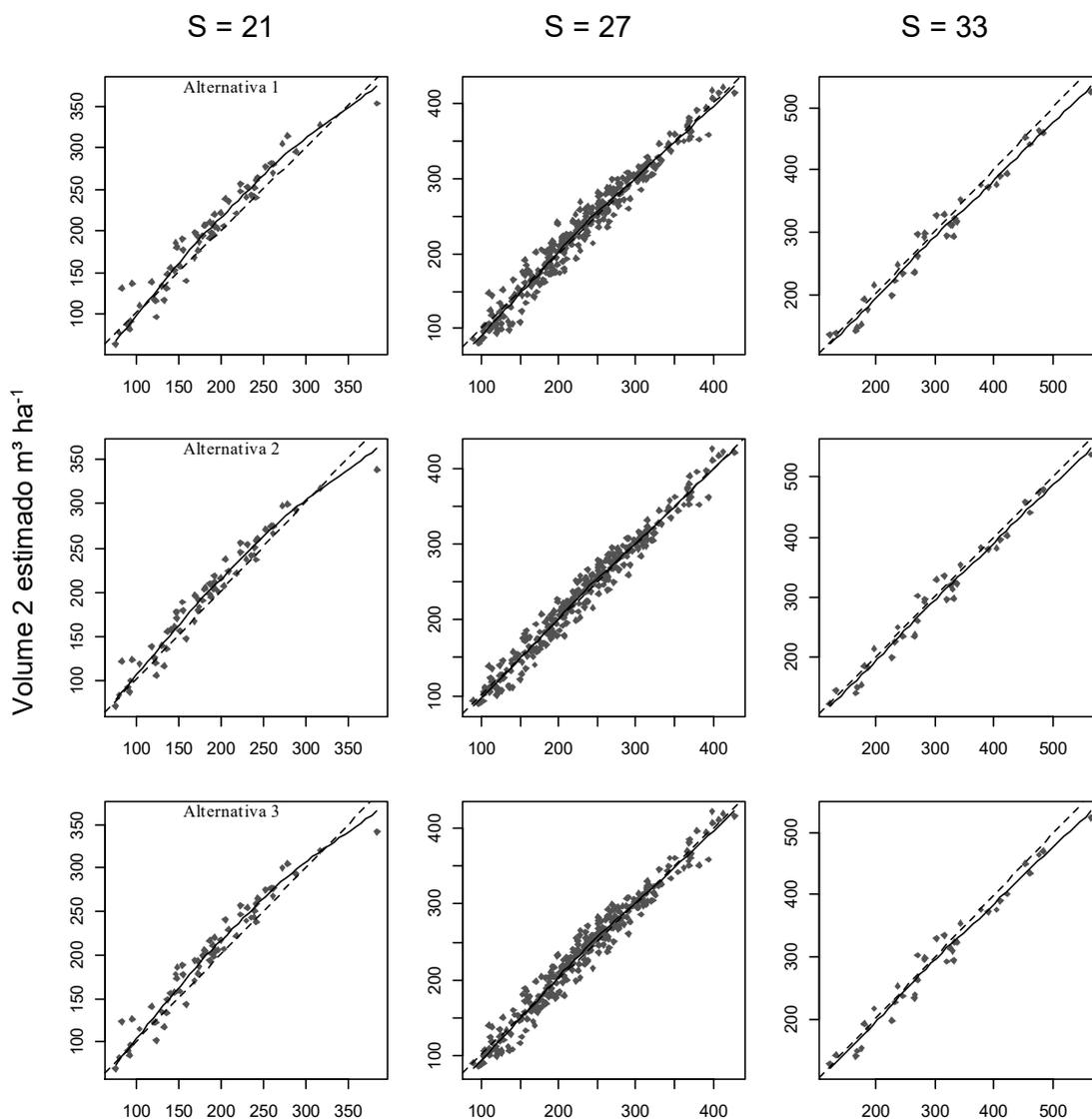
Salienta-se ainda na validação, que a alternativa 7 apresentou comportamento discrepante comparada às demais alternativas, com baixa acurácia e mais viesada nas classes de produtividade baixa e alta.

Tabela 12 – Estatísticas utilizadas para avaliar o volume obtido a partir da aplicação das sete alternativas de crescimento em diâmetro e altura, em uma base de dados de validação, com dados independentes dos utilizados no ajuste, por classe de produtividade.

Alternativa	Classe	r_{yy}	RMSE (%)	BIAS (%)
1	S = 21	0,9682	11,10	6,7505
	S = 27	0,9791	6,62	0,6349
	S = 33	0,9860	6,80	-3,4590
2	S = 21	0,9739	9,91	6,4053
	S = 27	0,9839	5,74	0,5775
	S = 33	0,9869	5,96	-2,1551
3	S = 21	0,9713	10,43	6,6469
	S = 27	0,9805	6,34	0,8780
	S = 33	0,9844	6,76	-2,8418
4	S = 21	0,9642	10,43	2,2578
	S = 27	0,9788	8,10	-1,8968
	S = 33	0,9851	7,16	-3,2151
5	S = 21	0,9820	7,62	3,9036
	S = 27	0,9854	6,11	-1,1506
	S = 33	0,9869	6,71	-3,4373
6	S = 21	0,9677	9,85	3,5969
	S = 27	0,9803	7,18	-0,9529
	S = 33	0,9856	6,68	-2,8465

7	S = 21	0,9527	16,45	12,6090
	S = 27	0,9592	9,86	1,2979
	S = 33	0,6809	28,70	-11,3406

Figura 11 – Volumes estimados *versus* observados pela aplicação das equações de crescimento em diâmetro e altura em uma base de dados independente do ajuste, em cada classe de produtividade.



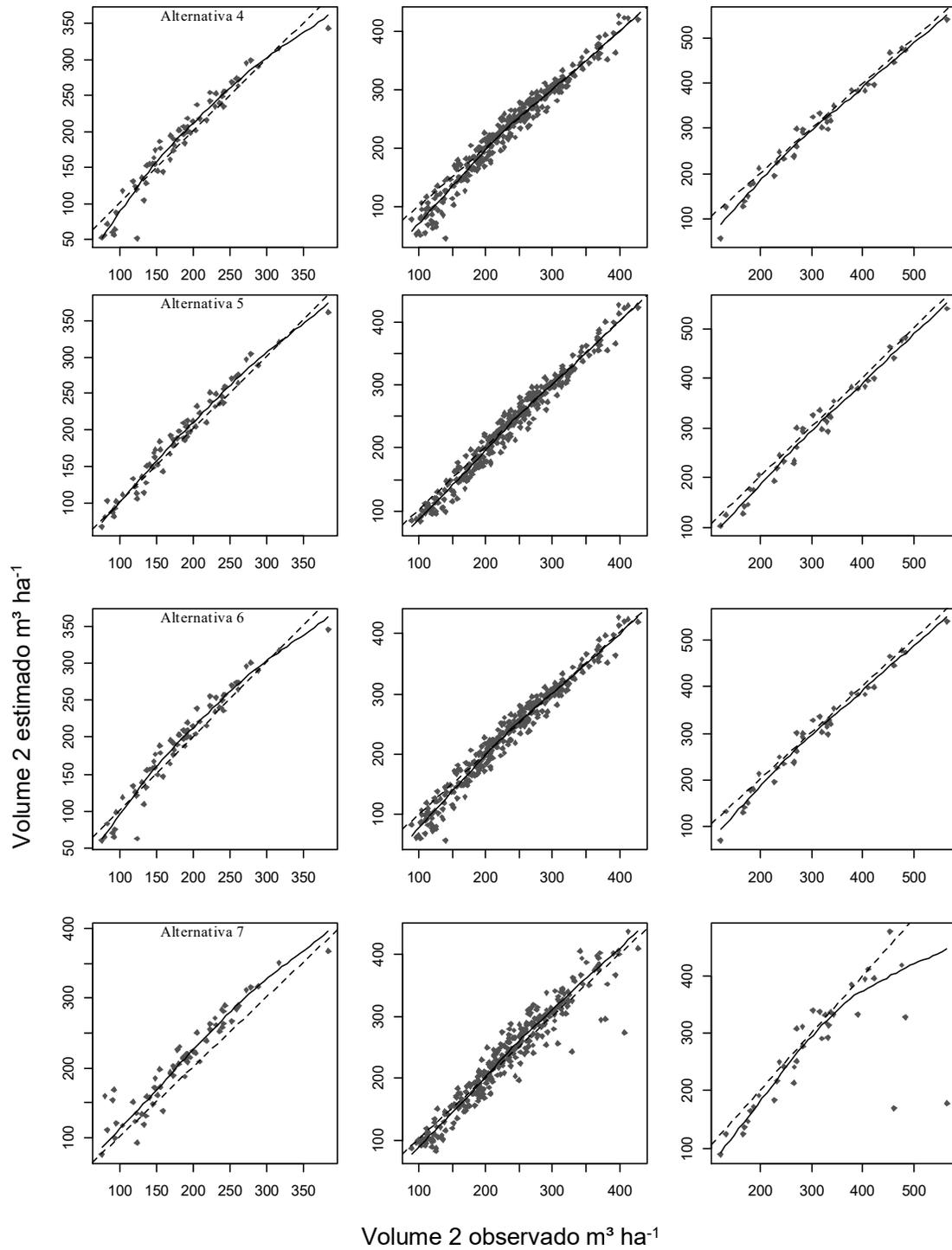
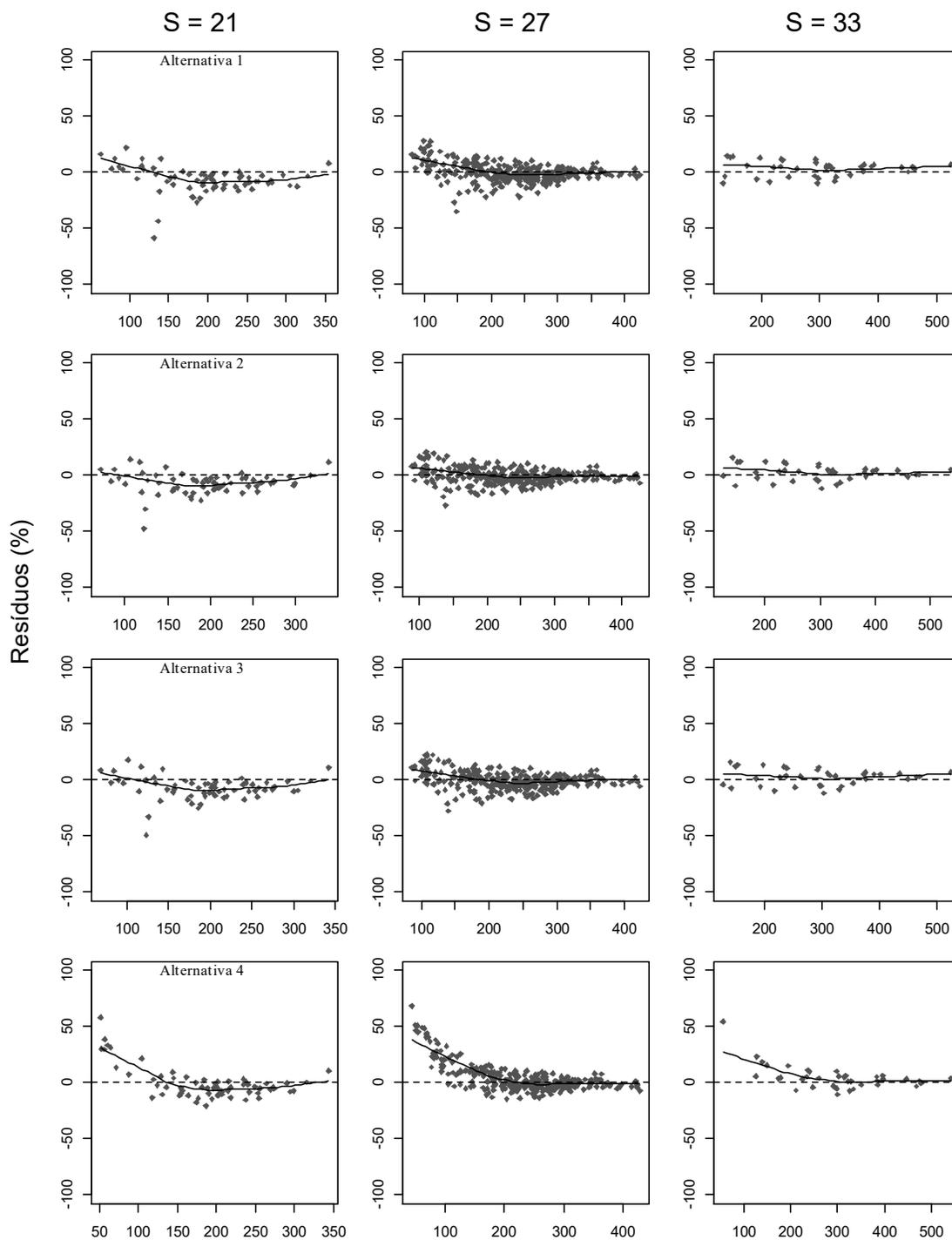
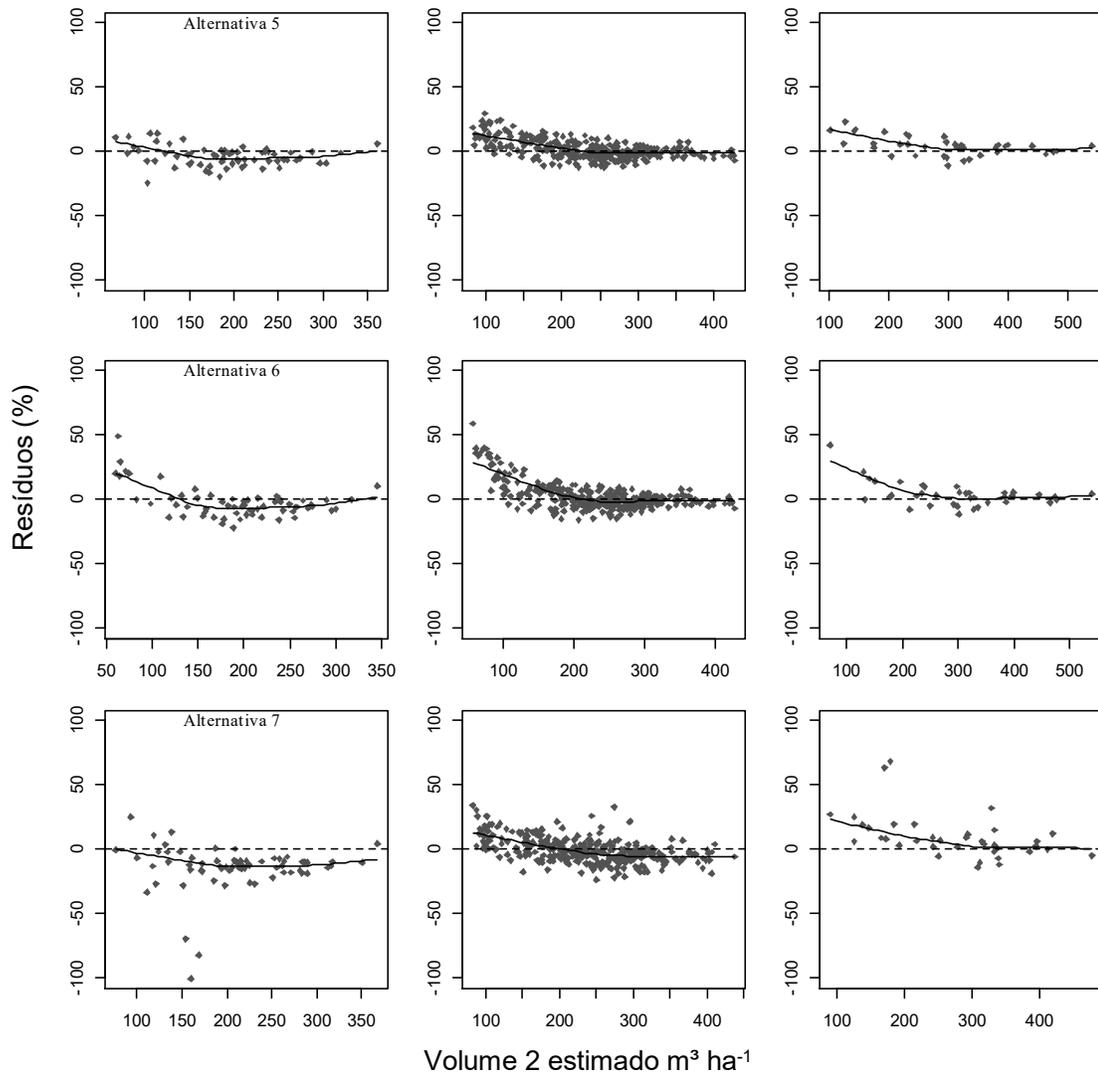


Figura 12 – Distribuição dos resíduos percentuais em função dos volumes estimados, em cada classe de produtividade pela aplicação dos modelos de crescimento em diâmetro e altura.





3.3. Critério 2 – Alternativas usando o identificador da árvore como grupos de efeitos aleatórios

De forma a completar o grupo de alternativas definidas na Tabela 3 foi realizada uma adaptação do modelo escolhido na seção 3.1, denominada alternativa 2_3, apresentada na Tabela 13. Além disso, a alternativa 2_2_0 foi escolhida na seção 3.1. Embora a modelagem adotada neste capítulo seja de árvore de individual, o ajuste desta alternativa utilizou o identificador da parcela como grupo de efeitos aleatórios. A decisão de utilizar a parcela como identificador de grupos de efeitos aleatórios ocorreu com base em testes preliminares, que indicaram que a variação entre grupos de parcela foi mais significativa quando comparada à variação entre árvores.

Tabela 13 – Definição das alternativas 2_2_0 e 2_3, complementando a Tabela 3.

Alternativa	Modelo
2_2_0	$Y_2 = Y_1 + \left(\beta_0 + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 IID3 + \beta_3 S \right) + \varepsilon, \text{ modelo fixo.}$
2_3	$Y_2 = Y_1 + \left(\beta_0 + (\beta_1 + b_1) \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + (\beta_2 + b_2) IID3 + \beta_3 S \right) + \varepsilon, \text{ modelo misto}$ adaptado do modelo 5, com a parcela como variável de agrupamento de efeitos aleatórios e inclinação na variável IID3.

em que: Y_2 = diâmetro (em cm) ou altura (m) em idade futura; Y_1 = diâmetro (cm) ou altura (m) em idade atual; I_2 = idade futura (meses); I_1 = idade atual (meses); S = índice de local; $IID3 = \frac{DAP^2}{D^2} \cdot \frac{HT}{H}$ = índice de competição independente de distância

(MARTINS, 2011); β_i = coeficientes do modelo; e ε = erro aleatório;

Os ajustes, validações e análises gráficas realizados para as equações ajustadas referentes às alternativas do grupo dois, com dados do critério 2 são apresentados nas seções adiante.

3.3.1. Ajuste das alternativas de modelagem para crescimento em diâmetro

No ajuste das alternativas de crescimento em diâmetro, observa-se Tabela 7 pelos resultados apresentados na Tabela 14 Tabela 4, que a alternativa 2_1 foi mais acurada em todas as classes de produtividade, entretanto, mais viesada também em todas as classes de produtividade. Vale ressaltar que a classe de produtividade Média (S = 27), na qual concentram-se a maior parte das parcelas é menos viesada, o que pode ser confirmado na Figura 13 e na Figura 14.

As alternativas 2_2 e 2_3 apresentaram comportamento semelhante no ajuste para o BIAS (%), embora a alternativa 2_2 apresente maior acurácia na classe de produtividade Média (S = 27).

Tabela 14 – Estatísticas utilizadas para avaliar o ajuste das alternativas para estimar o crescimento de diâmetro na altura do peito (DAP), por classe de produtividade.

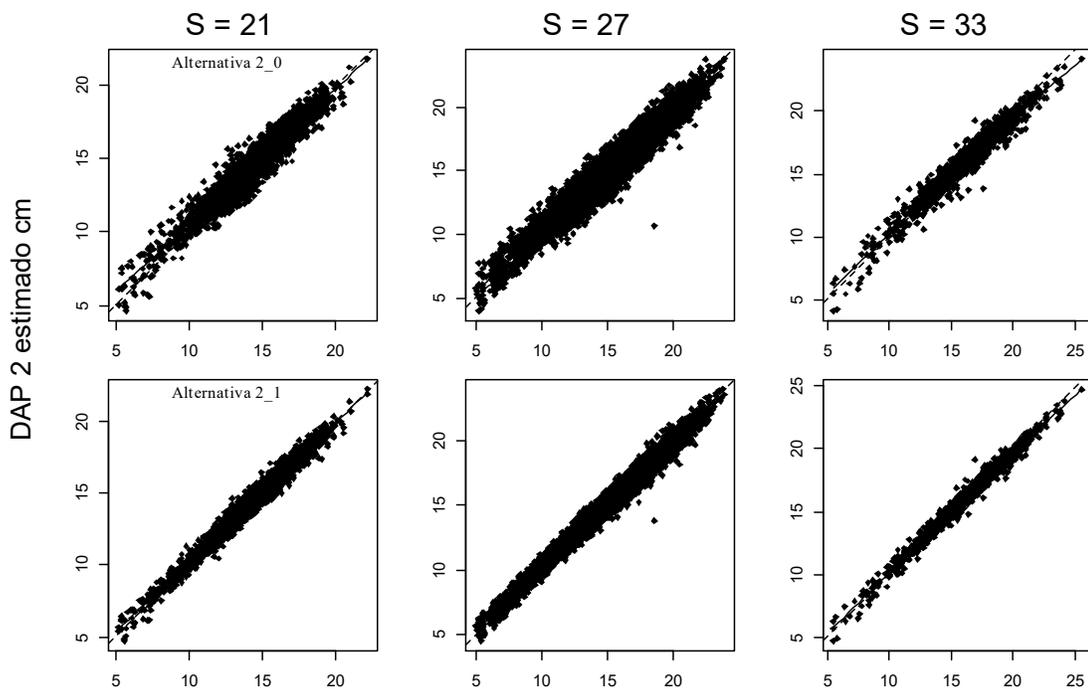
Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
2_0	S = 21	0,9642	4,48	1,0700
	S = 27	0,9672	4,19	-0,1404
	S = 33	0,9737	4,33	-0,6671
2_1	S = 21	0,9868	2,71	0,5038
	S = 27	0,9885	2,50	-0,0602
	S = 33	0,9909	2,58	-0,4042
2_2_0	S = 21	0,9235	19,89	18,0900
	S = 27	0,9273	16,84	15,0400
	S = 33	0,9420	14,02	12,1400
2_2	S = 21	0,9805	3,23	0,0328
	S = 27	0,9825	3,06	-0,0041
	S = 33	0,9844	3,28	-0,0232
2_3	S = 21	0,9741	3,75	0,0176
	S = 27	0,9783	3,43	-0,0061
	S = 33	0,9842	3,29	0,0461

Apesar da maior acurácia da alternativa 2_1 no ajuste, ao avaliar as estatísticas de validação (Tabela 15), esta alternativa foi menos acurada e mais viesada. Evidencia-se a superioridade da alternativa 2_3 (estratégia 1), mais acurada e menos viesada em todas as classes de produtividade, com significativa diferença para as demais quando se leva em conta a classe de produtividade Média.

Tabela 15 – Estatísticas utilizadas para avaliar a aplicação das alternativas de crescimento em diâmetro (DAP) em uma base de dados de validação.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
2_0	S = 21	0,9631	5,99	-4,3850
	S = 27	0,9733	6,34	-5,0030
	S = 33	0,9834	7,48	-6,4920
2_1	S = 21	0,9744	4,80	-3,4090
	S = 27	0,9809	5,07	-3,8840
	S = 33	0,9879	5,30	-5,1190
2_2_0	S = 21	0,9591	126,46	121,82
	S = 27	0,9611	136,05	130,23
	S = 33	0,9641	149,95	141,42
2_2	S = 21	0,9855	2,92	-0,6426
	S = 27	0,9870	3,01	-0,5238
	S = 33	0,9918	3,33	-1,8250
2_3	S = 21	0,9874	2,44	-0,3161
	S = 27	0,9901	2,37	-0,1314
	S = 33	0,9937	2,29	-0,1279

Figura 13 – Diâmetros na altura do peito estimados *versus* observados pelas equações ajustadas para o conjunto de alternativas dois com o conjunto de dados denominado critério 2.



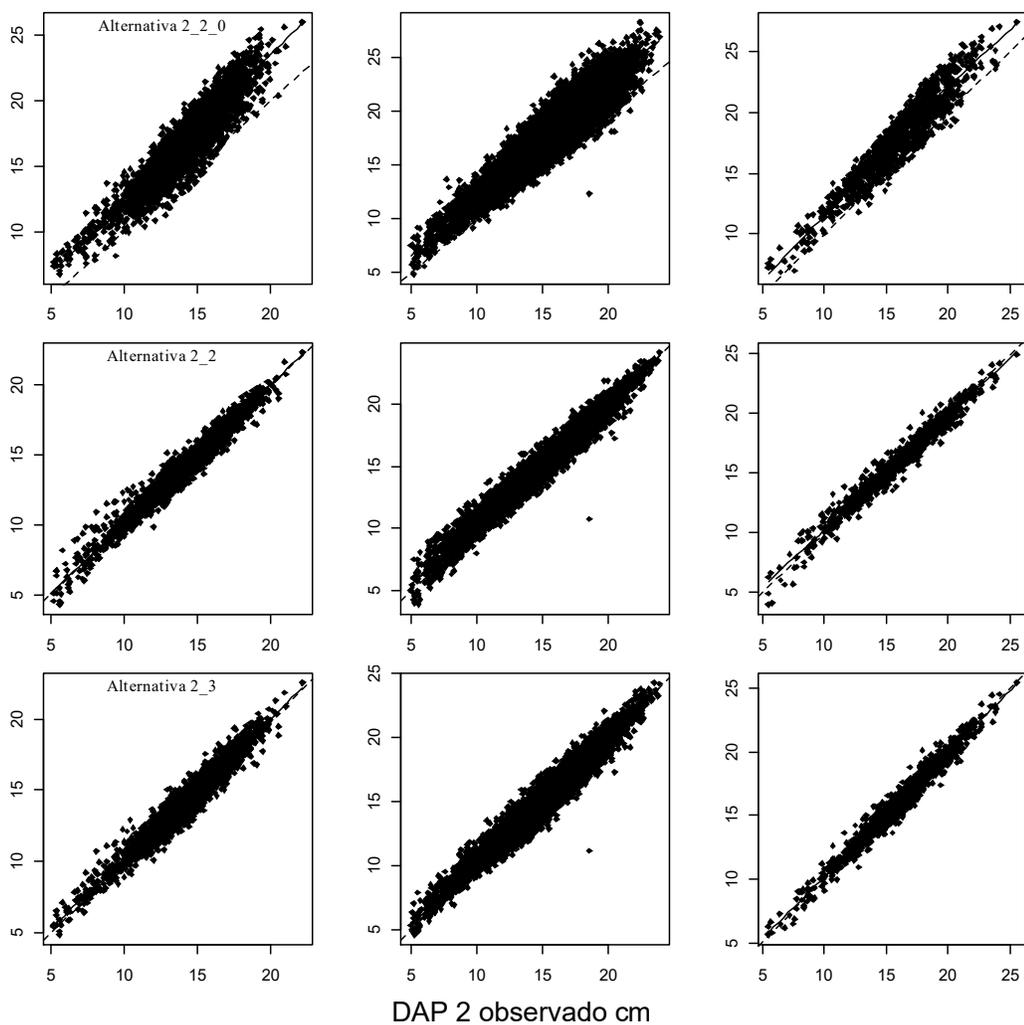
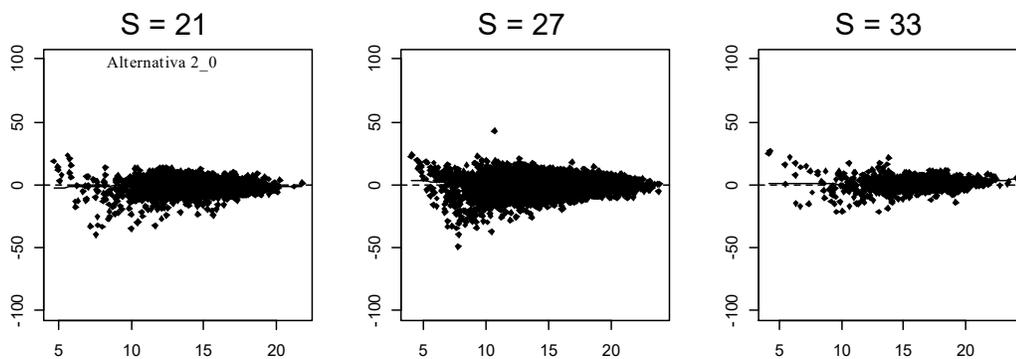
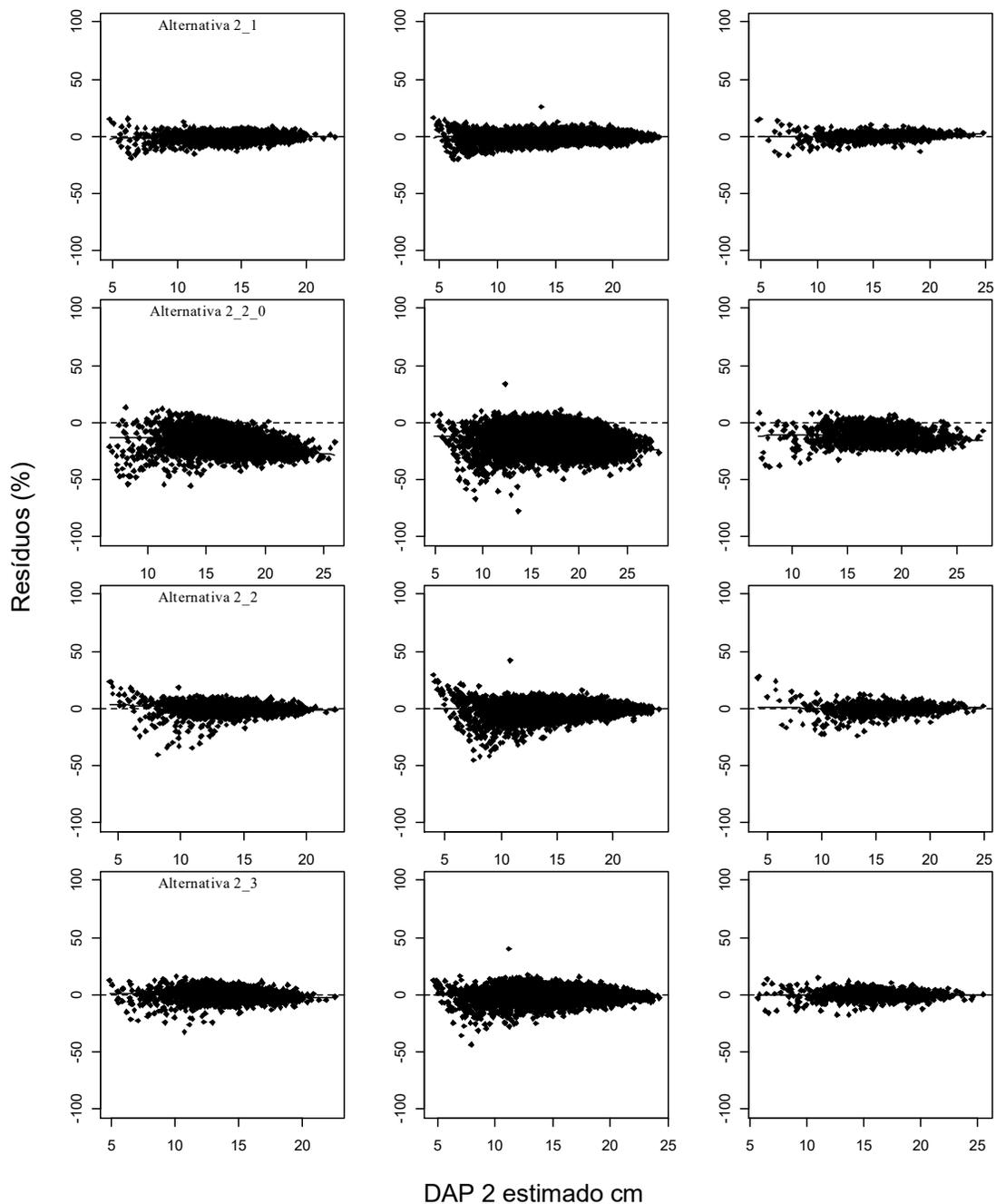


Figura 14 – Distribuição dos resíduos porcentuais em função dos diâmetros (DAP) estimados, em cada classe de produtividade para os modelos de crescimento em diâmetro do conjunto dois.





Nos modelos de crescimento em diâmetro avaliados no segundo conjunto de alternativas que usam os dados do critério 2, após a avaliação das estatísticas e da análise gráfica dos resíduos, é possível afirmar que a alternativa 2_3 foi mais acurada e menos viesada em todas as classes de produtividade.

Comparando-se os resultados com os encontrados no critério 1, verifica-se que a calibração de modelos de crescimento em diâmetro para parcelas e árvores individuais são mais acuradas e menos viesadas.

3.3.2. Ajuste das alternativas de modelagem para crescimento em altura

De acordo com o apresentado na Tabela 16, as alternativas propostas utilizando modelagem mista apresentam acurácia similar aos modelos de crescimento em altura não lineares nas alternativas 2_1 e 2_3. A alternativa 2_2 apresenta-se como mais acurada e menos tendenciosa, com RMSE (%) de 3.9760 na classe de produtividade Média (S = 27) e BIAS (%) de -0,0074 e a dispersão constante dos pontos em torno da linha 0:1 na Figura 15. Destaca-se a correlação superior em todas as classes de produtividade.

Nas classes de produtividade Baixa (S = 21) e Alta (S = 33), a alternativa 2_1 apresenta um viés pronunciado, comparado com demais alternativas, fato constatado ao observarmos os valores de RMSE (%) e BIAS (%) para esta alternativa. Na validação (Tabela 17), houve um aumento significativo da tendência de superestimação na classe Média (S = 27).

Os resíduos percentuais referentes às projeções de altura (Figura 16) demonstram uma leve tendência de subestimar a altura de árvores maiores na alternativa 2_2 para a classe de produtividade Média (S = 27). Evidencia-se na mesma classe que a alternativa 2_3 apresenta os resíduos agrupados em $\pm 20\%$, enquanto que nas alternativas 2_1 e 2_2 observa-se uma tendência maior de subestimação.

Tabela 16 – Estatísticas utilizadas para avaliar o ajuste das alternativas para estimar o crescimento em altura (HT), por classe de produtividade.

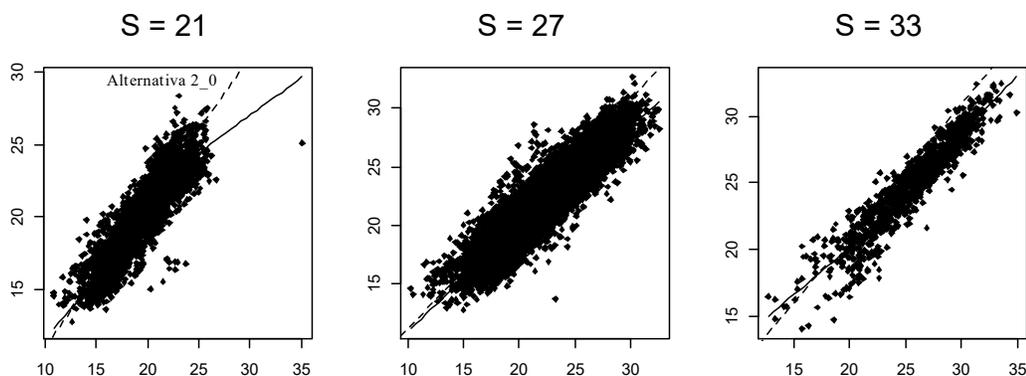
Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
2_0	S = 21	0,8880	7,54	0,0000
	S = 27	0,9254	5,46	3,3800
	S = 33	0,9418	5,51	-0,3460
2_1	S = 21	0,8958	7,23	3,1800
	S = 27	0,9314	5,23	-0,3232
	S = 33	0,9464	5,26	-2,6360
2_2_0	S = 21	0,8925	11,08	5,1280
	S = 27	0,9145	7,90	1,6620
	S = 33	0,9402	6,55	-1,2130
2_2	S = 21	0,9390	4,96	0,0674
	S = 27	0,9608	3,98	-0,0074
	S = 33	0,9696	3,47	-0,0480
2_3	S = 21	0,9074	6,37	0,0459
	S = 27	0,9424	4,93	-0,0148
	S = 33	0,9640	3,80	0,1032

Na validação das demais alternativas, pôde-se constatar que o comportamento apresentado no ajuste foi semelhante, embora a classe de produtividade Média (S = 27) da alternativa 2_2 tenha sido menos acurada. A alternativa 2_3 apresentou menor variação da correlação em todas as classes de produtividade, e apesar de mais acurada na classe de produtividade Média, apresenta maior viés comparando-se com a mesma classe da alternativa 2_2.

Tabela 17 – Estatísticas utilizadas para avaliar a aplicação das alternativas de crescimento em altura total (HT) em uma base de dados de validação.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
2_0	S = 21	0,8472	6,74	-3,4650
	S = 27	0,8868	7,52	-5,4150
	S = 33	0,8973	8,20	-6,1280
2_1	S = 21	0,8557	6,59	-3,3430
	S = 27	0,8895	7,37	-5,2590
	S = 33	0,9005	7,95	-5,8720
2_2_0	S = 21	0,8093	87,78	84,3930
	S = 27	0,8404	88,11	84,1840
	S = 33	0,9072	92,24	87,1370
2_2	S = 21	0,9019	4,99	1,8430
	S = 27	0,8762	5,36	-0,0597
	S = 33	0,9459	4,04	0,9131
2_3	S = 21	0,9170	4,33	-0,1048
	S = 27	0,9179	4,45	0,2750
	S = 33	0,9515	4,50	2,5550

Figura 15 – Alturas totais estimadas *versus* observadas pelas equações ajustadas referentes aos sete modelos de crescimento em altura total, em cada classe de produtividade.



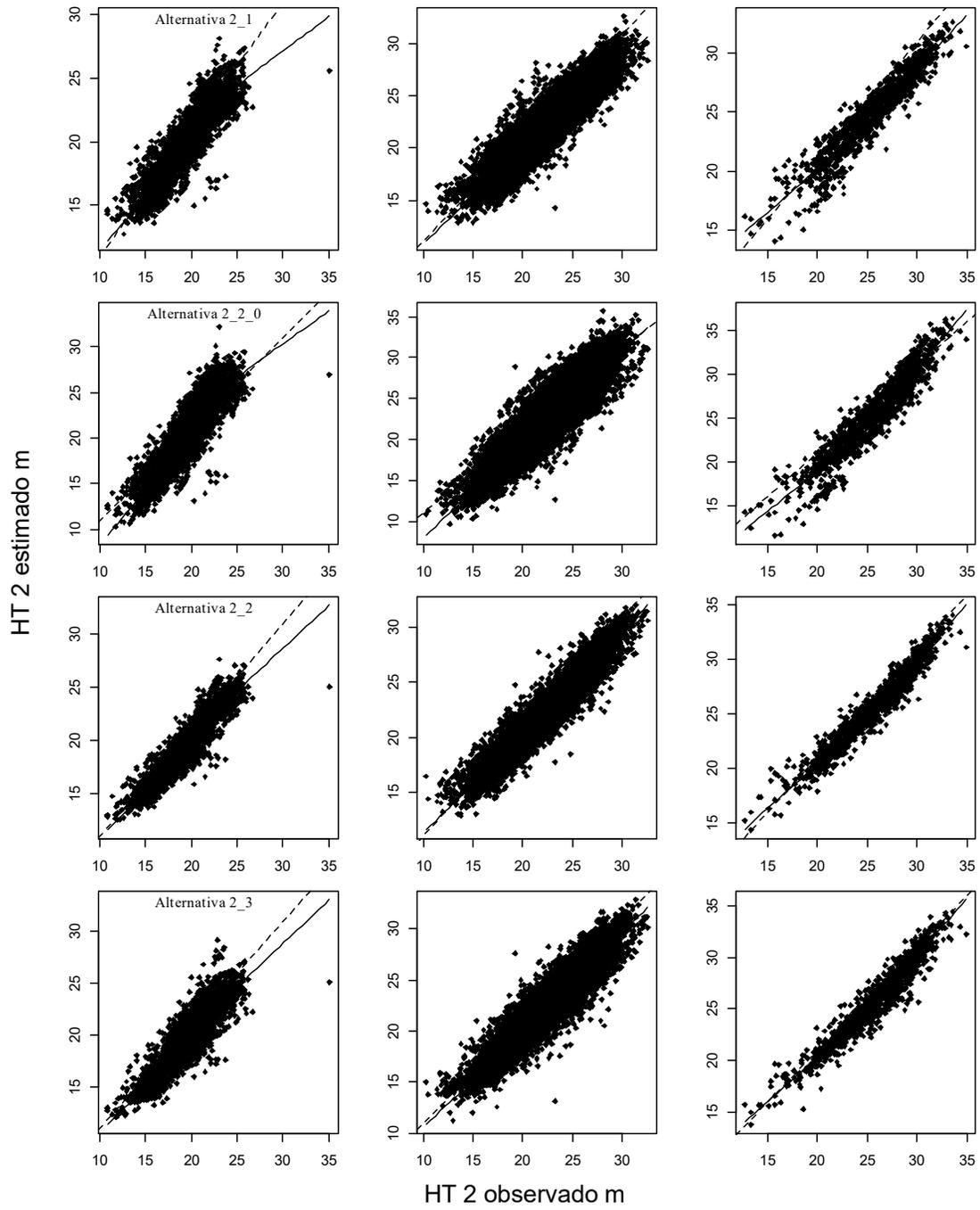
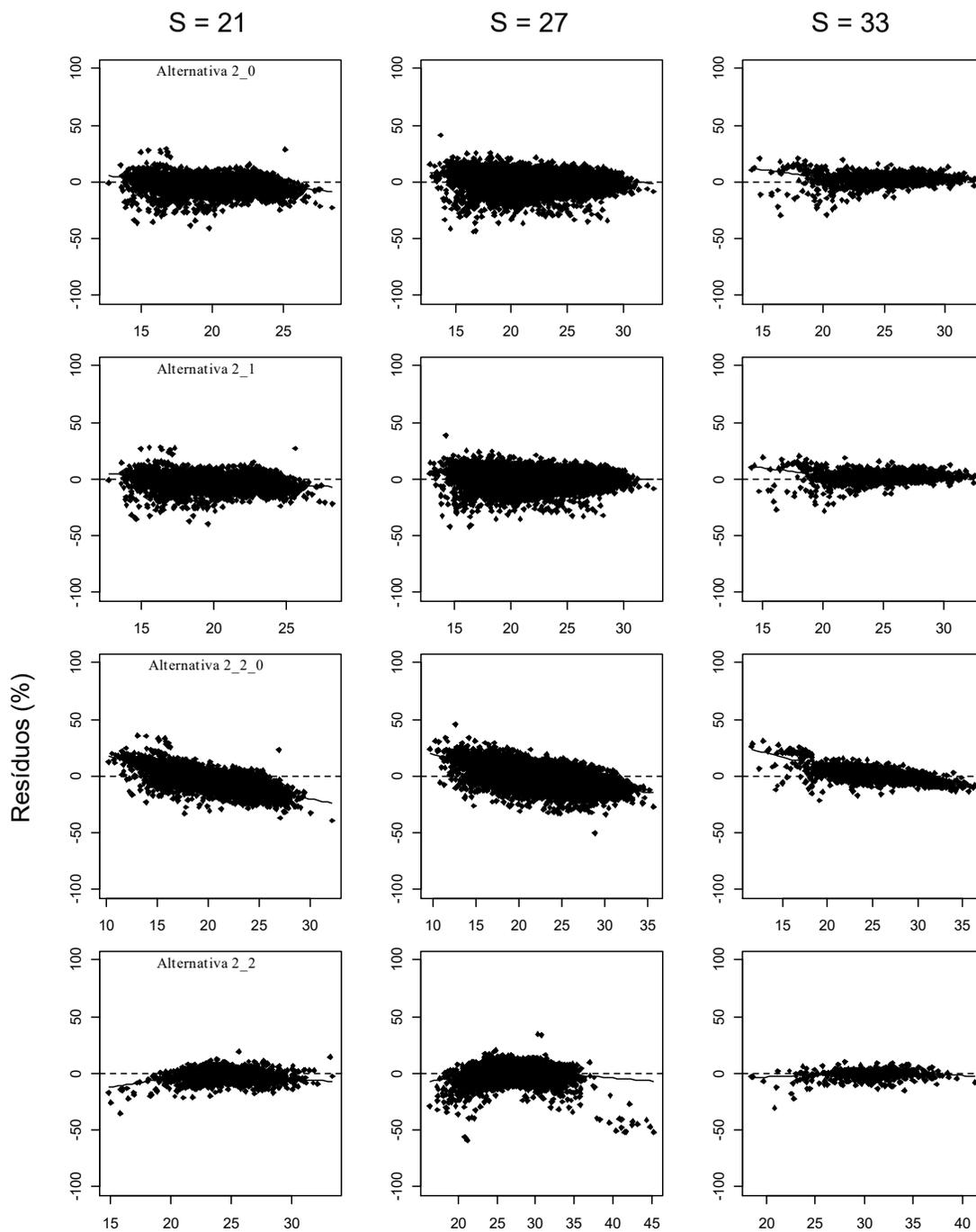
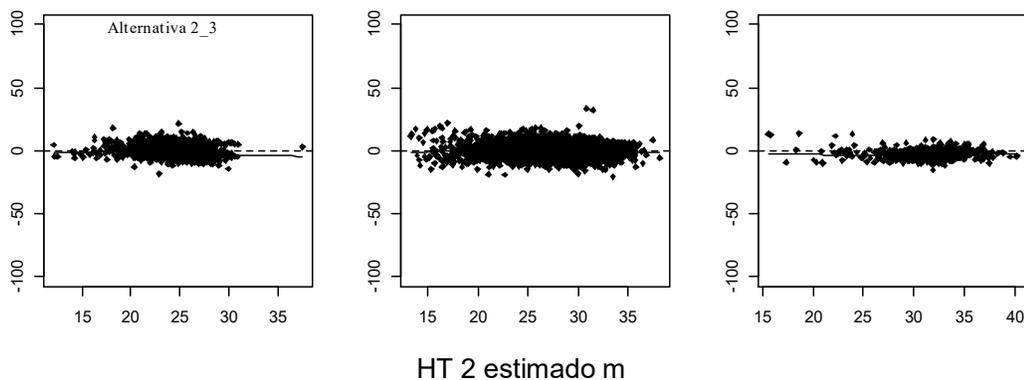


Figura 16 – Distribuição dos resíduos percentuais em função das alturas estimadas, em cada classe de produtividade para os modelos de crescimento em altura total.





Verificou-se no ajuste que a alternativa 2_2 apresentou menor viés e foi mais acurada, porém, pela análise gráfica de resíduos, a alternativa 2_3 apresentou menos dispersão residual em todas as classes de produtividade, o que pode ser confirmado pelas estatísticas de validação, embora na classe Alta ($S = 33$), a alternativa 2_2 tenha sido mais acurada e menos viesada.

3.3.3. Resultados de volume do modelo de árvore individual

Duas novas alternativas foram acrescentadas, a 2_4, resultantes da combinação das alternativas 2_2 para diâmetro (estratégia 2) e 2_3 para altura (estratégia 1); e a alternativa 2_5, resultado da combinação das alternativas 2_3 (estratégia 1) para diâmetro e 2_2 para altura (estratégia 2).

Pela análise das estatísticas apresentadas na Tabela 18 e observando a Figura 17 com volumes observados *versus* estimados, a alternativa 2_2 apresenta-se como a mais acurada com RMSE (%) de 3.37 na classe de produtividade Média ($S=27$), a mais representativa em número de parcelas, embora não seja a menos viesada. O menor BIAS (%) na mesma classe é produzido pela alternativa 2_5, com o valor de 0.02.

A alternativa 2_2_0 se manteve como a mais viesada, como ocorreu para o diâmetro a 1,3m de altura e altura total, o que é evidenciado tanto na análise das estatísticas de ajuste (Tabela 18) quanto na validação (Tabela 19). Além disso, ao observar o RMSE (%) na Tabela 19 para esta alternativa, nota-se uma piora significativa da acurácia, o que também pode ser observado na Figura 19.

Tabela 18 – Estatísticas utilizadas para avaliar o volume obtido a partir das alternativas de crescimento em diâmetro e altura, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
2_0	S = 21	0,9557	10,66	5,4257
	S = 27	0,9581	8,10	-0,8783
	S = 33	0,9678	8,70	-4,9032
2_1	S = 21	0,9738	8,30	4,1940
	S = 27	0,9751	6,27	-0,7431
	S = 33	0,9833	6,76	-4,1812
2_2_0	S = 21	0,9082	55,80	46,30
	S = 27	0,8804	43,16	34,44
	S = 33	0,8961	31,35	23,87
2_2	S = 21	0,9912	4,18	0,1569
	S = 27	0,9928	3,37	-0,2148
	S = 33	0,9975	1,98	-0,5085
2_3	S = 21	0,9681	8,43	0,4518
	S = 27	0,9759	6,56	0,1225
	S = 33	0,9889	4,30	0,0805
2_4	S = 21	0,9843	5,65	0,1984
	S = 27	0,9878	4,51	-0,1083
	S = 33	0,9952	2,75	-0,3040
2_5	S = 21	0,9820	6,26	0,3350
	S = 27	0,9860	4,81	0,0229
	S = 33	0,9938	3,02	-0,1455

Ao avaliar as estatísticas obtidas para a validação, observa-se que na classe de produtividade Média (S = 27), a alternativa 2_3 é a menos viesada e mais acurada. Entretanto, este não foi seu comportamento no ajuste. Embora a alternativa 2_3 apresente RMSE (%) ligeiramente superior à alternativa 2_5 na classe de produtividade Baixa (S = 21), ela se apresenta como menos viesada.

A alternativa 2_5 apresenta-se mais acurada e menos viesada e com maior correlação dos valores observados e estimados na classe de produtividade Alta (S = 33), com uma leve tendência de superestimação.

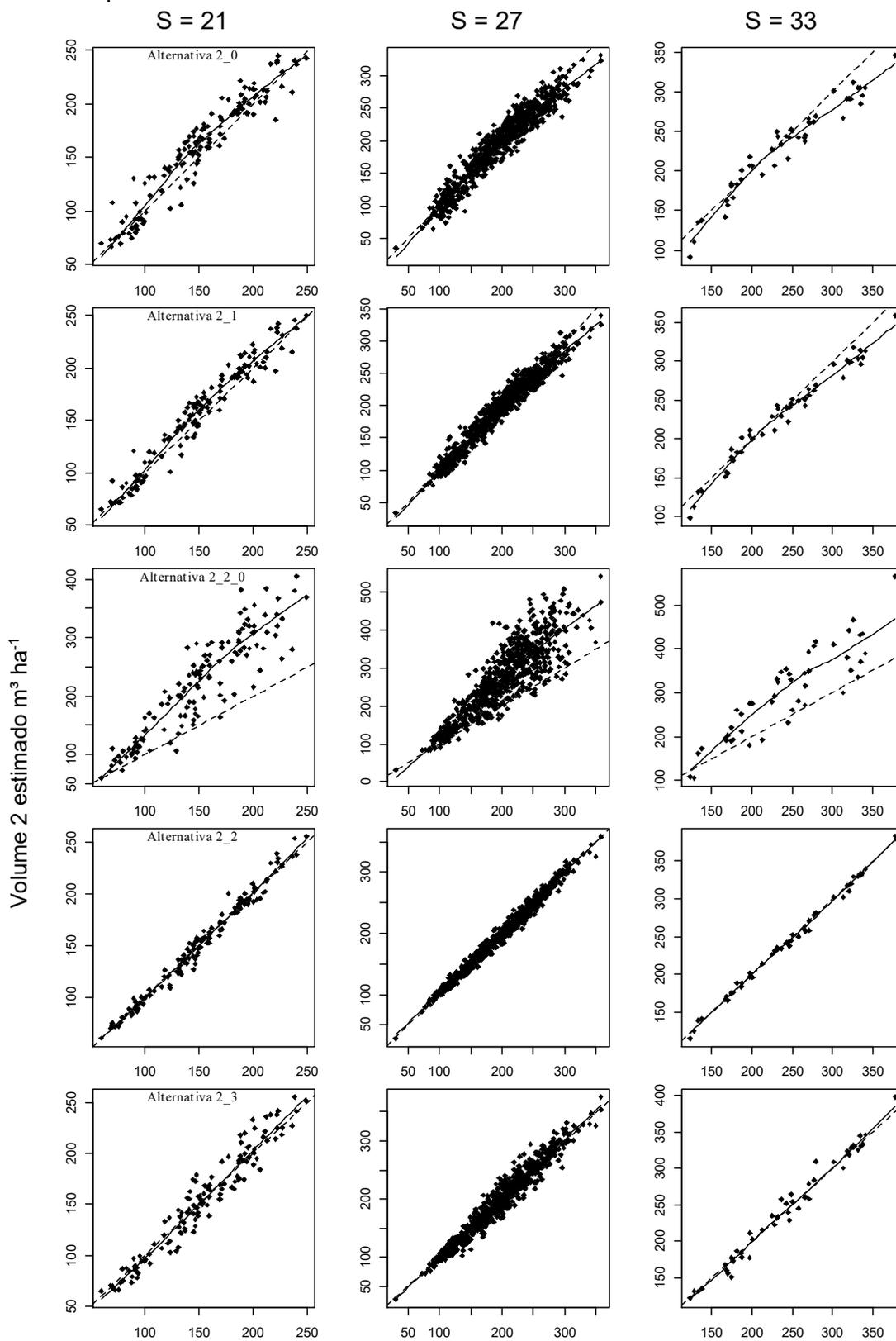
Tabela 19 – Estatísticas utilizadas para avaliar o volume obtido a partir da aplicação das alternativas de crescimento em diâmetro e altura no critério 2, em uma base de dados de validação, por classe de produtividade.

Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
2_0	S = 21	0,7883	16,46	-11,6225
	S = 27	0,8654	17,49	-14,2105
	S = 33	0,9330	20,52	-17,2601
2_1	S = 21	0,8518	14,11	-9,8251
	S = 27	0,8934	15,44	-12,4947
	S = 33	0,9617	17,96	-15,0042
2_2_0	S = 21	0,9045	849,99	813,89
	S = 27	0,9104	938,95	891,02
	S = 33	0,9409	1091,71	1030,04
2_2	S = 21	0,9441	6,25	1,3929
	S = 27	0,9363	6,41	-0,7105
	S = 33	0,9869	4,79	-2,2353
2_3	S = 21	0,9600	4,95	-0,3781
	S = 27	0,9648	4,72	0,2206
	S = 33	0,9877	4,11	2,3146
2_4	S = 21	0,9447	5,87	-0,6180
	S = 27	0,9570	5,25	-0,2249
	S = 33	0,9820	4,33	-0,4914
2_5	S = 21	0,9658	4,93	1,6222
	S = 27	0,9450	5,88	-0,2626
	S = 33	0,9921	3,29	0,5257

Observa-se que os volumes estimados e observados (Figura 17) se distribuem próximos à linha de tendência, com destaque para as alternativas 2_2 e 2_5 em todas as classes de produtividade, com maior proximidade, indicando maior acurácia no ajuste. Entretanto, na validação (Figura 19) a alternativa 2_2 apresenta uma ligeira tendência de superestimação para maiores volumes na classe de produtividade Alta (S = 33), e de subestimar para maiores volumes na classe de produtividade Baixa (S = 21), sendo verificado maior distanciamento dos pontos em relação a linha de tendência na classe de produtividade Média (S = 27).

Na validação da alternativa 2_5, o mesmo não ocorre, embora a distância média da linha de tendência tenha aumentado na classe de produtividade Média (S = 27). Corroborando estes resultados, a dispersão residual é menor nas alternativas 2_2, 2_4 e 2_5.

Figura 17 – Volumes estimados *versus* observados pelas equações ajustadas referentes aos sete modelos de crescimento em diâmetro e altura total, em cada classe de produtividade.



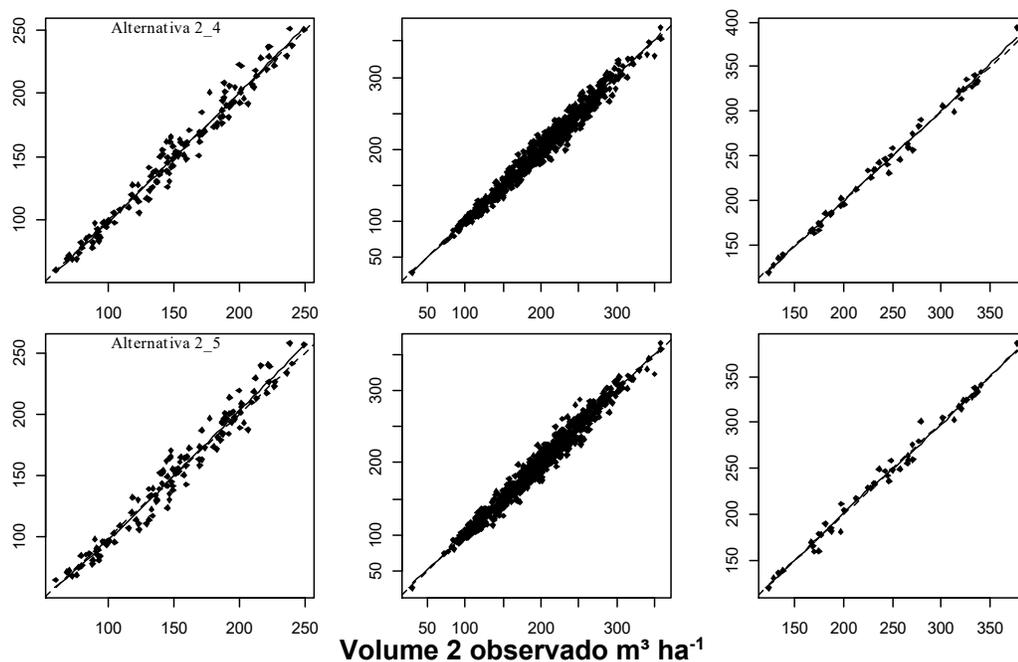
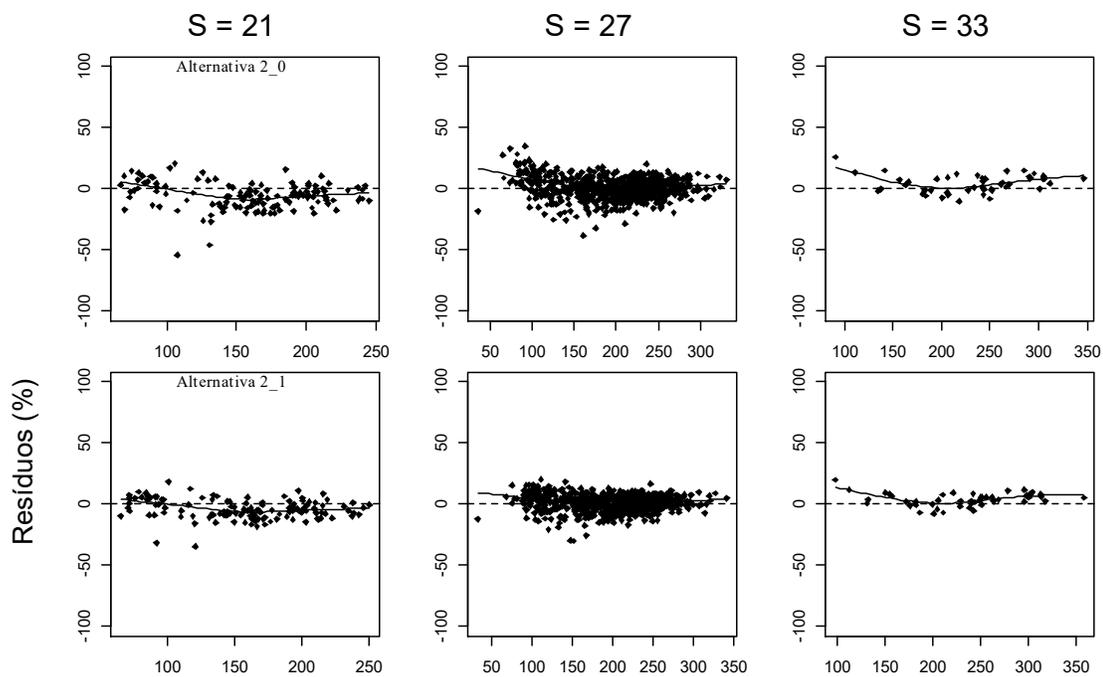
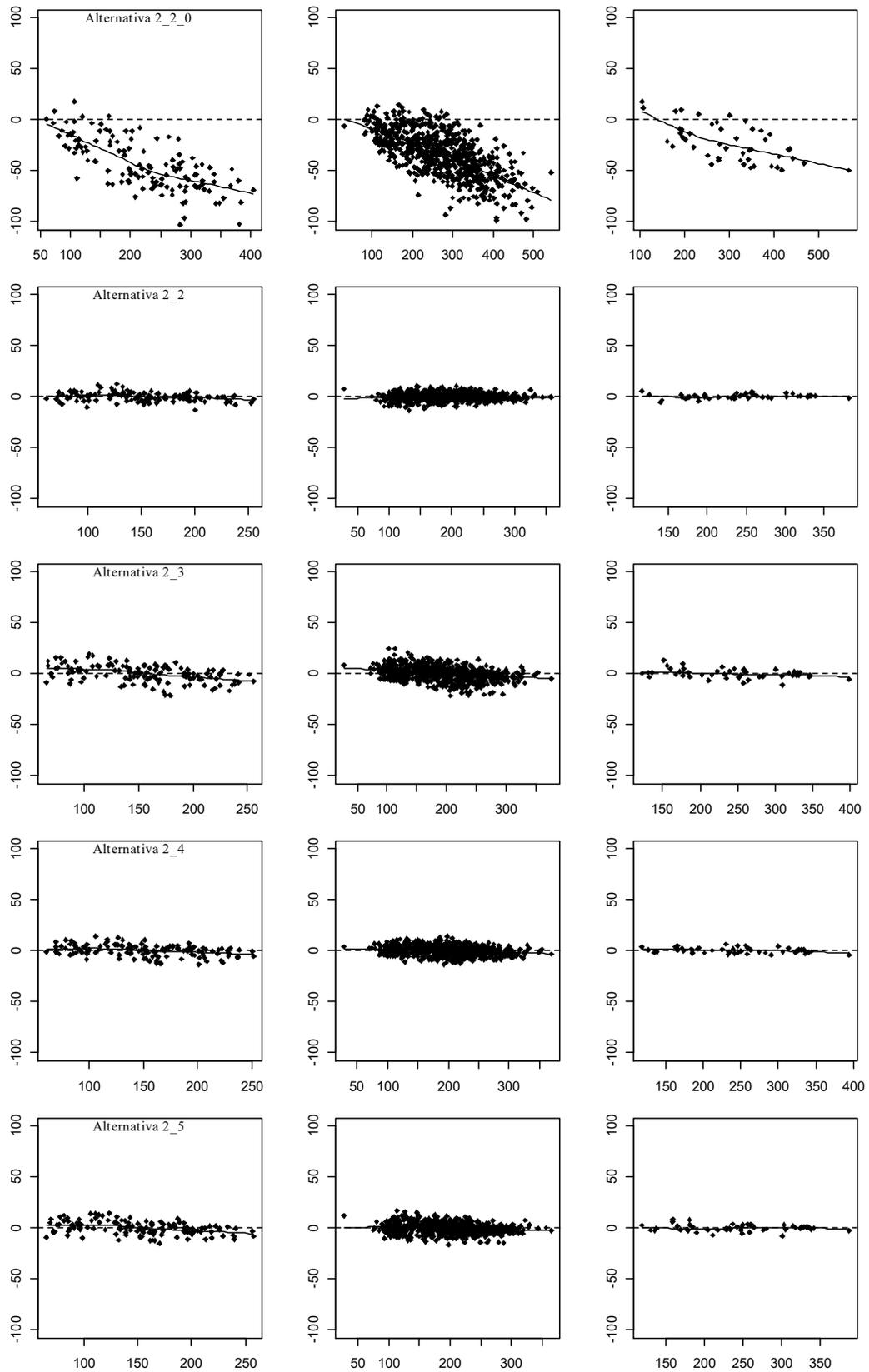


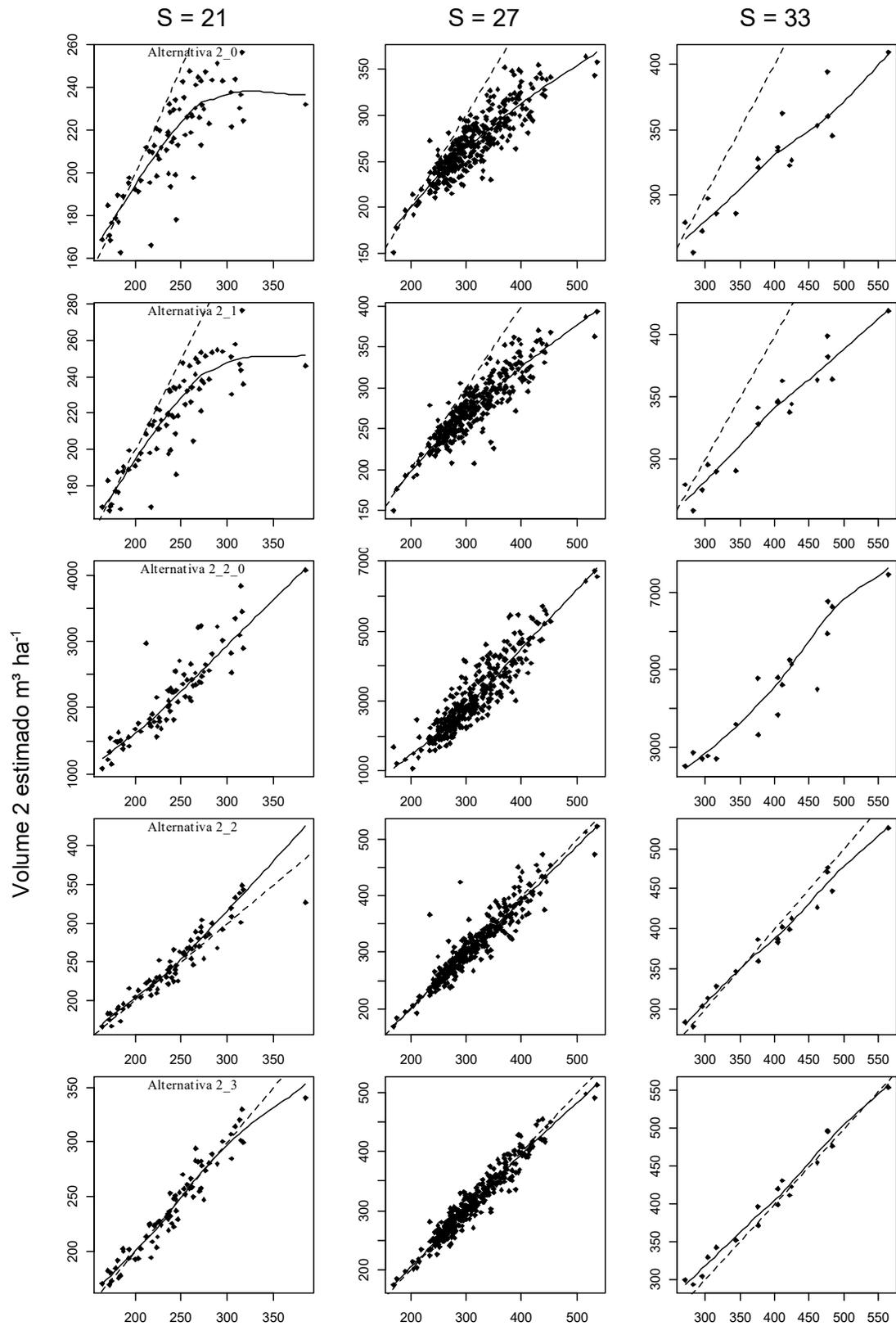
Figura 18 – Distribuição dos resíduos percentuais em função dos volumes totais estimados, em cada classe de produtividade para os modelos de crescimento em diâmetro e altura.





Volume 2 estimado $\text{m}^3 \text{ha}^{-1}$

Figura 19 – Volumes estimados *versus* observados pela aplicação das equações de crescimento em diâmetro e altura em uma base de dados independente do ajuste, em cada classe de produtividade.



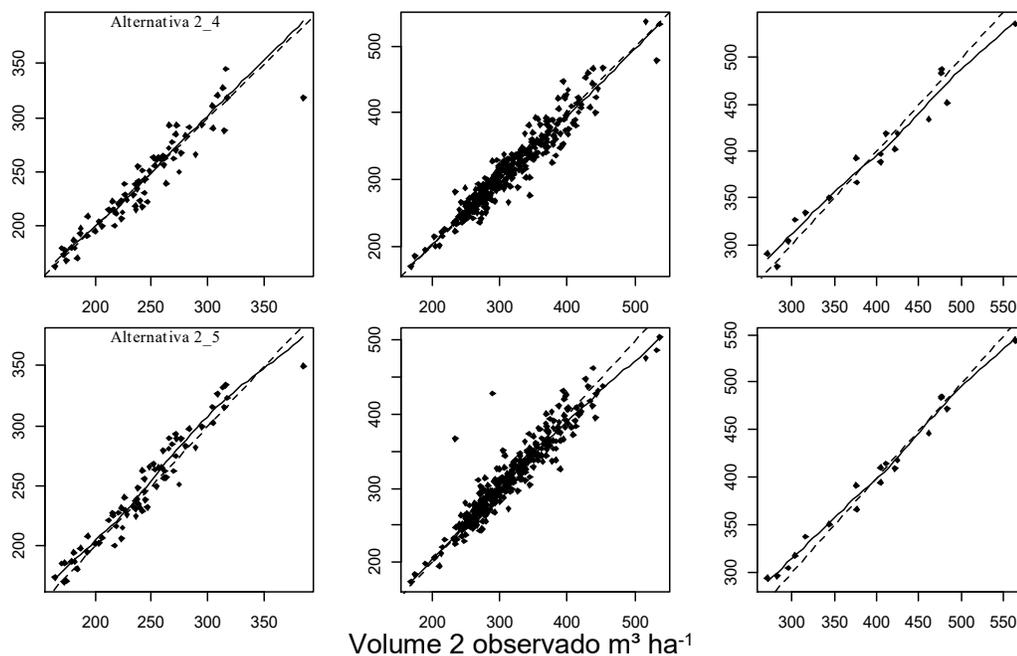
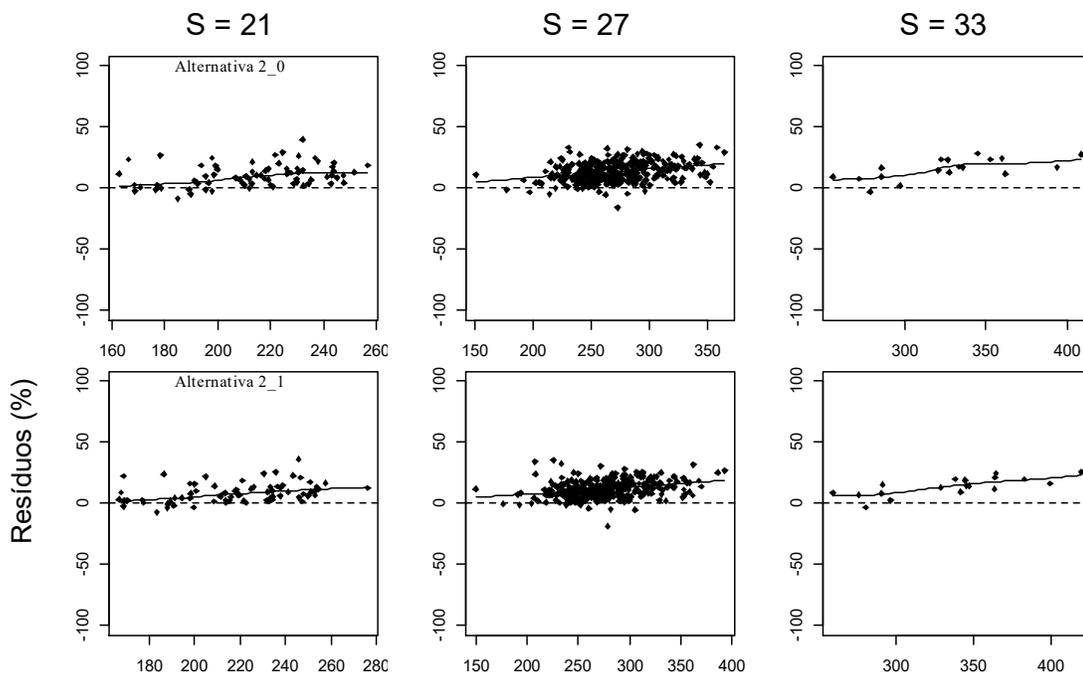
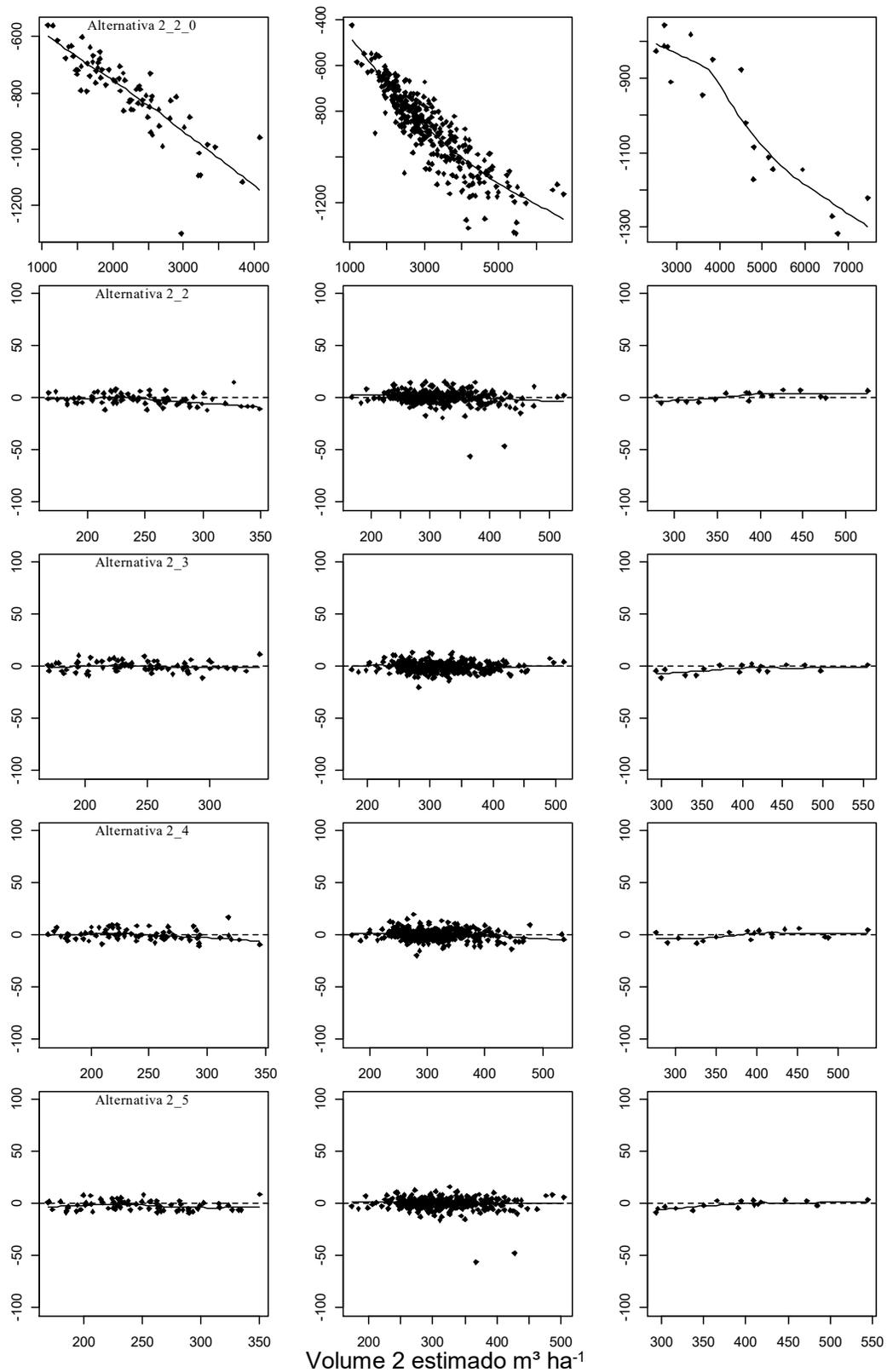


Figura 20 – Distribuição dos resíduos percentuais em função dos volumes estimados, em cada classe de produtividade para os modelos de crescimento em diâmetro e altura em uma base independente.





3.4. Comparação das alternativas ajustadas e validadas para os critérios definidos

Ao comparar-se a calibração realizada tendo como grupos de efeitos aleatórios o identificador da árvore ou da parcela (grupo de alternativas 2) à calibração utilizando outras variáveis disponíveis, verifica-se resultados semelhantes tanto em acurácia, quanto em tendência. Porém, para a classe de produtividade Baixa ($S = 21$), a alternativa 2_5 foi a mais acurada quando se leva em conta o RMSE (%) e a menos viesada.

Na classe de produtividade Média ($S = 27$), mais representativa em quantidade de parcelas, as alternativas 2 (critério 1), 2_3 e 2_5 (critério 2) foram acuradas e apresentaram baixa tendenciosidade. A alternativa 2_3 foi a mais acurada, com RMSE relativo de 4.72, seguido de 5.74 na alternativa 2 e 5.88 na alternativa 2_5. A alternativa 2_3 foi também a menos viesada, com BIAS relativo de 0.2206, leve tendência de superestimação. A alternativa 2_5 apresenta tendência contrária, mas igualmente leve com valor de -0.262. A mais viesada nesta classe foi a alternativa 2, com BIAS (%) de 0.57, todos considerando a validação.

Para a classe de produtividade Alta ($S = 33$), verifica-se que a alternativa 5 é mais acurada e menos viesada que as demais na mesma classe, apresentando 3.28 de RMSE relativo, seguido de 4.11% na mesma estatística para a alternativa 2_3 e 5.96% na alternativa 2. O BIAS relativo é expressivamente inferior, sendo de 0.52 nesta alternativa, 2.31% na alternativa 2_3 e -2.15 na alternativa 2.

Ao validar os modelos de crescimento em diâmetro e altura, obtendo volumes totais para plantios clonais não desbastados para 33 parcelas, Martins (2011) encontrou RMSE relativo de 31.7% para a classe de produtividade Baixa, e BIAS (%) de 7.5%. Para a classe de produtividade Média, foi obtido em seu trabalho um RMSE (%) de 23.6%, com BIAS (%) 4.1%, seguido dos valores obtidos na classe de produtividade Alta, de 18.6% de RMSE (%) e 1.7% de BIAS relativo.

4. CONCLUSÕES

As estratégias lineares e não lineares deste estudo permitem projetar o diâmetro a 1,3 m de altura e altura total de modo eficiente., incluindo ou não efeitos aleatórios, com exceção da alternativa linear fixa (alternativa 2_0,).

Dentre os modelos fixos avaliados, a alternativa 1, ajustada com dados do critério 1 é superior à alternativa 2_0 (critério 2). Já para os modelos mistos, a alternativa 2_5 (critério 2) é mais vantajosa pela avaliação das estatísticas de ajuste, validação e análise gráfica dos resíduos quando comparada à alternativa 2 (critério 1).

As equações representadas por alternativas que utilizaram a modelagem mista são acuradas e menos viesadas, tanto no ajuste quanto na validação, em ambos os critérios de divisão de dados adotados.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ARIAS-RODIL, M.; CASTEDO-DORADO, F.; CÁMARA-OBREGÓN, A.; DIÉGUEZ-ARANDA, U. Fitting and calibrating a multilevel mixed-effects stem taper model for *Maritime Pine* in NW Spain. **PloS one**, v. 10, n. 12, p. e0143521, 2015. Public Library of Science. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143521>>. Acesso em: 31/12/2015.

BINOTI, M. L. M. DA S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, v. 37, n. 4, p. 639–645, 2013. Sociedade de Investigações Florestais. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000400007&lng=pt&nrm=iso&tlng=en>. Acesso em: 31/12/2015.

BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **CERNE**, v. 21, n. 1, p. 97–105, 2015. UFLA - Universidade Federal de Lavras. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-77602015000100097&lng=en&nrm=iso&tlng=pt>. Acesso em: 31/12/2015.

CALAMA, R.; MONTERO, G. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. **Silva Fennica**, v. 39, n. 1, p. 37–54, 2005.

CARVALHO, S. DE P. C. E.; CALEGARIO, N.; SILVA, F. F. E.; et al. Modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* clonal. **CERNE**, v. 17, n. 4, p. 541–548, 2011. Universidade Federal de Lavras.

CASTRO, R.; SOARES, C.; LEITE, H.; et al. Competição em Nível de Árvore Individual em uma Floresta Estacional Semidecidual. **Silva Lusitana**, v. 22, n. 1, p. 43–66, 2014. Unidade de Silvicultura e Produtos Florestais. Disponível em: <http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S0870-63522014000100003&lng=pt&nrm=iso&tlng=pt>. Acesso em: 4/10/2015.

CASTRO, R. V. O.; SOARES, C. P. B.; LEITE, H. G.; et al. Artificial Neural Networks Effectiveness to Estimate Mortality in a Semi-Deciduous Seasonal Forest. **Australian Journal of Basic and Applied Sciences**, v. 9, n. 5, p. 435–444, 2015.

CASTRO, R. V. O.; SOARES, C. P. B.; MARTINS, F. B.; LEITE, H. G. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos. **Pesquisa Agropecuária Brasileira**, v. 48, n. 3, p. 287–295, 2013. SciELO Brasil.

FRAGA FILHO, C. V.; SIMIQUELI, A. P. .; ALTOE, W. . Fgmutils: Forest Growth Model Utilities. Disponível em: <<https://cran.r-project.org/web/packages/Fgmutils/index.html>>. Acesso em: 31/12/2015.

GÖRGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. **Computers and Electronics in Agriculture**, v. 116, p. 221–227, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169915001994>>. Acesso em:

9/12/2015.

ISLAM, M. N.; KURTTILA, M.; MEHTÄTALO, L.; HAARA, A. Analyzing the effects of inventory errors on holding-level forest plans: the case of measurement error in the basal area of the dominated tree species. **Silva Fennica**, v. 43, n. 1, p. 71–85, 2009. Finnish Society of Forest Science.

MARTINS, F. B. **Modelagem de crescimento em nível de árvore individual para plantios comerciais de eucalipto**, 2011. 143f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2011.

MARTINS, F. B.; SOARES, C. P. B.; SILVA, G. F. DA. Individual tree growth models for eucalyptus in northern Brazil. **Scientia Agricola**, v. 71, n. 3, p. 212–225, 2014. SciELO Brasil.

MÔRA, R. **Funções de afilamento de forma variável e modelagem de efeitos mistos em fustes de *Pinus taeda* e *Eucalyptus saligna***, 2015. 277f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2015.

PIETILÄ, I.; KANGAS, A.; MÄKINEN, A.; MEHTÄTALO, L. Influence of growth prediction errors on the expected losses from forest decisions. **Silva Fennica**, v. 44, n. 5, p. 829–843, 2010.

UZOH, F. C. C. C.; OLIVER, W. W. Individual tree diameter increment model for managed even-aged stands of ponderosa pine throughout the western United States using a multilevel linear mixed effects model. **Forest ecology and management**, v. 256, n. 3, p. 438–445, 2008. Elsevier. Disponível em: <<http://www.treearch.fs.fed.us/pubs/40984#.VoV-xYIcjD4.mendeley>>. Acesso em: 31/12/2015.

VISMARA, E. DE S.; MEHTÄTALO, L.; BATISTA, J. L. F. Linear mixed-effects models and calibration applied to volume models in two rotations of *Eucalyptus grandis* plantations. **Canadian Journal of Forest Research**, p. 132–141, 2015. NRC Research Press. Disponível em: <<http://www.nrcresearchpress.com/doi/abs/10.1139/cjfr-2014-0435#.VoWA8-JYWug>>. Acesso em: 31/12/2015.

ZHAO, L.; LI, C.; TANG, S. Individual-tree diameter growth model for fir plantations based on multi-level linear mixed effects models across southeast China. **Journal of Forest Research**, v. 18, n. 4, p. 305–315, 2013.

CAPÍTULO 4 – MODELAGEM DE CRESCIMENTO E PRODUÇÃO EM DIFERENTES NÍVEIS DE ABORDAGEM PARA POVOAMENTOS DE EUCALIPTO, COM E SEM EFEITOS ALEATÓRIOS

1. INTRODUÇÃO

Em mensuração florestal, a modelagem de crescimento e produção consiste em construir, realizar o ajustamento e aplicar equações com o objetivo de prognose, para uso em uma série de atividades no setor florestal, tais como a simulação da dinâmica natural do povoamento, a previsão da produção a cada período, dentre outras (BINOTI et al., 2015).

Modelos de crescimento e produção são, provavelmente, o método mais eficiente e confiável para descrever a dinâmica de florestas e realizar a previsão da produção a longo prazo. Os resultados destes modelos são essenciais para o planejamento de atividades de manejo florestal, contribuindo na tomada de decisão. De acordo com Campos e Leite (2013), entre outras coisas, os modelos de crescimento e produção têm um papel importante no manejo florestal e elaboração de políticas florestais.

Existem na literatura inúmeras classificações para os modelos de crescimento e produção. Vanclay (1994) classifica esses modelos como: modelo de povoamento total ou em nível de povoamento, modelos de classe de tamanho ou distribuição de diâmetros e modelos de árvores individuais.

Modelos em nível de povoamento são aqueles que consideram a produção por unidade de área. Segundo Burkhart (1977), estes modelos são normalmente inflexíveis para avaliar multiproduto e grande número de tratamentos, embora eficaz quando o objetivo é limitado a produção total.

Os modelos de crescimento e produção em nível de árvore individual simulam o crescimento de cada árvore em diâmetro, altura e copa, decidindo se a árvore vive ou morre, calculando o crescimento e o volume, de forma individual (MARTINS et al., 2014). A partir dos volumes individuais calculados, obtém-se os volumes totais por parcela, por meio da agregação dos mesmos. São três os submodelos componentes deste nível de modelo: probabilidade de mortalidade, regeneração ou ingresso e crescimento dimensional (altura e diâmetro).

Embora representem um salto significativo, pode ser dispendioso controlar e manter modelos de árvores individuais dependendo do objetivo do negócio florestal,

uma vez que as medidas detalhadas necessárias para a implementação destes modelos são onerosas (KANKARE et al., 2015).

Modelos de distribuição de diâmetros não são tão detalhadas e caros como os modelos de árvore individual e nem inflexíveis como os modelo no nível de povoamento. Assim, tornam-se uma alternativa atraente para modelar o crescimento e a produção de florestas. Para gerar esses modelos, é necessário o uso de funções de densidade de probabilidade (fdp), que caracterizam a distribuição de diâmetro da floresta. Vários estudos na literatura podem ser encontradas avaliando métodos científicos para aplicar este tipo de modelo, com diferentes distribuições de probabilidade, métodos de ajuste e recuperação das distribuições. Alguns exemplos de aplicação desses modelos podem ser encontrados em Bailey e Dell (1973), Leite et al. (1990), Cunha Neto et al (1994), Guimarães et al. (1994), Nogueira et al (2005), Soares et al (2007), Schneider et al. (2008), Oliveira Neto et al (2010), Soares et al. (2010a), Soares et al (2010b), Binoti et al (2011) e Santos (2012).

Como a principal vantagem destes modelos, tem-se a possibilidade de detalhamento das estimativas de produção, o que pode facilitar o planejamento florestal, tais como a produção de floresta multiproduto.

Tendo disponíveis os resultados dos modelos de em nível de árvore individual, é possível obter informações obtidas por modelos de distribuição de diâmetros, como também para modelos em nível de povoamento. A compatibilidade dos resultados de modelos de crescimento em nível de árvore individual para modelos de distribuição de diâmetros ocorre quando as árvores dentro das parcelas são classificadas de acordo com classes de diâmetro. Com o número de árvores em cada classe diamétrica, é possível obter a produção em cada classe de diâmetro, dentre outras informações necessárias, obtidas pelos modelos de distribuição de diâmetros.

O modelo de árvore individual permite ainda a obtenção de resultados dos modelos em nível de povoamento, bastando ao usuário totalizar os volumes obtidos em cada parcela.

Como alternativa aos modelos de regressão empregados na modelagem de distribuição de diâmetros, o uso de modelos mistos permite identificar as diferentes fontes de variabilidade estocástico que não são explicadas pela parte fixa do modelo dividindo-se a variância residual em diferentes componentes (CALAMA; MONTERO, 2005). Neste tipo de modelo, as medidas dos indivíduos não precisam ser igualmente espaçadas e balanceadas e as análises podem ser conduzidas com os dados de indivíduos que tiveram medidas perdidas (parcelas que não foram medidas, árvores

que não foram medidas dentro de parcelas, por exemplo) ou que apresentam ausência de informação em algum momento do estudo. Portanto, modelos mistos oferecem mais uma vantagem de utilização frente aos outros tipos de modelos para dados longitudinais e de medidas repetidas (UZOH; OLIVER, 2008).

Na modelagem de efeitos aleatórios, os parâmetros de um modelo ajustado podem variar entre os grupos, sendo assim, o modelo pode ser calibrado para um determinado grupo afim melhorar a qualidade das previsões dentro do referido grupo (VISMARA et al., 2015).

Diante do que fora discutido, o objetivo deste capítulo foi avaliar e comparar os volumes totais obtidos por modelos de crescimento e produção em nível de povoamento, distribuição de diâmetros e árvore individual, ajustados e validados com e sem efeito aleatório, utilizando dois critérios de divisão de banco de dados.

2. MATERIAL E MÉTODOS

2.1. Descrição dos dados

Os dados utilizados neste capítulo foram obtidos a partir dos dados descritos na seção 3.4 da tese.

A base de dados para cada nível de modelagem avaliado neste capítulo com as variáveis pareadas na idade 1 e idade 2 foi criada com a função *createPairedData* do pacote *Fgmutils* (FRAGA FILHO et al., 2015), criado durante o desenvolvimento deste trabalho e disponibilizado nos repositórios da linguagem R.

2.2. Divisão da base de dados para ajuste e validação dos modelos

Com o objetivo de avaliar diferentes grupos na calibração de modelos mistos, os dados foram divididos utilizando-se dois critérios: no primeiro (critério 1), o objetivo foi avaliar o comportamento dos modelos para idades futuras, definindo a variável parcela como grupo. No segundo critério (critério 2), foram definidos grupos formados por variáveis relacionadas ao povoamento, disponíveis na base de dados do inventário.

De acordo com o critério 1, os dados foram divididos em dois conjuntos: o primeiro para ajuste, e o segundo para validação dos modelos. Do conjunto de ajuste, participaram todas as parcelas com idade futura (idade 2) até 60 meses, e no segundo, de projeção (ou validação), foram selecionadas as mesmas parcelas do grupo de ajuste, porém com idade futura igual ou superior a 72 meses.

De acordo com o critério 2, o conjunto de parcelas foi dividido aleatoriamente também em dois grupos, sendo 70% das parcelas utilizadas para ajuste dos modelos e 30% para validação, ou seja, de acordo com este critério, as parcelas do conjunto de ajuste não estavam presentes no conjunto de validação.

2.3. Alternativas de modelagem sem efeito aleatório

Alternativas de modelagem que não utilizam efeito aleatório para crescimento e produção foram avaliadas nos níveis de povoamento, distribuição de diâmetros e árvore individual, de acordo com os critérios apresentados. Foram selecionadas as alternativas que resultaram em maior exatidão, conforme capítulos anteriores. O primeiro grupo de alternativas para os diferentes níveis de modelagem foi ajustado e validado para o conjunto de dados definidos no critério 1, e o segundo grupo de alternativas que também considerou os diferentes níveis foi ajustado e validado para o critério 2.

As alternativas avaliadas usando o conjunto de dados do critério 1 são apresentadas na Tabela 1.

Tabela 1 – Alternativas de modelagem fixa de crescimento e produção avaliadas em diferentes níveis, ajustadas e validadas com dados do critério 1.

Nível de modelagem	Alternativa	Modelo
Povoamento	7_0	$V_2 = \frac{\beta_{00} + \beta_{01}S}{1 + \exp\{[(\beta_{10} + \beta_{11}B_2) - l_2] / \beta_2\}} + \varepsilon_i$
Distribuição de diâmetros	1_0	$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) dmax_2 + \varepsilon$ $\gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$ $dmax_2 = dmax_1 \left(\frac{l_1}{l_2} \right) + \alpha_0 \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$
Árvore individual	2_0	<p>Modelos lineares fixos, para projeção de DAP e HT</p> $\widehat{DAP}_2 = \beta_0 + \beta_1 DAP_1 + \beta_2 l_2 + \beta_3 \bar{D}_1 + \beta_4 DMIN_1 + \beta_5 DMAX_1 + \varepsilon$ $\widehat{HT}_2 = \beta_0 + \beta_1 HT_1 + \beta_2 l_1 + \beta_3 IID3_1 + \beta_4 l_2 + \varepsilon$

em que: S é o índice de local; V_2 é o volume estimado na idade futura; B_2 é a área basal projetada para a idade futura ($m^2 ha^{-1}$); DAP_2 é o diâmetro estimado na idade futura; DAP_1 é o diâmetro observado na idade atual; \bar{D}_1 é o diâmetro médio observado na idade atual; $DMIN_1$ é o diâmetro mínimo observado na idade atual; $DMAX_1$ é o diâmetro máximo observado na idade atual; $IID3 = \frac{DAP^2}{\bar{D}^2} \cdot \frac{HT}{H}$ é o índice de competição independente de distância (MARTINS, 2011); γ_1 e γ_2 são os parâmetros de forma da função Weibull nas idades atual e futura, β_1 e β_2 são os parâmetros de escala da função Weibull nas idades atual e futura; $dmax_1$ e $dmax_2$ são os diâmetros máximos nas idades atual e futura, em cm; α_i , com $i=0$ a 1 são parâmetros dos modelos β são os parâmetros fixos do modelo, b parâmetros aleatórios do modelo misto; a_0 o coeficiente aleatório estimado para cada parcela;

As alternativas avaliadas para o conjunto de dados do critério 2 para os diferentes níveis de modelos de crescimento e produção são apresentadas na Tabela 2.

Tabela 2 – Alternativas de modelagem fixa de crescimento e produção avaliadas em diferentes níveis, ajustadas e validadas com dados do critério 2.

Nível de modelagem	Alternativa	Modelo
Povoamento	7	$V_2 = \frac{\beta_{00} + \beta_{01}S}{1 + \exp\{[(\beta_{10} + \beta_{11}B_2) - I_2] / \beta_2\}} + \varepsilon_i$
Distribuição de diâmetros	1	$\beta_2 = \beta_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) d \max_2 + \varepsilon$ $\gamma_2 = \gamma_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \varepsilon$ $d \max_2 = d \max_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \varepsilon$ $N_2 = N_1 \cdot e^{(-\alpha_0)(I_2^{a_1} - I_1^{a_1})} + \varepsilon$
Árvore individual	1	$Y_2 = Y_1 + \left(\beta_0 + \beta_1 \left(\frac{1}{I_2} - \frac{1}{I_1} \right) + \beta_2 BAI + \beta_3 S \right) + \varepsilon, \quad \text{Linear,}$ <p>adaptado de Bella (1971) e Campos e Leite (2009), (MARTINS et al., 2014)</p>

em que: S – índice de local; V_2 – Volume estimado na idade futura; B_2 – área basal projetada para a idade futura; DAP_2 – Diâmetro estimado na idade futura; DAP_1 – Diâmetro observado na idade atual; \bar{D}_1 – Diâmetro médio observado na idade atual; $DMIN_1$ – Diâmetro mínimo observado na idade atual; $DMAX_1$ – Diâmetro máximo observado na idade atual; $IID3 = \frac{DAP^2}{\bar{D}^2} \cdot \frac{HT}{H}$ = índice de competição independente de distância (MARTINS, 2011); Y_2 = diâmetro (em cm) ou altura (m) em idade futura; Y_1 = diâmetro (cm) ou altura (m) em idade atual; β - parâmetros fixos do modelo, b parâmetros aleatórios do modelo misto; a_0 o coeficiente aleatório estimado para cada parcela;

2.4. Alternativas de modelagem com efeito aleatório

Foram avaliadas alternativas de modelagem de crescimento e produção nos níveis de povoamento, distribuição de diâmetros e árvore individual, de acordo com os critérios apresentados. Sendo assim, o primeiro grupo de alternativas para os diferentes níveis de modelagem foi ajustado e validado para o conjunto de dados definidos no critério 1, e o segundo grupo de alternativas que também considerou os diferentes níveis foi ajustado e validado para o critério 2. Foram selecionadas as

alternativas mais acuradas e que apresentaram os melhores resultados na análise gráfica dos resíduos realizada nos três capítulos anteriores.

As alternativas avaliadas usando o conjunto de dados do critério 1 são apresentadas na Tabela 3.

Tabela 3 – Alternativas de modelagem mista de crescimento e produção avaliadas em diferentes níveis utilizando a parcela como grupo de efeitos aleatórios.

Nível de modelagem	Alternativa	Modelo
Povoamento	7_1	$V_2 = \frac{\beta_{00} + \beta_{01}S}{1 + \exp\{[(b_0 + \beta_{10}) + (b_1 + \beta_{11})B_2] - l_2\} / \beta_2} + \varepsilon_i$, com a variável parcela como variável de agrupamento de efeitos aleatórios e área basal na idade futura como inclinação aleatória
Distribuição de diâmetros	1_1	$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) d \max_2 + \varepsilon$ $\gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$ $d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$
Árvore individual	2_2	$DAP_2 = (\beta_0 + b_0) + \beta_1 DAP_1 + \beta_2 l_2 + (\beta_3 + b_3) \bar{D}_1 + \beta_4 DMIN_1 + \beta_5 DMAX_1 + \varepsilon$, com a parcela como variável de agrupamento de efeitos aleatórios, e diâmetro médio atual como inclinação aleatória. $\widehat{HT}_2 = (\beta_0 + b_0) + \beta_1 HT_1 + \beta_2 l_1 + (\beta_3 + b_3) IID3_1 + \beta_4 l_2 + \varepsilon$, com a parcela como variável de agrupamento de efeitos aleatórios, e IID3 na idade atual como inclinação aleatória.

em que: S – índice de local; V_2 – Volume estimado na idade futura; B_2 – área basal projetada para a idade futura; DAP_2 – Diâmetro estimado na idade futura; DAP_1 – Diâmetro observado na idade atual; \bar{D}_1 – Diâmetro médio observado na idade atual; $DMIN_1$ – Diâmetro mínimo observado na idade atual; $DMAX_1$ – Diâmetro máximo observado na idade atual; $IID3 = \frac{DAP^2}{\bar{D}^2} \cdot \frac{HT}{H}$ = índice de competição independente de distância (MARTINS, 2011); β - parâmetros fixos do modelo, b parâmetros aleatórios do modelo misto; a_0 o coeficiente aleatório estimado para cada parcela;

As alternativas avaliadas para o conjunto de dados do critério 2, não consideraram a parcela como grupo de efeitos aleatórios, mas outras variáveis do povoamento, ou variáveis categorias definidas com tal propósito. As alternativas para os diferentes níveis de modelos de crescimento e produção são apresentadas na Tabela 4.

Tabela 4 – Alternativas de modelagem mista para crescimento e produção avaliadas em diferentes níveis utilizando outras variáveis do povoamento como grupo de efeitos aleatórios.

Nível de modelagem	Alternativa	Modelo
Povoamento	11	$\text{Ln}V_2 = \beta_0 + b_0 + \frac{\beta_1}{l_2} + \beta_2 S + \beta_3 \text{Ln}B_2 + b_1 \hat{H}_2 + b_2 \ln(V1) + \varepsilon$, com a variável <i>classeldade1</i> como variável de agrupamento de efeitos aleatórios, altura média da parcela estimada na idade futura e log do volume m ³ ha ⁻¹ na idade atual como variáveis de inclinação aleatória.
Distribuição de diâmetros	2	$\beta_2 = \beta_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) d \max_2 + \varepsilon$ $\gamma_2 = \gamma_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$ $d \max_2 = d \max_1 \left(\frac{l_1}{l_2} \right) + (\alpha_0 + a_0) \left(1 - \frac{l_1}{l_2} \right) + \varepsilon$, utilizando como grupo de efeitos aleatórios a variável <i>classificacaoWeibull</i> , formada pela concatenação das variáveis <i>classeForma</i> e <i>classeEscala</i> , ambas formadas por classes de amplitude 1 dos parâmetros forma e escala estimados previamente para a função Weibull 2 parâmetros.
Árvore individual	5	$Y_2 = Y_1 + \left((\beta_0 + b_0) + \beta_1 \left(\frac{1}{l_2} - \frac{1}{l_1} \right) + \beta_2 BAI + \beta_3 S + \beta_4 \left(\frac{1}{l_2} - \frac{1}{l_1} \right) S \right) +$, modelo misto definido a partir do modelo Linear / adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014), tendo a variável ‘Chave’ como variável de agrupamento de efeitos aleatórios, e o índice de local como inclinação aleatória

em que: S – índice de local; V_2 – Volume estimado na idade futura; B_2 – área basal projetada para a idade futura; DAP_2 – Diâmetro estimado na idade futura; DAP_1 – Diâmetro

observado na idade atual; \bar{D}_1 – Diâmetro médio observado na idade atual; $DMIN_1$ – Diâmetro mínimo observado na idade atual; $DMAX_1$ – Diâmetro máximo observado na idade atual; $IID3 = \frac{DAP^2}{\bar{D}^2} \cdot \frac{HT}{H}$ = índice de competição independente de distância (MARTINS, 2011); β - parâmetros fixos do modelo, b parâmetros aleatórios do modelo misto; a_0 o coeficiente aleatório estimado para cada parcela;

2.5. Avaliação dos resultados

A avaliação das alternativas foi realizada pelas seguintes estatísticas:

a) Coeficiente de correlação, dado pela expressão (PIETILÄ et al., 2010; MÔRA, 2015):

$$r_{\hat{y}y} = \frac{n^{-1}(\sum(\hat{y}_i - \hat{y}_m)(y_i - \bar{y}))}{\sqrt{(n^{-1}\sum(\hat{y}_i - \hat{y}_m)^2)(n^{-1}\sum(y_i - \bar{y})^2)}} \quad (1)$$

b) BIAS(%) (UZOH; OLIVER, 2008):

$$BIAS = \frac{\sum_{i=1}^n (\hat{Y} - Y)}{n} \quad (2)$$

$$BIAS(\%) = \frac{BIAS}{\sum_{i=1}^n Y / n} \cdot 100 \quad (3)$$

c) Raiz quadrada do erro quadrático médio (RMSE %) (*Root Mean Square Error*), (ISLAM et al., 2009; ZHAO et al., 2013; ARIAS-RODIL et al., 2015):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n}} \quad (4)$$

$$RMSE(\%) = \frac{RMSE}{\sum_{i=1}^n Y / n} \cdot 100 \quad (5)$$

De forma a complementar a análise foram elaborados gráficos relacionando os volumes observados com os volumes estimados pelas quatro alternativas e gráficos de resíduos percentuais, obtidos pela expressão (MENDONÇA, 2010; MARTINS et al., 2014):

$$res(\%) = \left(\frac{y_i - \hat{y}_i}{y_i} \right) \cdot 100 \quad (6)$$

Toda a avaliação é realizada para cada classe de produtividade apresentada, como Baixa (S = 21), Média (S = 27), e Alta (S = 33). Na classe de produtividade média (S = 27) está a maior parte das parcelas utilizadas no ajuste (81% das parcelas), seguido pela classe de produtividade baixa (S = 21) com 14% das parcelas e a alta (S = 33) com 5% do total de parcelas.

3. RESULTADOS E DISCUSSÃO

Nesta seção são apresentados os resultados das alternativas fixas e mistas, ajustadas e validadas com dados divididos usando dois diferentes critérios e nos diferentes níveis de modelagem do crescimento e da produção.

3.1. Avaliação de modelos de crescimento e produção em diferentes níveis, utilizando modelagem fixa

Nesta seção são apresentados os resultados do ajuste e validação das alternativas fixas, tanto para o critério 1, quanto para o critério 2, avaliadas neste trabalho.

3.1.1. Volumes totais obtidos por equações ajustados e validados com o conjunto de dados do critério 1, com modelagem fixa

Os resultados apresentados nesta seção referem-se a alternativas fixas, ajustadas e validadas com dados divididos pelo critério 1, no qual para do conjunto de ajuste, participaram todas as parcelas com idade futura (idade 2) até 60 meses, e no segundo (de validação), foram selecionadas as mesmas parcelas do grupo de ajuste, porém com idade futura igual ou superior a 72 meses.

Os valores de RMSE (%) e BIAS (%) foram distintos nos três níveis de modelagem no ajuste, embora na classe de produtividade média (S = 27) tenham sido próximos para as alternativas no nível de povoamento (7_0) e de árvore individual (2_0), neste último foi mais acurado, contudo ligeiramente mais viesado (Tabela 5).

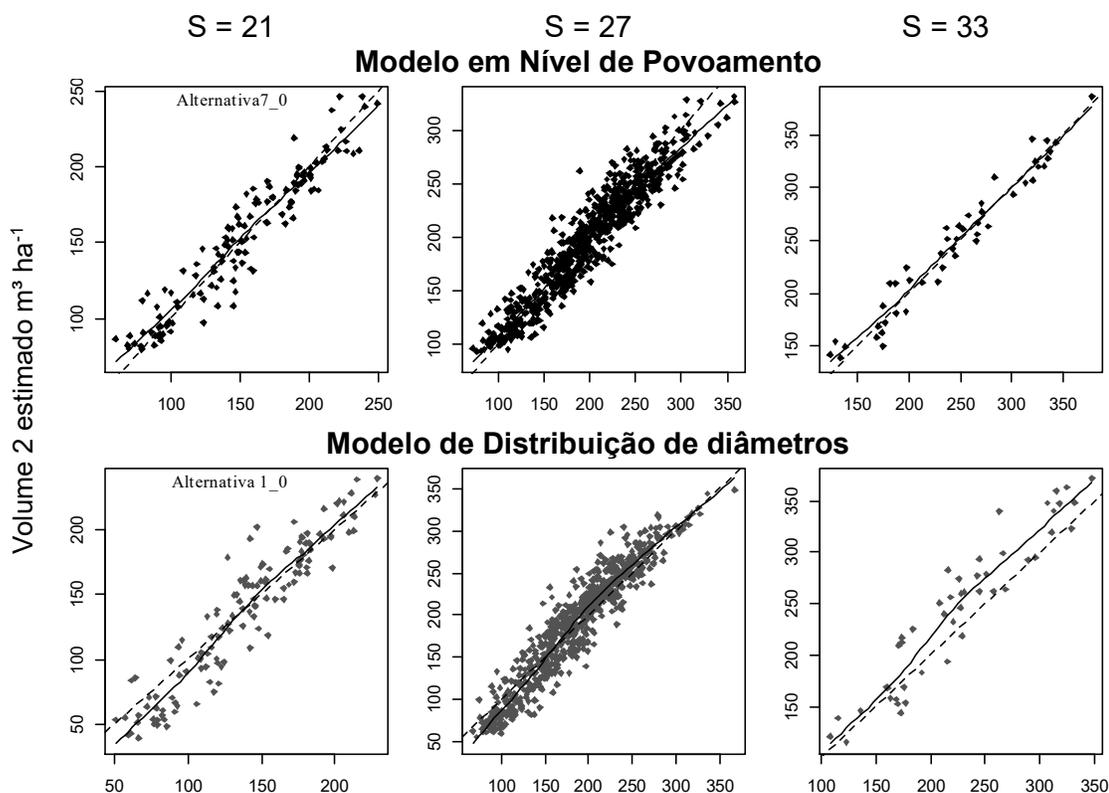
Tabela 5 – Estatísticas obtidas, por classe de produtividade, para o ajuste e validação de modelos de crescimento e produção em nível de povoamento, distribuição de diâmetros e árvore individual, com modelos fixos.

Nível de modelagem	Alternativa	Classe	r_{yy}	RMSE (%)	BIAS (%)
Ajuste					
Povoamento	7_0	S = 21	0,9506	9,40	0,8781
		S = 27	0,9514	8,65	-0,2067
		S = 33	0,977	5,85	1,2171
Distribuição de diâmetros	1_0	S = 21	0,9397	13,79	-1,3080
		S = 27	0,9475	11,75	2,2890
		S = 33	0,9466	13,30	8,0690
Árvore individual	2_0	S = 21	0,9557	10,66	5,4257
		S = 27	0,9581	8,10	-0,8783
		S = 33	0,9678	8,70	-4,9032

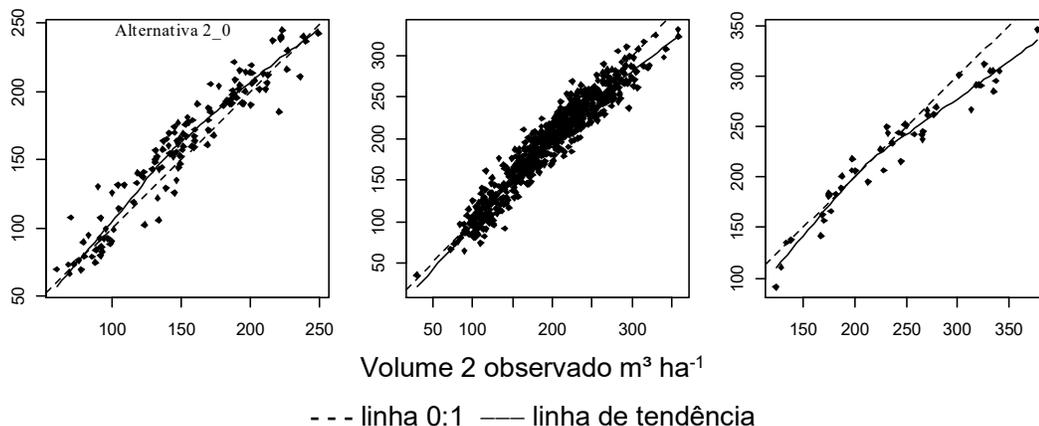
		Validação			
Povoamento	7_0	S = 21	0,9513	5,61	1,3222
		S = 27	0,9346	6,41	0,4151
		S = 33	0,9883	4,58	2,6142
Distribuição de diâmetros	1_0	S = 21	0,9670	6,56	4,3190
		S = 27	0,9662	8,31	6,8490
		S = 33	0,9885	11,88	11,228
Árvore individual	2_0	S = 21	0,7883	16,46	-11,6225
		S = 27	0,8654	17,49	-14,2105
		S = 33	0,9330	20,52	-17,2601

Pela análise gráfica da Figura 1 é possível confirmar que a dispersão entre os volumes totais estimados e observados é maior nesta classe quando consideramos a alternativa 7_0, apresentando tendência de subestimação para parcelas de maior volume, e ligeira superestimação em povoamentos de menor volume, o que também pode ser confirmado ao examinarmos a dispersão residual apresentada na Figura 3 para esta alternativa.

Figura 1 – Volumes estimados versus observados obtidos pelo ajuste de modelos de crescimento e produção em diferentes níveis, utilizando modelos fixos.

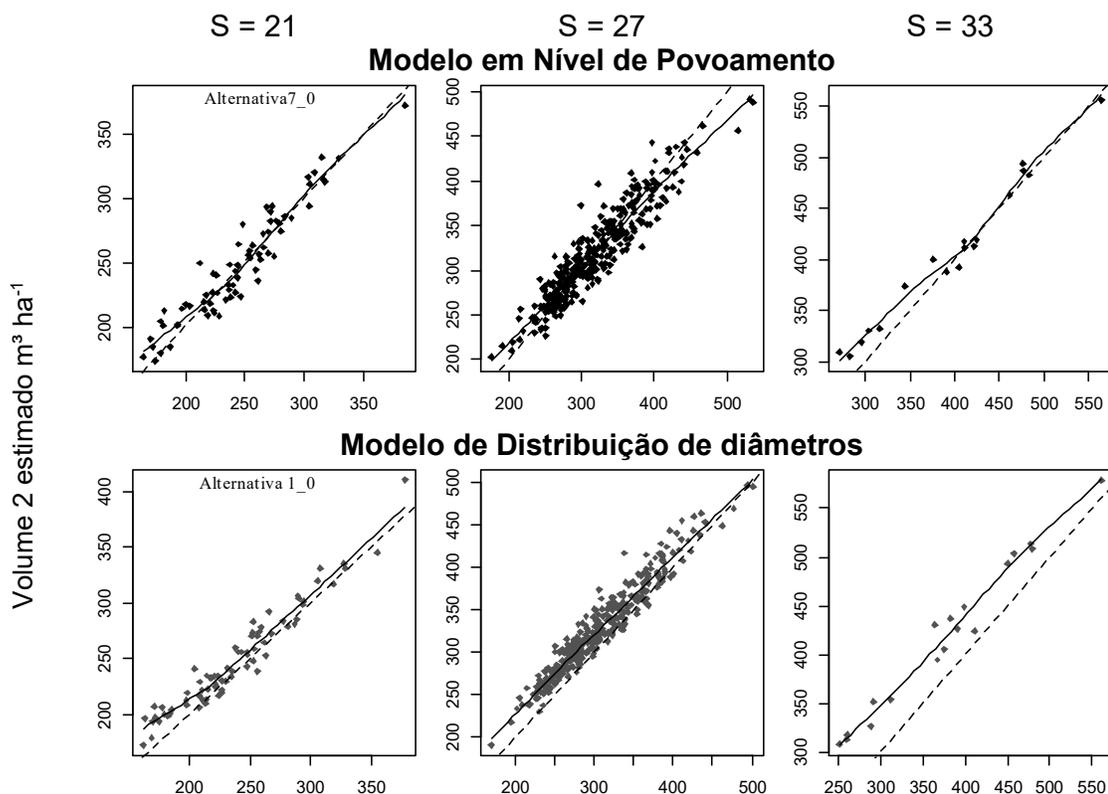


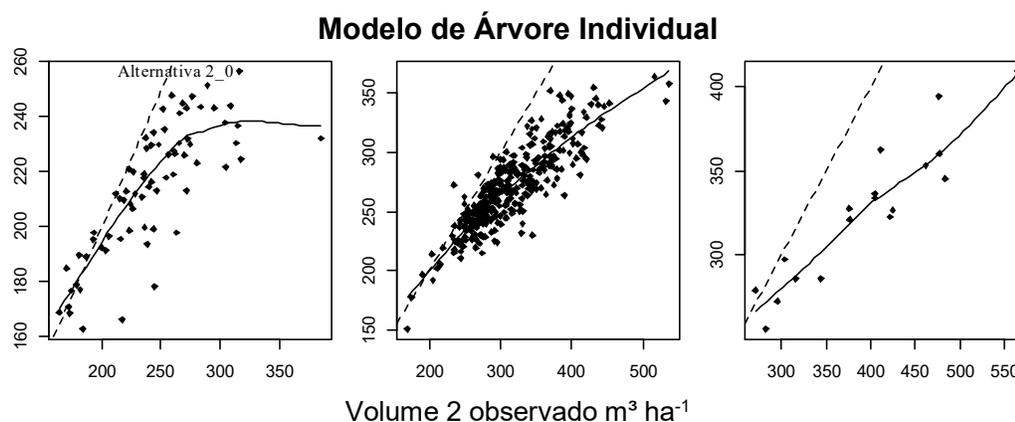
Modelo de Árvore Individual



O modelo no nível de povoamento (alternativa 7_0) apresentou comportamento semelhante na validação, conforme estatísticas apresentadas para validação deste nível na Tabela 1. Fato corroborado pela análise da Figura 2, na qual apesar de apresentar maior tendência de subestimação em parcelas com maior volume na classe de produtividade Média ($S = 27$), é mais acurada que as alternativas avaliadas para distribuição de diâmetros (1_0) e árvore individual (2_0).

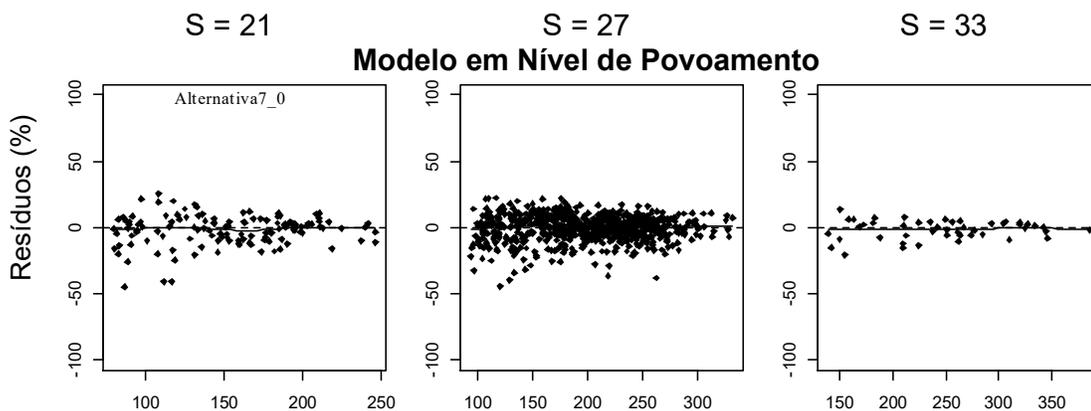
Figura 2 – Volumes estimados versus observados obtidos pela aplicação dos modelos de crescimento em diferentes níveis a uma base de dados independente, para os modelos fixos avaliados.





Salienta-se que os modelos de árvore individual e distribuição de diâmetros são constituídos de submodelos e por este motivo espera-se que erros destes submodelos sejam somados ao erro global das equações ajustadas. Sendo assim, o modelo de distribuição de diâmetros foi menos acurado, e mais tendencioso quando comparados aos dois níveis, no ajuste. Contudo, na validação, observa-se na Tabela 1 e na Figura 2 que o modelo de árvore individual apresentou uma tendência de subestimação incompatível com o que foi apresentado no ajuste, para todas as classes de produtividade. Tal fato também é revelado na dispersão do resíduo relativo de acordo com a Figura 4.

Figura 3 – Distribuição dos resíduos percentuais das equações ajustadas para os modelos de crescimento e produção em diferentes níveis utilizando modelagem fixa, para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em $\text{m}^3 \text{ha}^{-1}$ e no eixo Y o erro em percentagem.



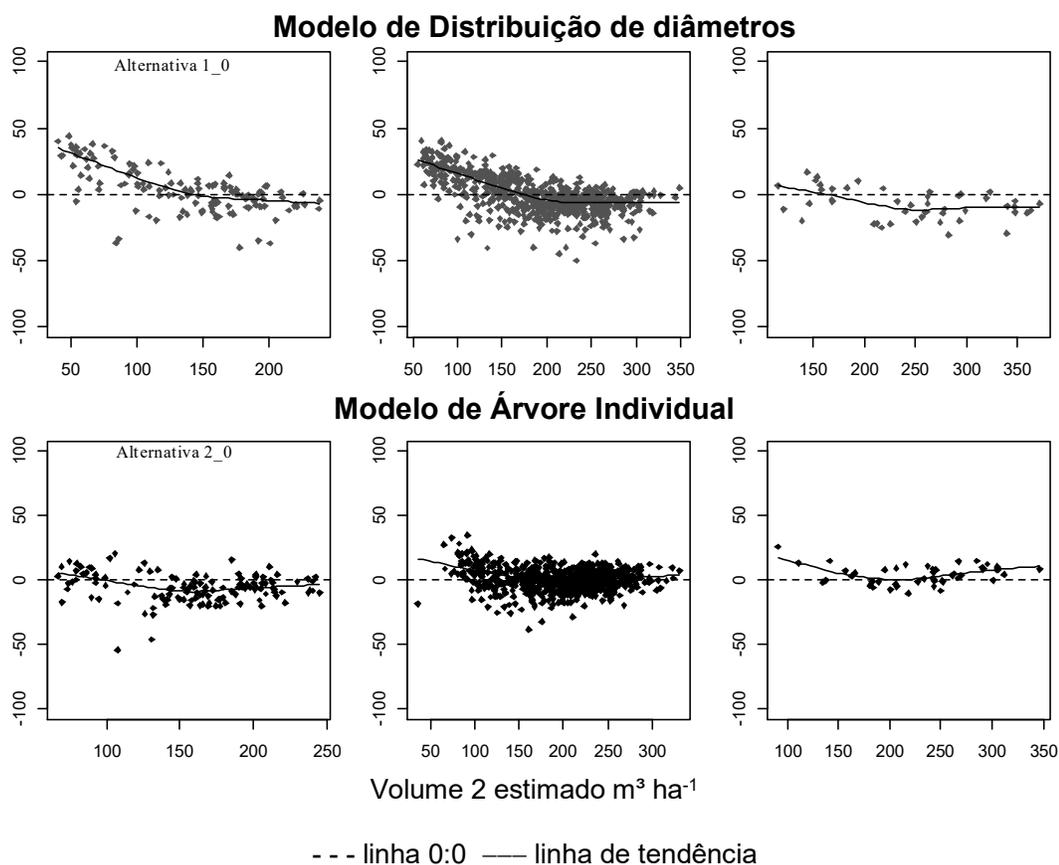
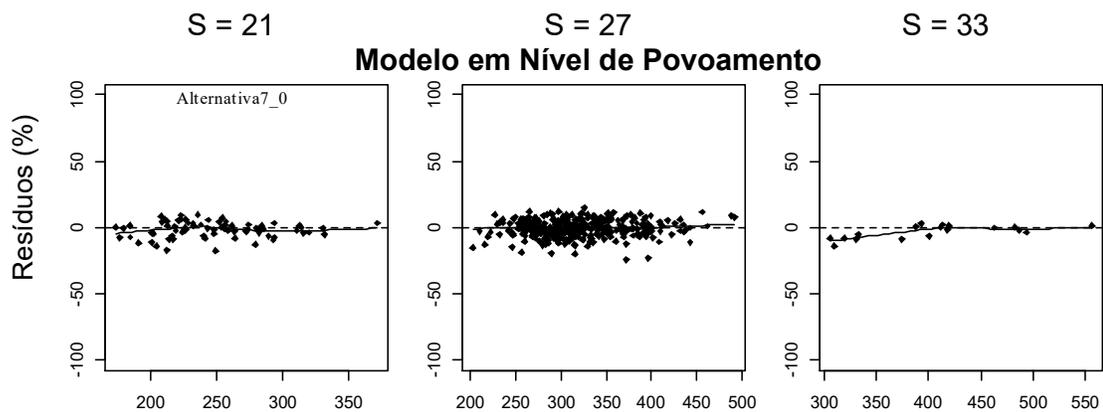
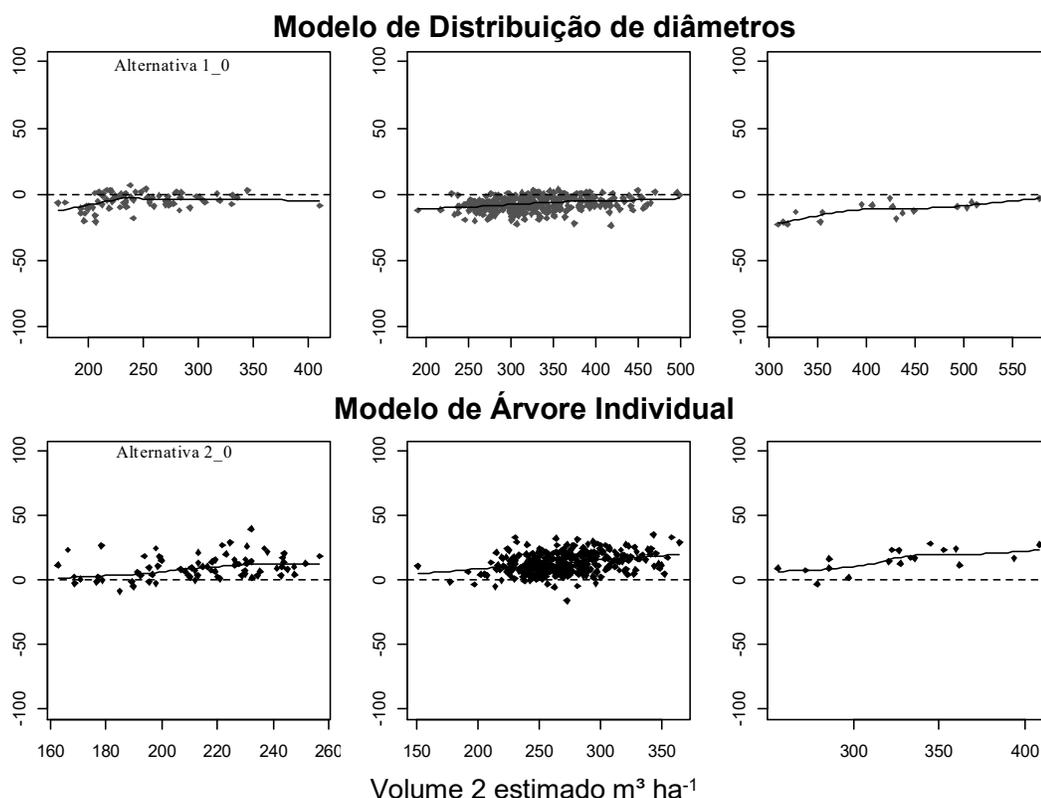


Figura 4 – Distribuição dos resíduos percentuais obtidos pela aplicação dos modelos de crescimento e produção no nível de povoamento (alternativa 7_1), distribuição de diâmetros (1_1) e árvore individual (2_2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em $m^3 ha^{-1}$ e no eixo Y o erro em porcentagem.





Para as alternativas avaliadas em diferentes níveis de modelagem de crescimento e produção, utilizando dados divididos segundo o critério 1, o modelo em nível de povoamento de Mendonça (2010) (alternativa 7_0) foi o que apresentou maior acurácia e menos viesado, seguido do modelo linear de crescimento em DAP e HT de árvore individual (alternativa 2_0), e por último o modelo de distribuição de diâmetros.

3.1.2. Volumes totais obtidos por equações ajustados e validados com o conjunto de dados do critério 2, com modelagem fixa

Para as alternativas fixas propostas para os diferentes níveis de modelagem, ajustadas e validadas com os dados denominados critério 2, é realizada a análise das estatísticas e análise gráfica nesta seção. Recorda-se que segundo este critério, o conjunto de parcelas foi dividido aleatoriamente em dois grupos, sendo 70% das parcelas utilizadas para ajuste dos modelos e 30% para validação, ou seja, as parcelas do conjunto de ajuste não estavam presentes no conjunto de validação.

A alternativa que utilizou o modelo de árvore individual apresentou maior acurácia na classe de produtividade Média ($S = 27$), que concentra 81% do total de parcelas. Embora tenha sido ligeiramente mais viesada, ao avaliar-se o gráfico (Figura

5) que apresenta os volumes observados versus os estimados, verifica-se que a tendência é menos acentuada quando comparada à alternativa 7 na mesma classe. A mesma tendência é mantida para o modelo de árvore individual (alternativa 1) na Figura 6, com o gráfico dos volumes observados e estimados, obtidos pela aplicação da alternativa na base de dados de validação.

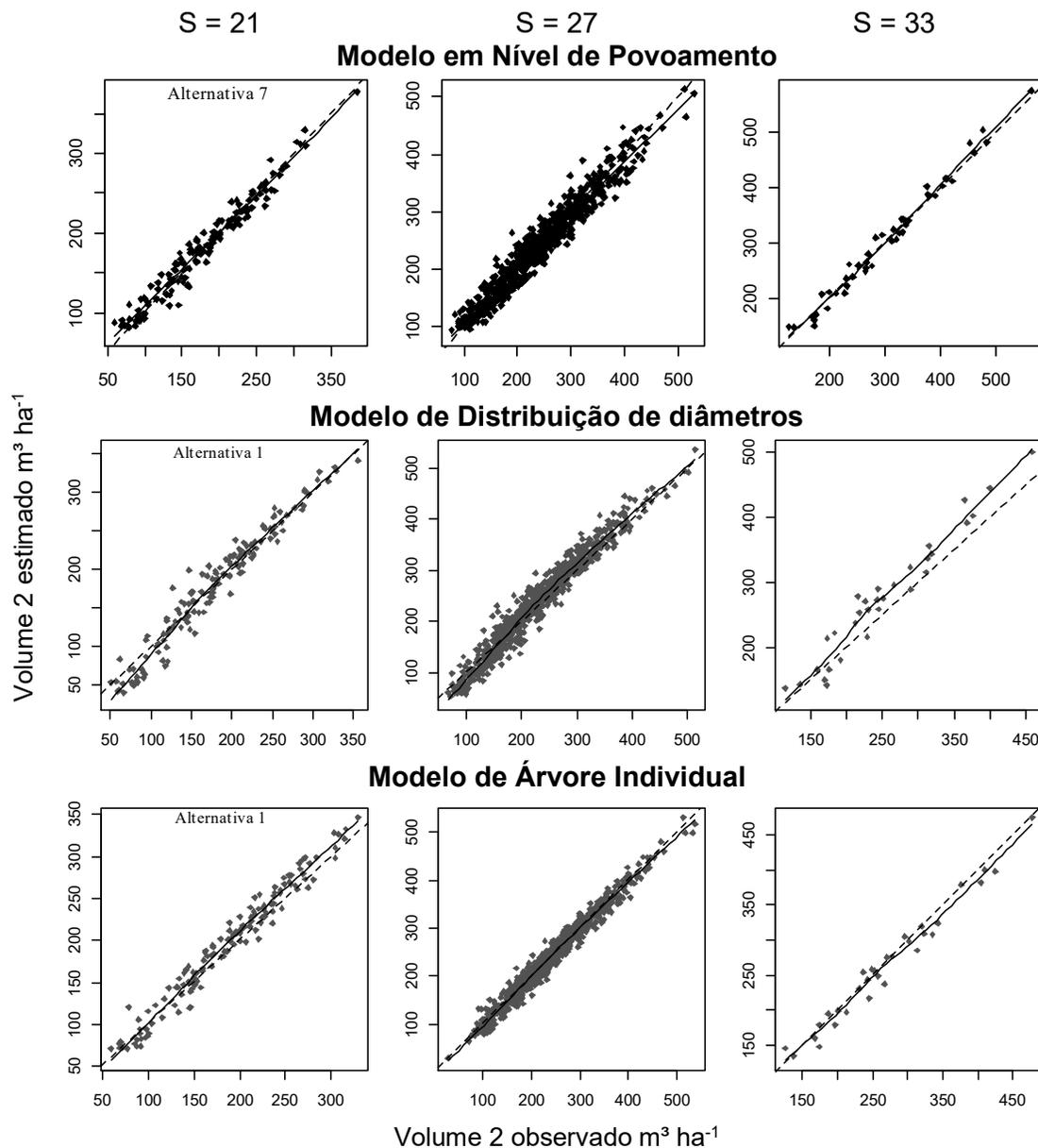
Tabela 6 – Estatísticas obtidas, por classe de produtividade, para o ajuste e validação de modelos de crescimento e produção em nível de povoamento, distribuição de diâmetros e árvore individual, utilizando a parcela como grupo de efeitos aleatórios, utilizando modelos mistos.

Nível de modelagem	Alternativa	Classe	$r_{y\hat{y}}$	RMSE (%)	BIAS (%)
Ajuste					
Povoamento	7	S = 21	0,9790	7,14	1,2230
		S = 27	0,9720	7,79	-0,2190
		S = 33	0,9910	4,71	0,8910
Distribuição de diâmetros	1	S = 21	0,9747	9,76	-0,0783
		S = 27	0,9771	9,30	3,0921
		S = 33	0,9746	12,59	8,8899
Árvore individual	1	S = 21	0,9829	8,30	4,4746
		S = 27	0,9855	5,94	-0,4373
		S = 33	0,9864	5,73	-2,3953
Validação					
Povoamento	7	S = 21	0,9660	8,08	0,0642
		S = 27	0,9800	6,84	-0,2613
		S = 33	0,9870	6,21	3,3755
Distribuição de diâmetros	1	S = 21	0,9730	10,05	1,3674
		S = 27	0,9710	10,34	3,6127
		S = 33	0,9793	11,09	7,3315
Árvore individual	1	S = 21	0,9682	11,10	6,7505
		S = 27	0,9791	6,62	0,6349
		S = 33	0,9860	6,80	-3,4590

Na classe de produtividade Baixa (S = 21), a alternativa 7 foi mais acurada de acordo com o RMSE relativo apresentado e pela análise gráfica (Figura 5). Ainda para esta alternativa, verifica-se que a classe de produtividade Alta (S = 33) é a mais acurada e menos viesada comprando-se às demais tanto no ajuste, quanto na validação. A alternativa de árvore de individual apresentou menor acurácia e um viés acentuado nesta classe.

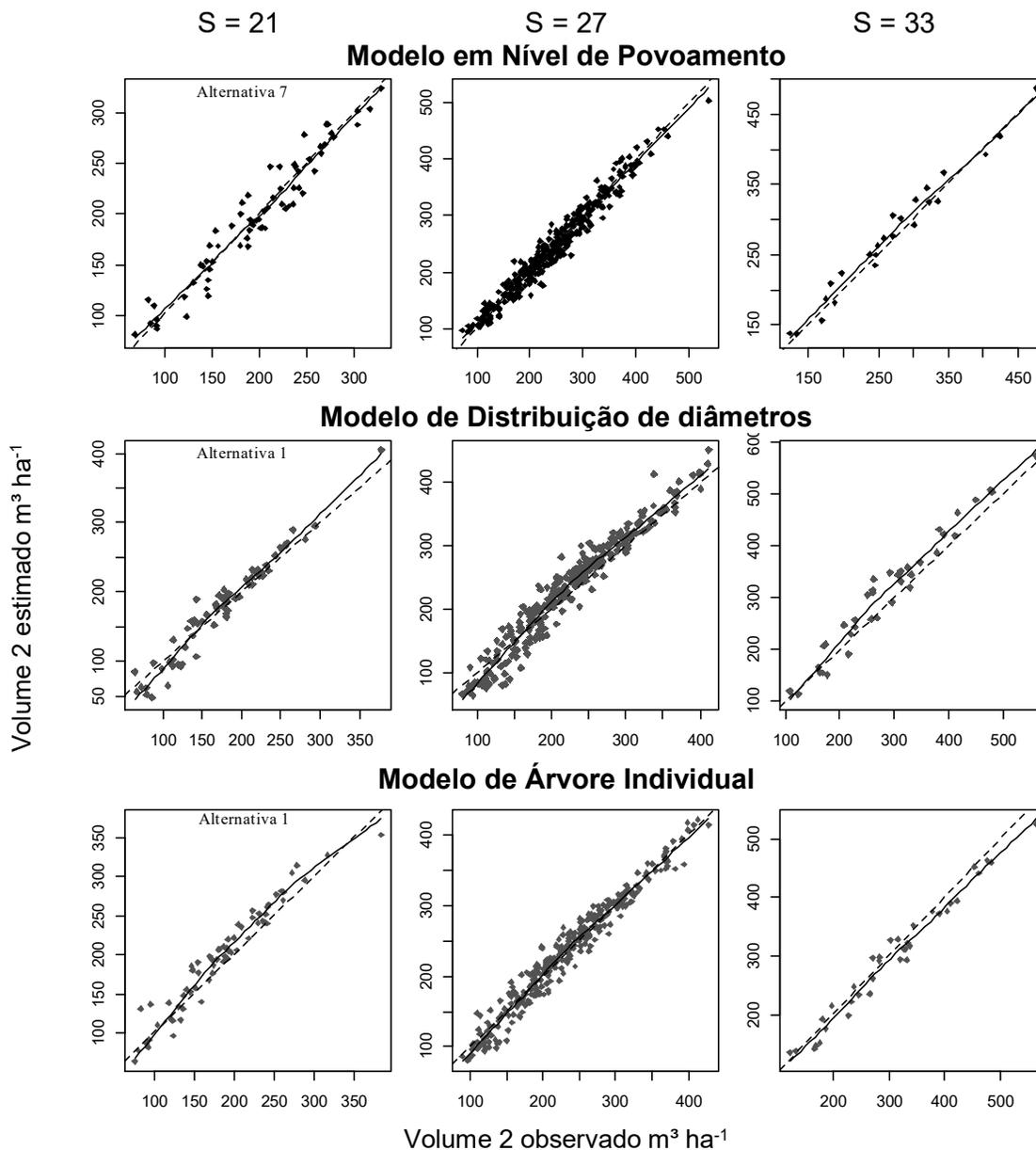
Tomando como base a classe de produtividade Alta (S = 33), a modelagem no nível de povoamento é menos viesada e mais acurada no ajuste, e na validação, contudo, com aumento do viés tendendo à superestimação.

Figura 5 – Volumes estimados versus observados obtidos pelo ajuste de modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (5) para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro observado, em $\text{m}^3 \text{ha}^{-1}$ e no eixo Y o volume estimado em $\text{m}^3 \text{ha}^{-1}$.



O comportamento dos volumes observados e estimados na validação é próximo com o que foi avaliado para o ajuste (Figura 5) das alternativas em diferentes níveis com dados do critério 2 pode ser observado na Figura 6. Verifica-se a presença da tendência de superestimação no modelo de distribuição de diâmetros (alternativa 1) em parcelas com maior volume, e de subestimação em parcelas de menor volume, mais acentuada na classe de produtividade Alta (S = 33).

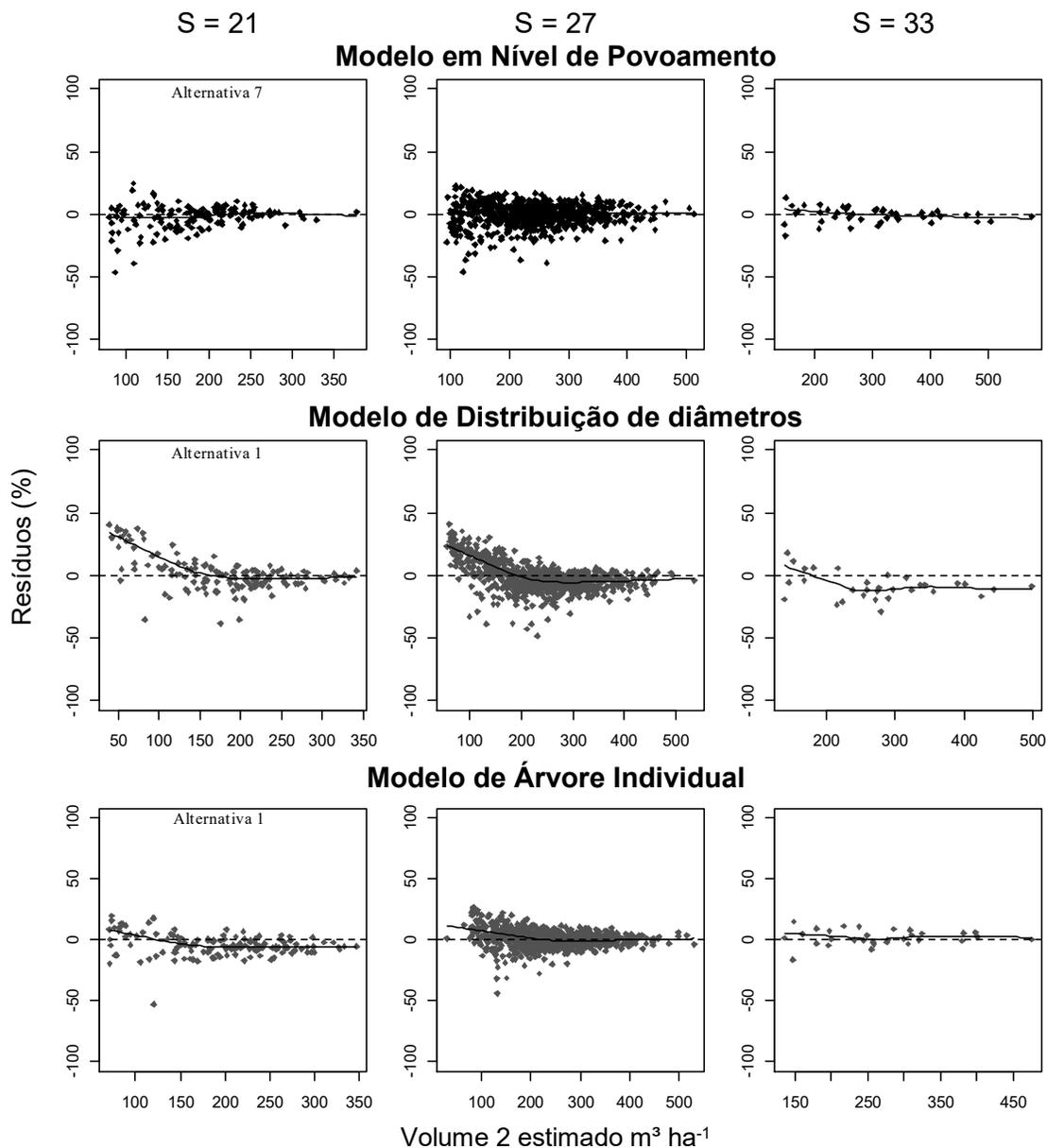
Figura 6 – Volumes estimados versus observados obtidos pela aplicação dos modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro observado, em $\text{m}^3 \text{ha}^{-1}$ e no eixo Y o volume estimado em $\text{m}^3 \text{ha}^{-1}$.



A dispersão residual (Figura 7) do ajuste confirma a análise feita para as alternativas apresentadas, com maior concentração de pontos entre $\pm 25\%$ na classe de produtividade Média de alternativa 1 de modelos de árvore individual, com ligeira tendência para subestimação em parcelas com menor volume nesta classe. A maior tendência de subestimação verificada ocorre nas classes de produtividade Baixa (S = 21) e Média (S = 27) para o modelo de distribuição de diâmetros.

Neste nível de modelagem, apresenta-se uma tendência de superestimar parcelas com volume superior a 200 m³ ha⁻¹.

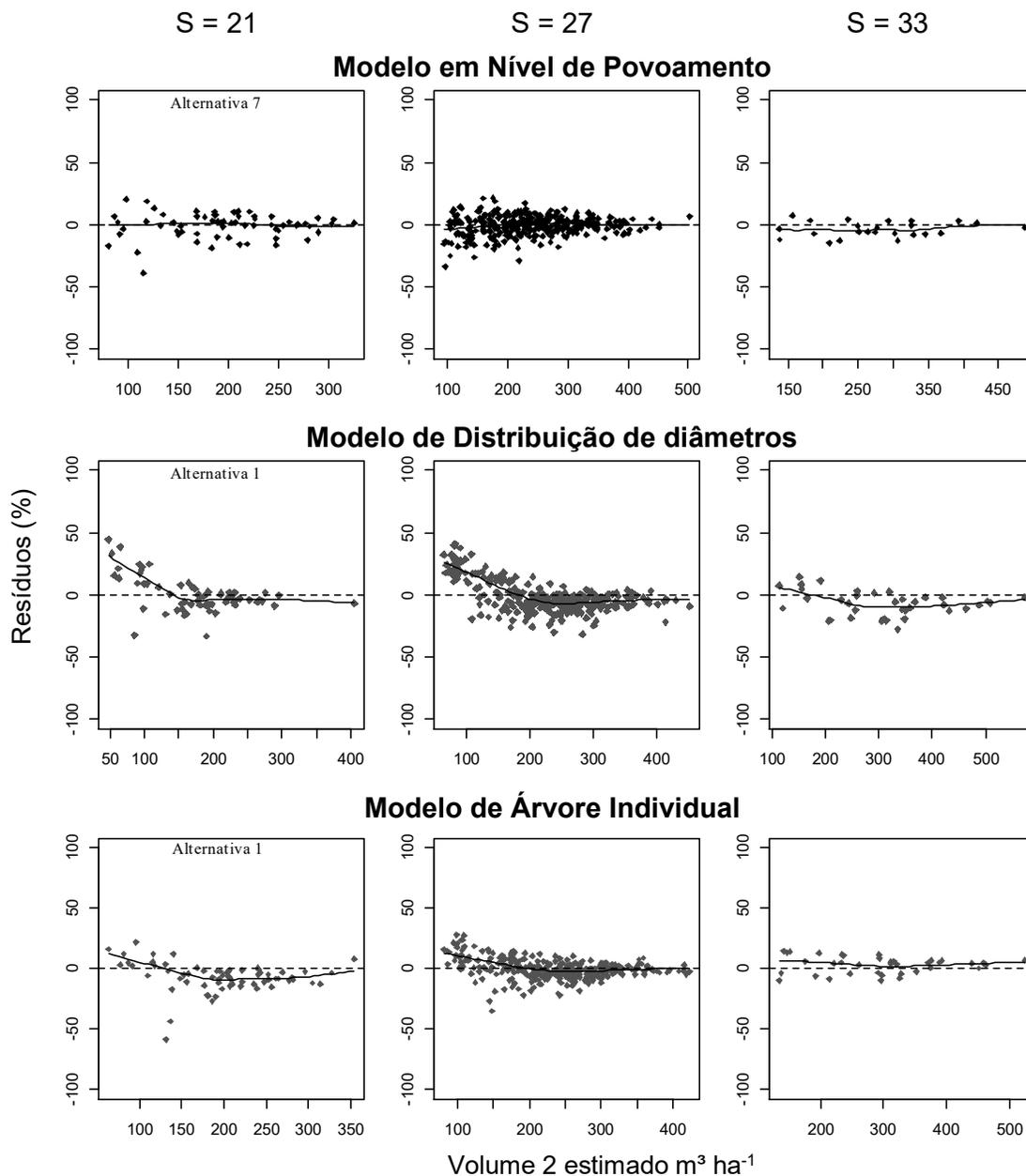
Figura 7 – Distribuição dos resíduos percentuais das equações ajustadas para os modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (5), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em m³ ha⁻¹ e no eixo Y o erro em porcentagem.



Comportamento similar ao avaliado na dispersão residual do ajuste é apresentado para a validação, como pode ser observado na Figura 8. Embora na

classe de produtividade Alta ($S = 33$) para o modelo em nível de povoamento a tendência passou a ser de superestimação em parcelas de até $400 \text{ m}^3 \text{ ha}^{-1}$.

Figura 8 – Distribuição dos resíduos percentuais obtidos pela aplicação dos modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em $\text{m}^3 \text{ ha}^{-1}$ e no eixo Y o erro em percentagem.



Para as alternativas avaliadas em diferentes níveis de modelagem de crescimento e produção, utilizando dados divididos segundo o critério 2, o modelo em

nível de povoamento de Mendonça (2010) (alternativa 7) foi o que apresentou maior acurácia e menor viés, seguido do modelo Linear adaptado de Bella (1971) e Campos e Leite (2009) (MARTINS et al., 2014) (alternativa 1), e do modelo de distribuição de diâmetros (alternativa 1), respectivamente. Corroborando os resultados apresentados, Gonçalves (2014) mostrou, dentro dos critérios adotados na pesquisa, que o modelo de Mendonça (2010) superou os resultados obtidos pelo modelo de Clutter (1963).

3.2. Avaliação de modelos de crescimento e produção em diferentes níveis, utilizando modelagem mista

Nesta seção são apresentados os resultados das alternativas que utilizaram modelagem mista, ajustadas e validadas para os dados divididos pelos critérios 1 e 2, avaliadas neste trabalho.

3.2.1. Volumes totais obtidos por equações calibradas com a parcela como grupo de efeitos aleatórios, com modelagem mista utilizando dados do critério 1

Os resultados apresentados nesta seção referem-se às alternativas que utilizaram modelagem mista com dados divididos segundo o critério 1, tanto no ajuste, quanto na aplicação (validação) em uma base de dados distinta. O modelo de árvore individual (alternativa 2_2) foi o que apresentou melhor resultado no ajuste, com elevada acurácia, e ausente de viés, confirmados pelas estatísticas RMSE e BIAS relativo e pela Figura 9.

Na validação, verifica-se que a modelagem de árvores individuais continua apresentando superioridade, embora para outra alternativa (2_5), embora na classe de produtividade Baixa o viés tenha sido ligeiramente superior que o viés apresentado na mesma classe para a alternativa 7_1. Cabe destacar que na validação o modelo de distribuição de diâmetros teve sua tendência acentuada quando se compara ao que foi obtido no ajuste. No caso da classe de produtividade Média (S = 27), a tendência foi contrária ao apresentado no ajuste, desta vez de superestimação.

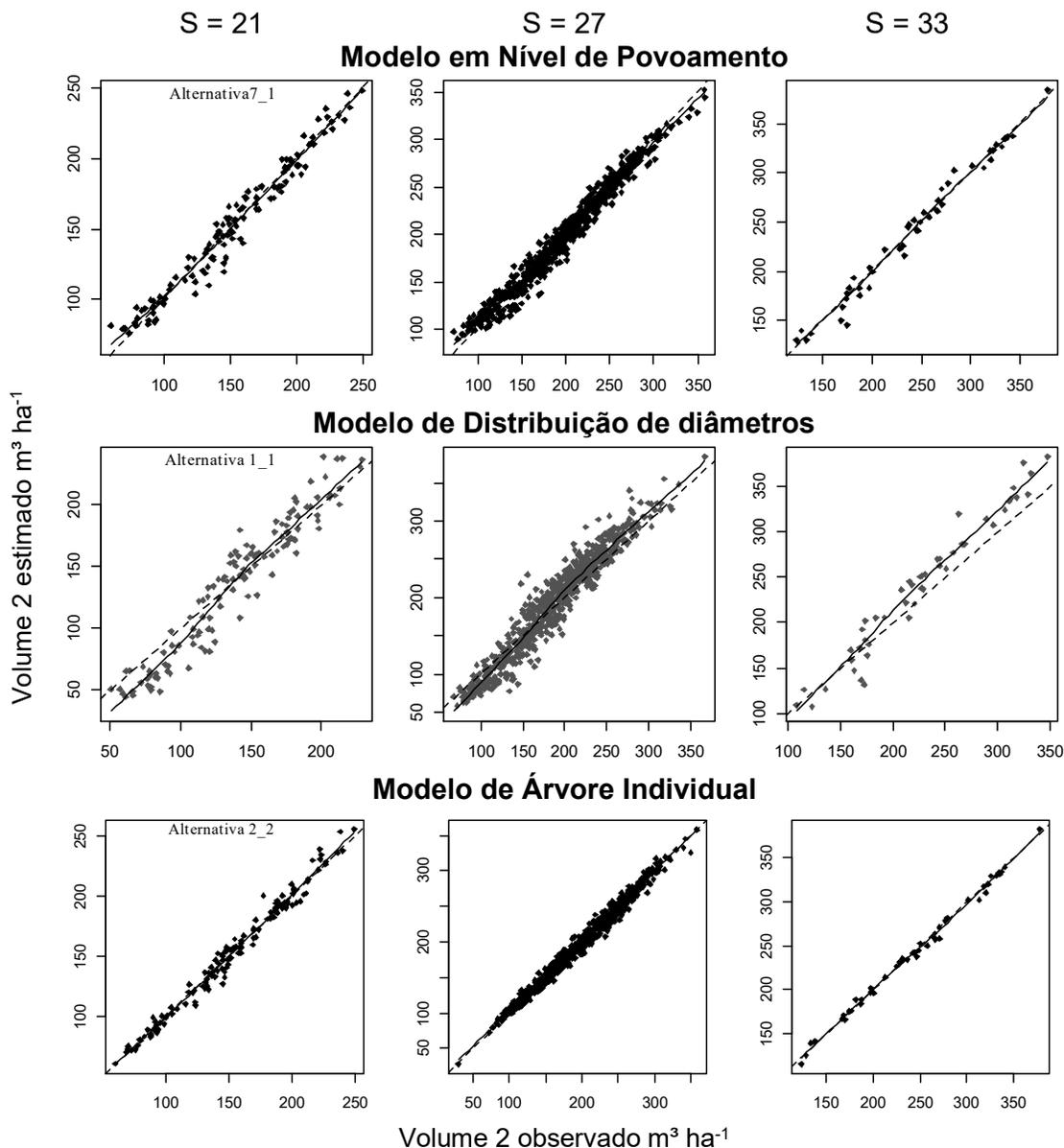
Tabela 7 – Estatísticas obtidas, por classe de produtividade, para o ajuste e validação de modelos de crescimento e produção em nível de povoamento, distribuição de diâmetros e árvore individual, utilizando a parcela como grupo de efeitos aleatórios, utilizando modelos mistos.

Nível de modelagem	Alternativa	Classe	$r_{\hat{y}y}$	RMSE (%)	BIAS (%)
Ajuste					
Povoamento	7_1	S = 21	0,9816	5,76	0,1168
		S = 27	0,9866	4,59	0,0092
		S = 33	0,9909	3,71	-0,0439
Distribuição de diâmetros	1_1	S = 21	0,9632	12,00	-2,257
		S = 27	0,9634	10,58	2,676
		S = 33	0,9799	10,14	5,944
Árvore individual	2_2	S = 21	0,9912	4,18	0,1569
		S = 27	0,9928	3,37	-0,2148
		S = 33	0,9975	1,99	-0,5085
Validação					
Povoamento	7_1	S = 21	0,9435	6,70	1,3544
		S = 27	0,9491	7,01	-1,0427
		S = 33	0,9829	7,80	-0,7443
Distribuição de diâmetros	1_1	S = 21	0,9660	6,62	4,239
		S = 27	0,9610	8,95	7,377
		S = 33	0,9921	11,45	11,098
Árvore individual	2_5	S = 21	0,9658	4,93	1,6222
		S = 27	0,9450	5,88	-0,2626
		S = 33	0,9921	3,29	0,5257

Na Figura 9 ainda é possível confirmar ausência perceptível de viés para o modelo de árvore individual em todas as classes de produtividade, como também a acurácia superior às demais, com a nuvem de pontos com leve dispersão.

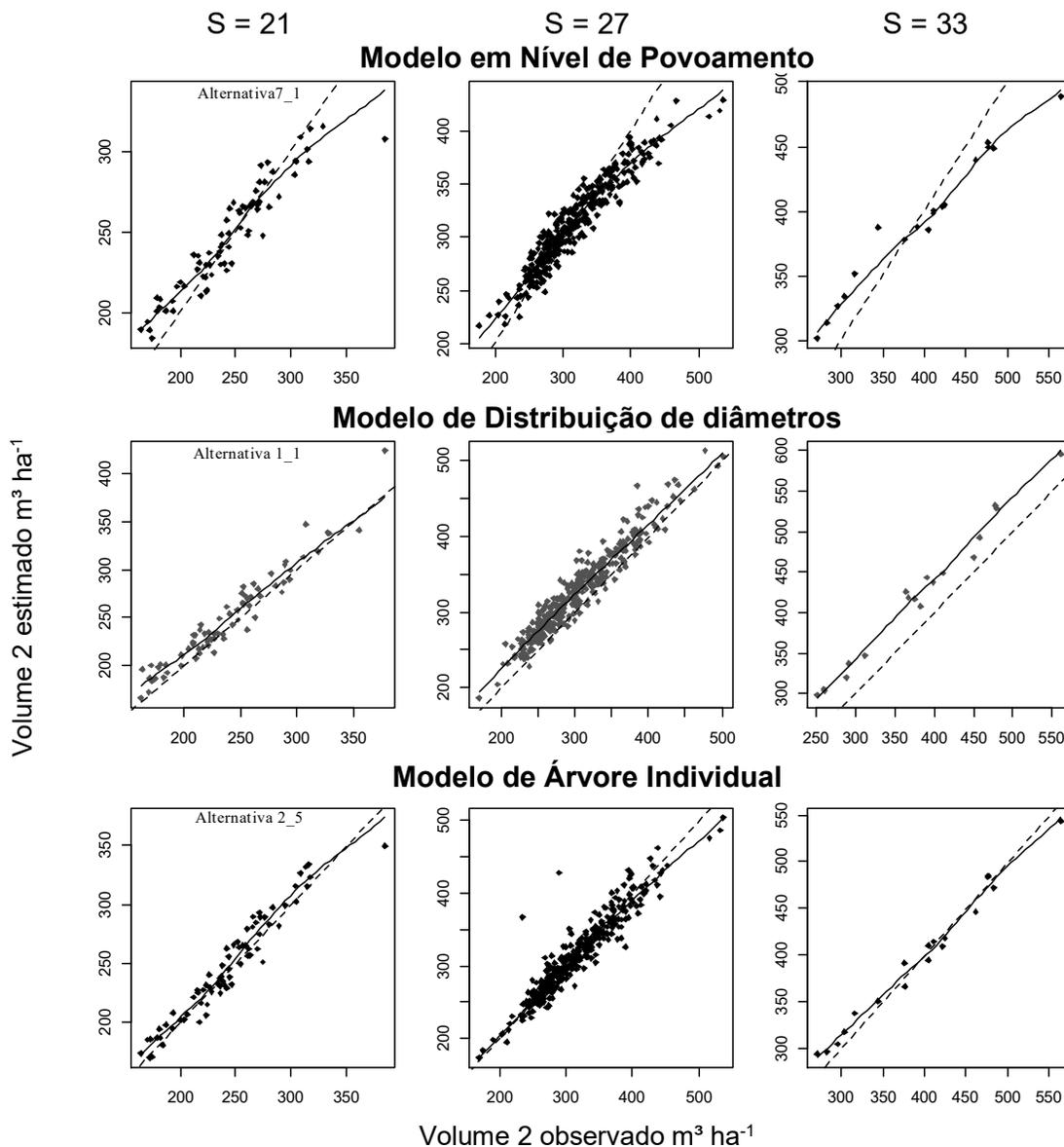
O modelo em nível de povoamento, ajustado pela alternativa 7_1 vem em seguida, mostrando-se acurado e não viesado, exceto quando avalia-se na classe de produtividade Baixa (S = 21), uma leve tendência de superestimação em parcelas com menor volume, o que pode ser confirmado na Figura 11. Ainda nesta alternativa, mas na classe de produtividade Média (S = 27) o que se verifica é uma leve tendência de subestimação em parcelas de maior volume.

Figura 9 – Volumes estimados versus observados obtidos pelo ajuste de modelos de crescimento e produção no nível de povoamento (alternativa 7_1), distribuição de diâmetros (1_1) e árvore individual (2_2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro observado, em m^3ha^{-1} e no eixo Y o volume estimado em m^3ha^{-1} .



O modelo de distribuição de diâmetros apresenta uma tendência de superestimação para parcelas com volume acima de $200 \text{ m}^3\text{ha}^{-1}$ nas classes de produtividade Média e Alta, e tendência contrária na classe de produtividade Baixa (Figura 9).

Figura 10 – Volumes estimados versus observados obtidos pela aplicação dos modelos de crescimento e produção no nível de povoamento (alternativa 7_1), distribuição de diâmetros (1_1) e árvore individual (2_2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro observado, em m^3ha^{-1} e no eixo Y o volume estimado em m^3ha^{-1} .



Os gráficos de resíduos (Figuras 11 e 12), confirmam o que foi discutido até aqui, porém, percebe-se que o modelo de distribuição de diâmetros apresenta um comportamento contrário na validação, passando a superestimar, e não subestimar como ocorreu no ajuste em todas as classes de produtividade. Verifica-se que a tendência é mais acentuada nas classes de produtividade Média e Alta, e nas parcelas com menores volumes.

Comparando as alternativas de modelagem em nível de povoamento e de árvore individual, percebeu-se que o comportamento da última é mais estável, embora a avaliação seja para alternativas diferentes no ajuste e na validação.

Figura 11 – Distribuição dos resíduos percentuais das equações ajustadas para os modelos de crescimento e produção no nível de povoamento (alternativa 7_1), distribuição de diâmetros (1_1) e árvore individual (2_2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em $m^3 ha^{-1}$ e no eixo Y o erro em porcentagem.

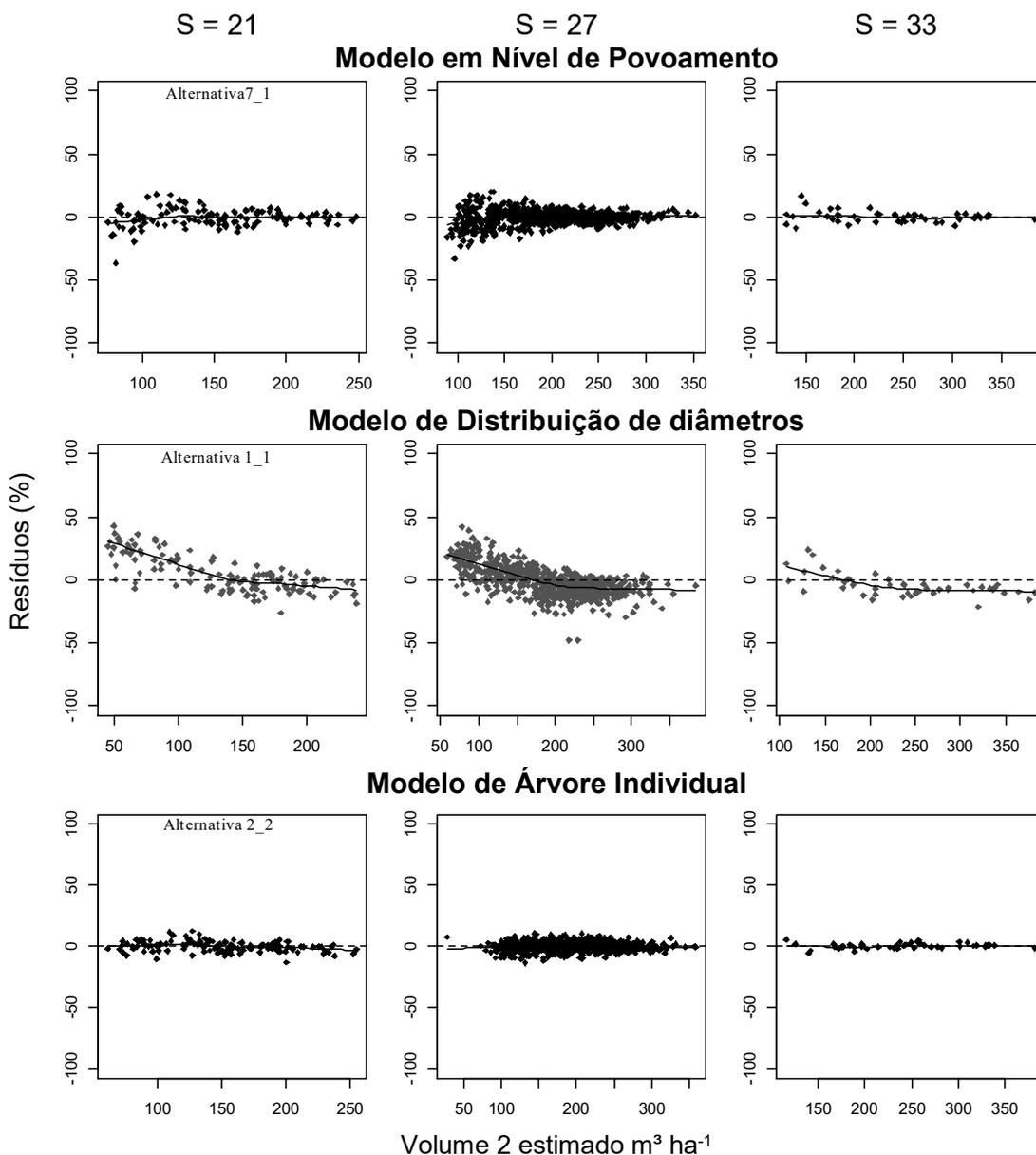
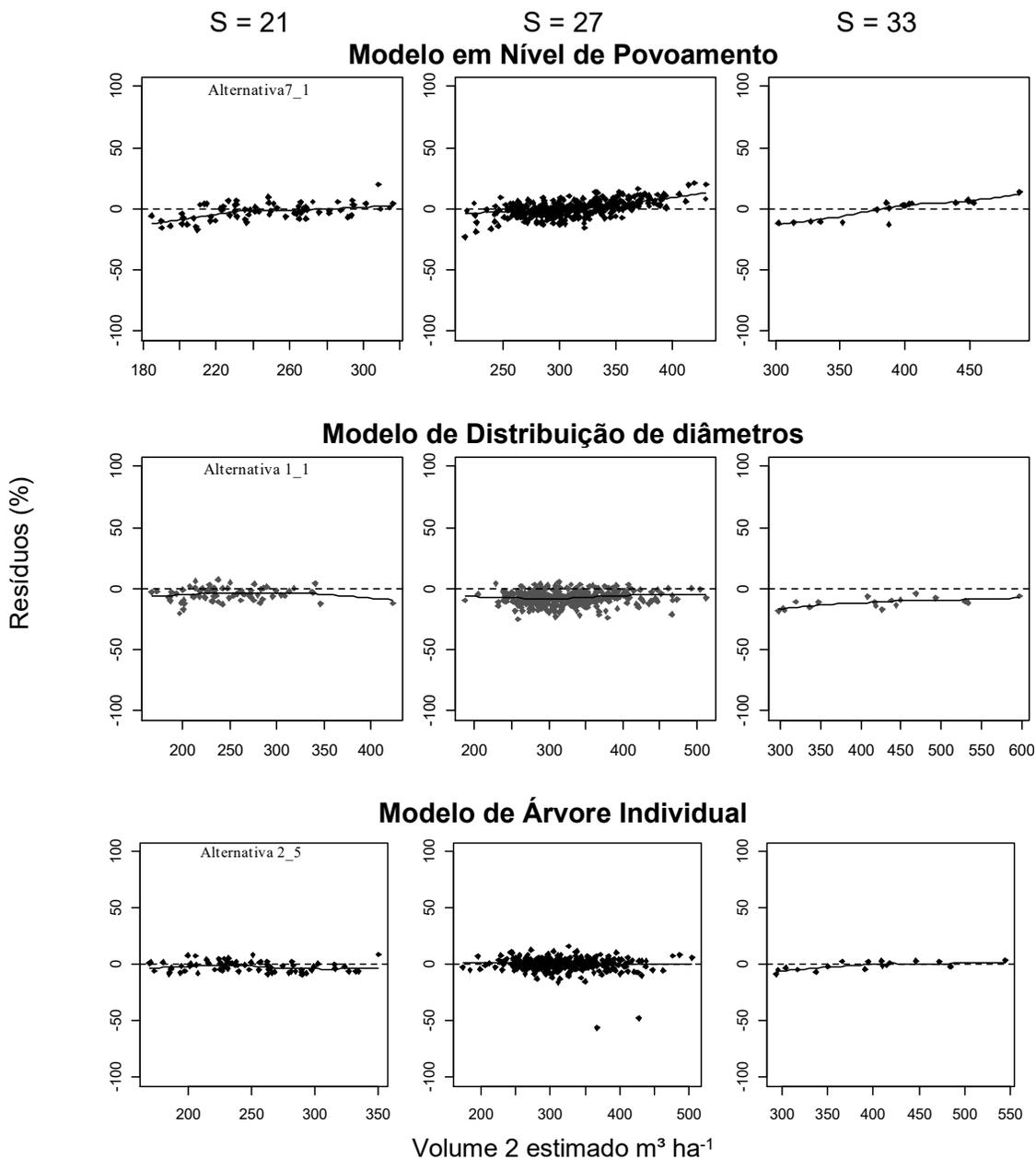


Figura 12 – Distribuição dos resíduos percentuais obtidos pela aplicação dos modelos de crescimento e produção no nível de povoamento (alternativa 7_1), distribuição de diâmetros (1_1) e árvore individual (2_2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em $\text{m}^3 \text{ha}^{-1}$ e no eixo Y o erro em porcentagem.



Ao avaliar alternativas mistas de modelagem de crescimento em diferentes níveis, utilizando o critério 1, no qual do conjunto de ajuste participaram todas as parcelas com idade futura (idade 2) até 60 meses, e no segundo (de validação), foram selecionadas as mesmas parcelas do grupo de ajuste, porém com idade futura igual ou superior a 72 meses e dentre elas destaca-se as alternativas utilizando modelagem

em nível de árvore individual, sendo mais acuradas e menos tendenciosas, no ajuste e na validação.

3.2.2. Volumes totais obtidos por equações calibradas para outras variáveis como grupo de efeitos aleatórios, utilizando modelagem mista para o conjunto de dados do critério 2

Nesta seção são apresentados e discutidos resultados de alternativas de modelagem mistas ajustadas e validadas com os dados do critério 2, que não utilizaram a parcela como grupo de efeitos aleatórios. Tal abordagem pode ser adequada quando as parcelas para as quais será feita a aplicação não participaram do ajuste.

De acordo com as estatísticas apresentadas na Tabela 8, os modelos em nível de povoamento apresentaram-se mais acurados e menos viesados quando comparados aos demais, embora apresentem resultado equivalente aos modelos de árvore individual no ajuste, sobretudo nas classes de produtividade Média e Alta.

A alternativa de modelagem de distribuição de diâmetros apresentou um comportamento mais acurado e menos viesado quando comparado ao resultado obtido pelo ajuste e validação usando dados do critério 1. Embora ligeira, a tendência de superestimar volumes é perceptível principalmente na classe de produtividade Alta (S = 33).

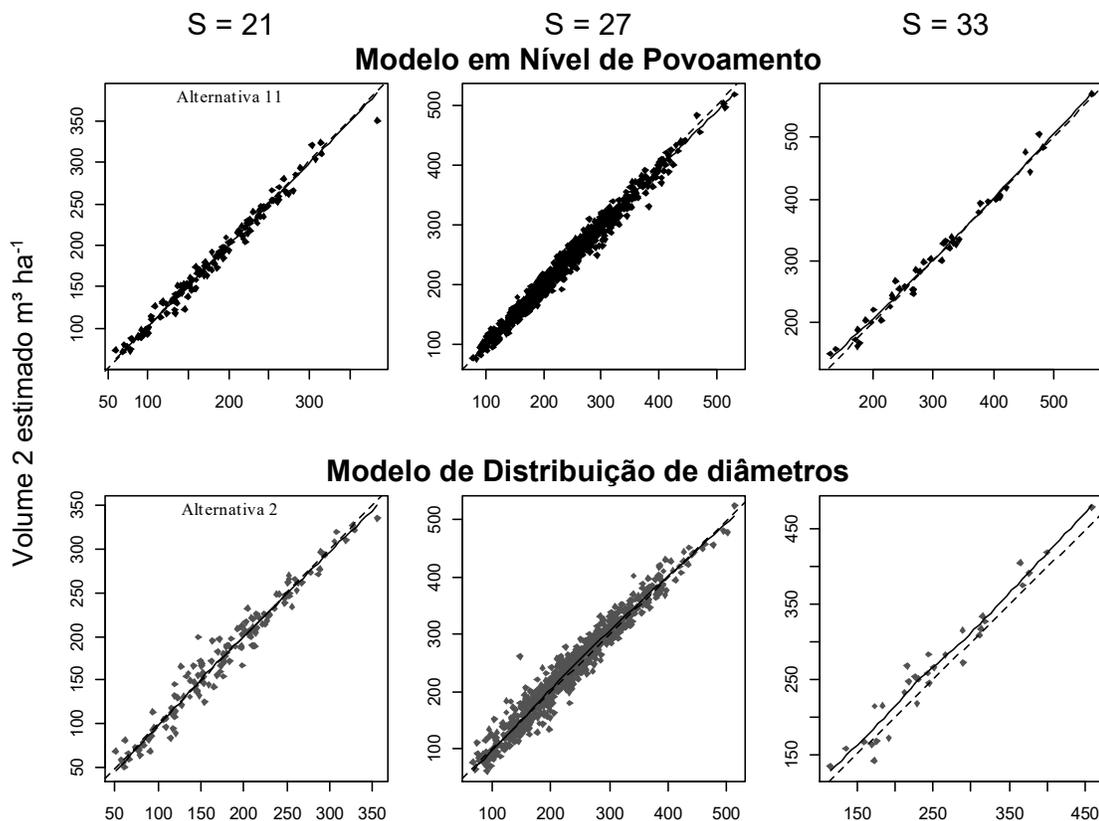
Tabela 8 – Estatísticas obtidas, por classe de produtividade, para o ajuste e validação de modelos de crescimento e produção em nível de povoamento, distribuição de diâmetros e árvore individual, utilizando a parcela como grupo de efeitos aleatórios, utilizando modelos mistos.

Nível de modelagem	Alternativa	Classe	$r_{\hat{y}y}$	RMSE (%)	BIAS (%)
Ajuste					
Povoamento	11	S = 21	0,9919	4,43	0,1957
		S = 27	0,9905	4,56	-0,2869
		S = 33	0,9921	4,41	1,3930
Distribuição de diâmetros	2	S = 21	0,9791	7,82	-0,0861
		S = 27	0,9778	7,82	1,7009
		S = 33	0,9754	9,18	5,4058
Árvore individual	5	S = 21	0,9817	6,97	2,4434
		S = 27	0,9863	5,71	-0,4715
		S = 33	0,9891	4,69	-0,7678

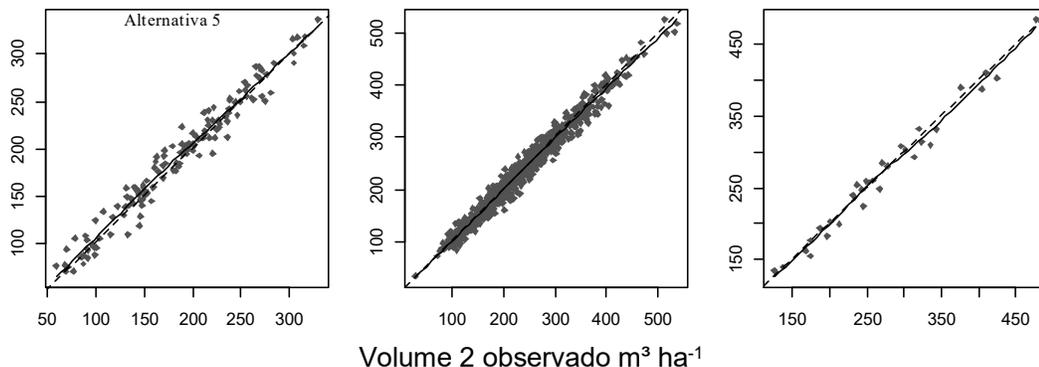
		Validação			
Povoamento	11	S = 21	0,9858	5,63	1,6071
		S = 27	0,9880	5,33	-0,4757
		S = 33	0,9849	6,66	3,4785
Distribuição de diâmetros	4	S = 21	0,9691	9,15	0,9870
		S = 27	0,9697	8,9	2,2548
		S = 33	0,9799	9,38	5,6704
Árvore individual	2	S = 21	0,9739	9,91	6,4053
		S = 27	0,9839	5,74	0,5775
		S = 33	0,9869	5,96	-2,1551

Corroborando com este resultado, a análise gráfica dos volumes observados versus estimados por estes modelos (Figura 13 e Figura 14) demonstram a dispersão similar e a ausência de tendência em todas as classes de produtividade destes níveis de modelagem.

Figura 13 – Volumes estimados versus observados obtidos pelo ajuste de modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (5) para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro observado, em m^3ha^{-1} e no eixo Y o volume estimado em m^3ha^{-1} .

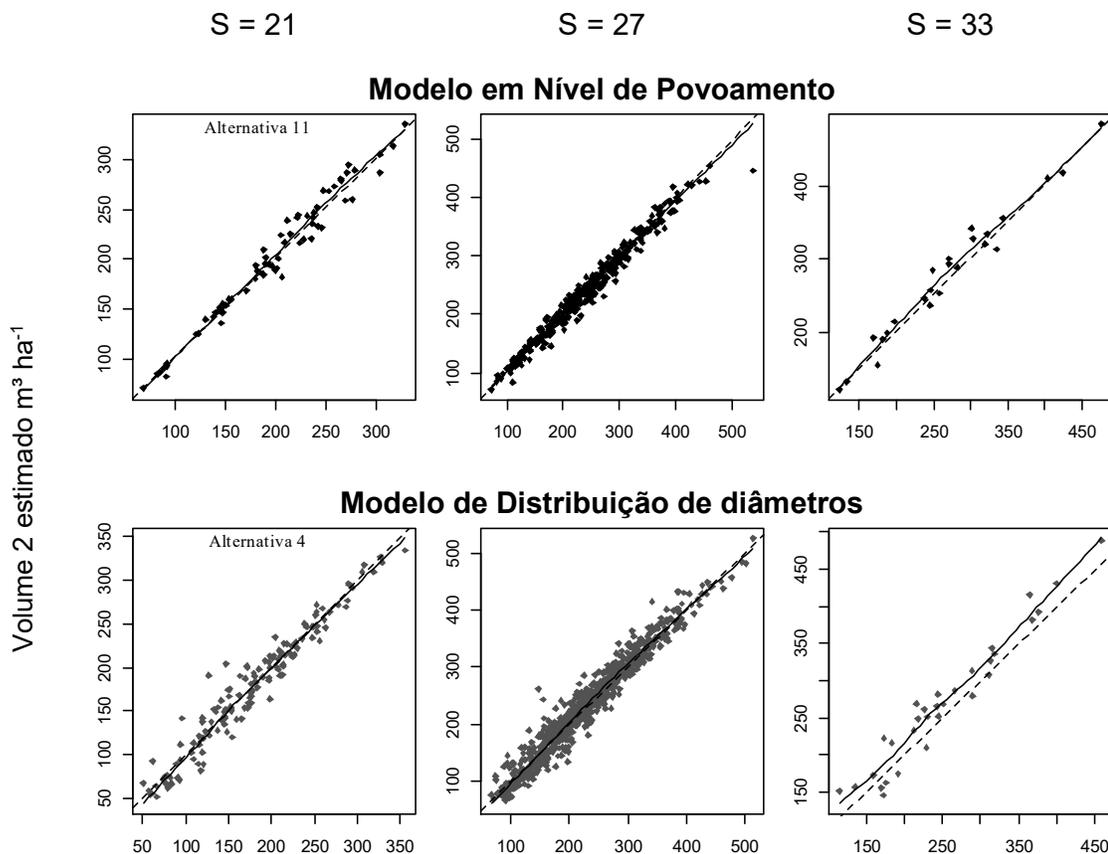


Modelo de Árvore Individual

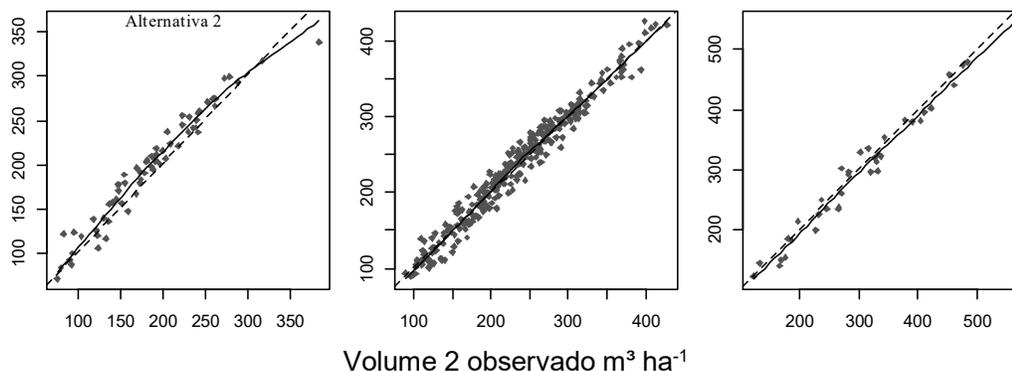


Nota-se que embora o modelo de árvore individual tenha sido acurado, a dispersão entre volumes observados e estimados (Figura 14) é maior em todas as classes de produtividade se comparado à Figura 10.

Figura 14 – Volumes estimados versus observados obtidos pela aplicação dos modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro observado, em m³ha⁻¹ e no eixo Y o volume estimado em m³ha⁻¹.

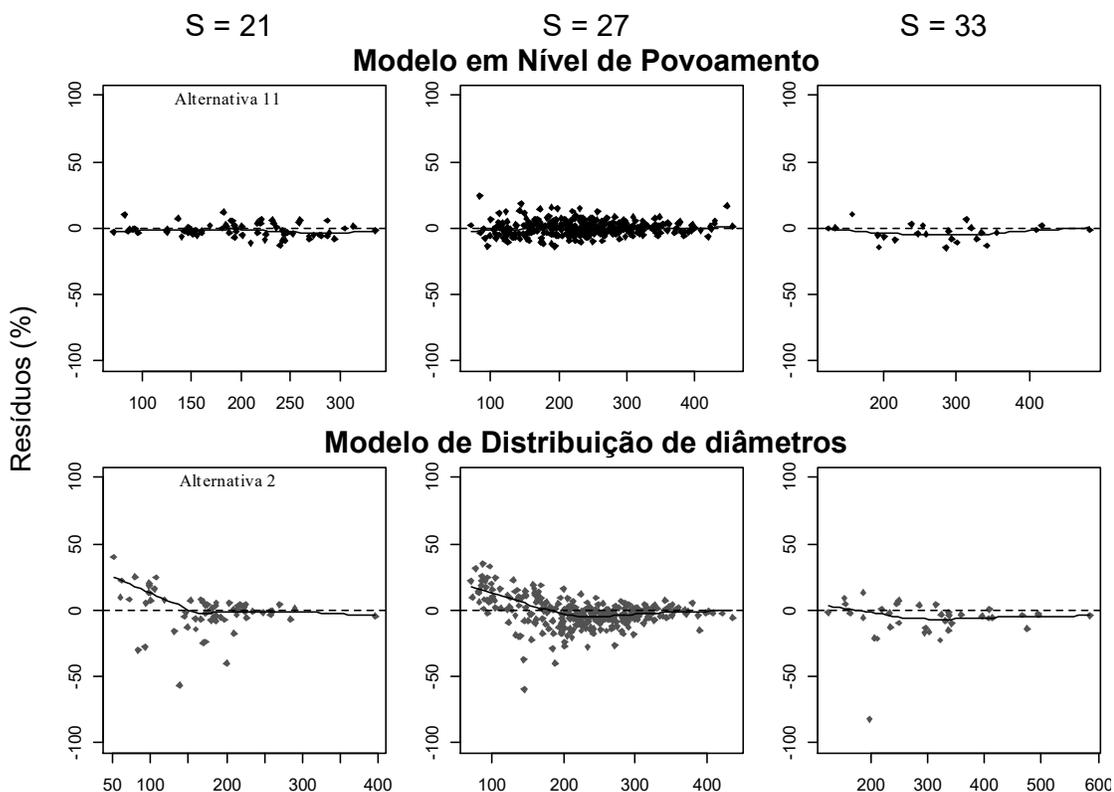


Modelo de Árvore Individual



O comportamento residual apresentado para o modelo de distribuição de diâmetros (Figura 15) é o mesmo comparado ao apresentado na seção anterior, quando os dados foram o do critério 1, na Figura 11 e na Figura 12.

Figura 15 – Distribuição dos resíduos percentuais das equações ajustadas para os modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (5), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em m³ ha⁻¹ e no eixo Y o erro em percentagem.



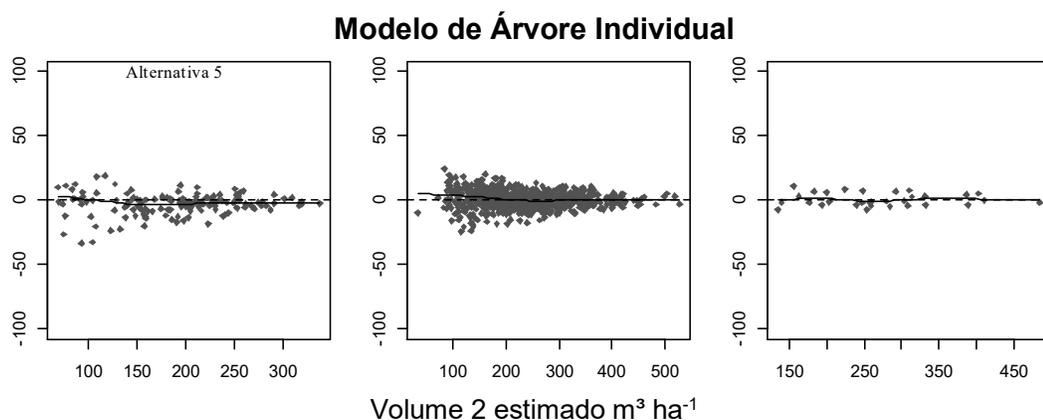
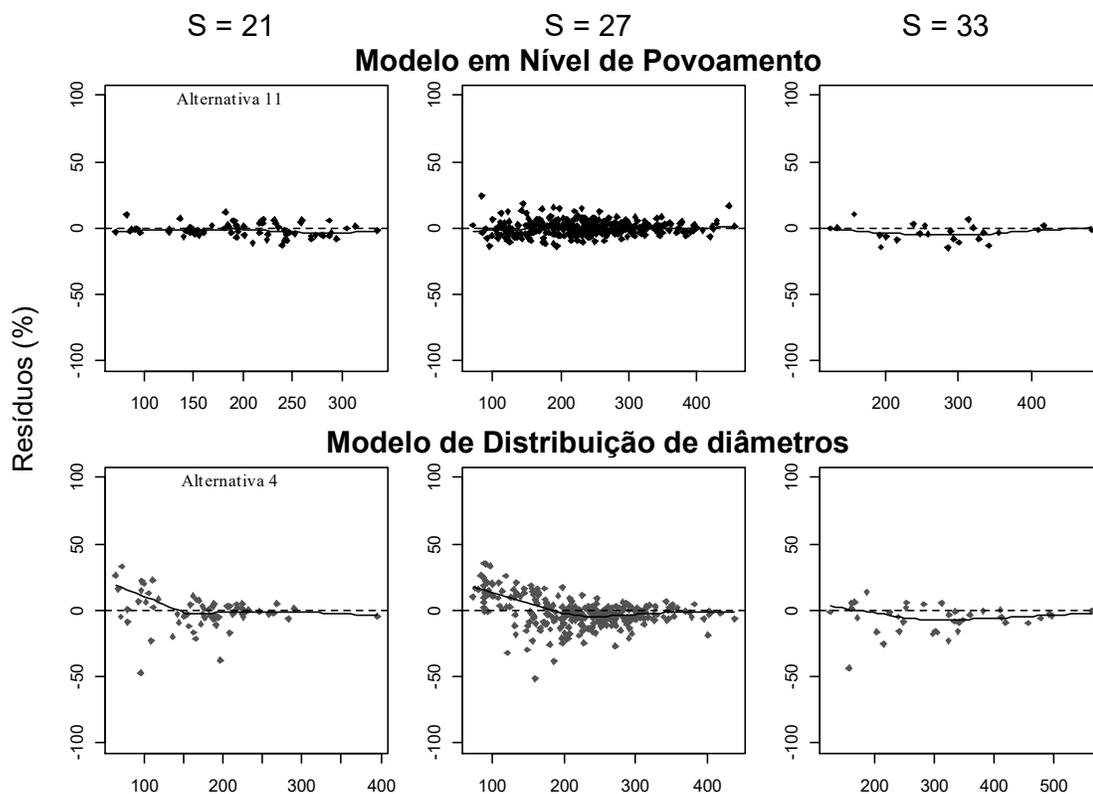
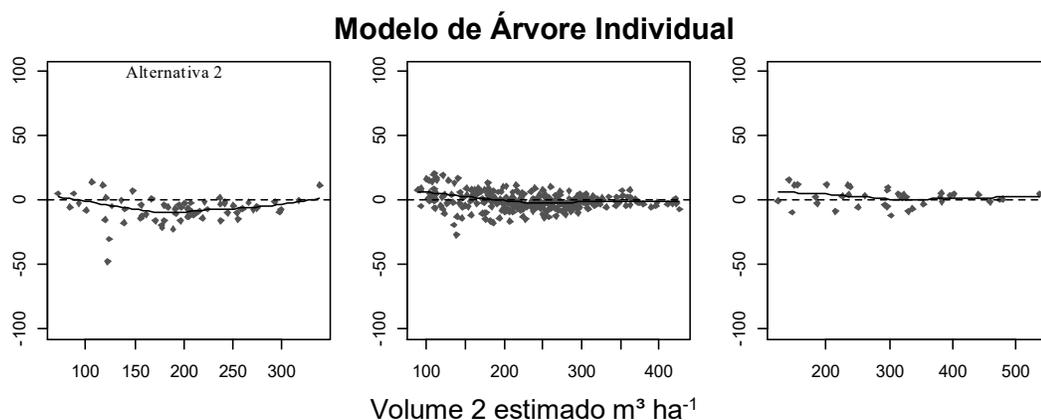


Figura 16 – Distribuição dos resíduos percentuais obtidos pela aplicação dos modelos de crescimento e produção no nível de povoamento (alternativa 11), distribuição de diâmetros (2) e árvore individual (2), para as diferentes classes de qualidade de local, em que no eixo X está o volume futuro estimado, em $\text{m}^3 \text{ha}^{-1}$ e no eixo Y o erro em percentagem.





Ao avaliar alternativas mistas de modelagem de crescimento em diferentes níveis, utilizando o critério 2, pode-se afirmar que as alternativas para o nível de povoamento (11) e árvore individual (2) tiveram resultados semelhantes em termos de acurácia e viés, tanto no ajuste, quanto na validação. Sendo assim, ambos podem ser utilizados para estimação do volume de povoamentos de eucalipto.

3.3. Comparação entre os diferentes níveis de modelagem para ambos os critérios com e sem efeito aleatório

No nível de povoamento, as alternativas mistas avaliadas apresentaram melhores resultados para acurácia e ausência de viés quando considerado o ajuste em todas classes de produtividade, tanto para os dados do critério 1, quanto para os dados do critério 2. Na validação, a superioridade foi confirmada nas classes de produtividade Baixa e Média pela alternativa mista em nível de povoamento que utilizou a parcela como grupo de efeitos aleatórios no critério 2, entretanto menos acurada e mais tendenciosa na classe de produtividade Alta.

Apesar de apresentar acurácia moderada no ajuste, a alternativa fixa em nível de povoamento que utilizou o critério 2, apresentou a pior acurácia dentre as demais avaliadas neste nível.

Para a distribuição de diâmetros a modelagem mista foi superior quando os dados de ajuste e validação foram do critério 2. Para o mesmo critério, a alternativa fixa foi superior à alternativa mista que utilizou o critério 1, entretanto, na validação foi menos acurada, embora menos viesada. Destaca-se ainda nos modelos de distribuição de diâmetros com dados do critério 2 a redução do viés comparado às demais alternativas fixas e mistas para este nível de modelagem.

Dentre todas as alternativas avaliadas nos três níveis de modelagem, aquelas de distribuição de diâmetros foram as menos acuradas no ajuste e na validação, seja

com modelos fixos ou mistos, e com tendências de superestimação dos volumes totais.

Na modelagem de árvore individual, as alternativas mistas que utilizaram dados do critério 1 foram as mais acuradas dentre todas as avaliadas em todos os níveis, no ajuste e na validação, corroborando com os resultados encontrados por Martins (2011) e Castro et al. (2013). Aquelas que usaram o critério 2 também apresentaram superioridade nestes quesitos, porém semelhantes à alternativa no nível de povoamento.

Embora em todos os níveis de modelagem estudados os efeitos aleatórios tenham trazido ganhos em acurácia, cabe ressaltar que muito ainda há que se pesquisar para que esta técnica estatística potencialize ainda mais os seus benefícios. Dentre as possibilidades de pesquisa, estão a modelagem das estruturas de correlação e da das estruturas heterocedástica intra-indivíduos, a avaliação de diferentes variáveis nas bases de dados de inventários florestais para formação de grupos de efeitos aleatórios e a avaliação de outras possibilidades de variáveis para formarem inclinações aleatórias podem contribuir com o avanço do uso de modelos mistos no manejo florestal.

Salienta-se que a partir dos resultados obtidos pelos modelos de árvore individual, é possível obter os mesmos resultados que na modelagem de distribuição de diâmetros e de em nível de povoamento. Apesar disso, faz-se necessário ponderar que seu ajuste é computacionalmente mais complexo, principalmente para bases de dados de grande número de observações (superiores a 100.000 observações).

Outra dificuldade apresentada é a manutenção ao longo do tempo de uma base de dados consistida em nível de árvore individual. Tal questão pode demandar processos específicos para gerenciamento dos dados de inventário florestal, ou seja, política e planejamento mais detalhado, aliado a adoção de tecnologias apropriadas para validação de dados mesmo em campo, como também comprometimento da equipe responsável.

4. CONCLUSÕES

Para realizar a projeção (critério 1) utilizando a modelagem fixa, o modelo de povoamento de Mendonça (2010) apresentou maior acurácia, e menos tendenciosidade. Entretanto, ao avaliar-se a modelagem mista, o modelo linear de árvore individual para crescimento em diâmetro a 1,3 m de altura e altura total proposto neste trabalho foi superior ao resultado obtido pelo modelo de Mendonça (2010).

Quando o objetivo foi aplicar as equações ajustadas em uma base de dados com parcelas diferentes das utilizadas no ajuste (validação), o modelo fixo de Mendonça (2010) (alternativa 7) foi superado pelo modelo misto de Clutter (1963), com classe de idade atual (com amplitude de 2 meses) escolhida como grupo de efeitos aleatórios, com inclinação na altura média na idade futura e logaritmo do volume na idade atual, em todos os critérios avaliados.

Com os resultados obtidos para as condições em que foi desenvolvido este estudo, é possível afirmar que os modelos mistos apresentaram maior acurácia e menor tendenciosidade, quando comparados aos modelos de efeitos fixos. Já a escolha do nível adequado de modelagem dependerá dos objetivos estratégicos, mas os resultados aqui alcançados demonstram que os modelos de árvore individual são uma opção mais acurada e menos tendenciosa para projeção, e oferecerem a possibilidade de obtenção dos resultados dos modelos de distribuição de diâmetros e povoamento.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ARIAS-RODIL, M.; CASTEDO-DORADO, F.; CÁMARA-OBREGÓN, A.; DIÉGUEZ-ARANDA, U. Fitting and calibrating a multilevel mixed-effects stem taper model for *Maritime Pine* in NW Spain. **PloS one**, v. 10, n. 12, p. e0143521, 2015. Public Library of Science. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143521>>. Acesso em: 31/12/2015.

BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **CERNE**, v. 21, n. 1, p. 97–105, 2015. UFLA - Universidade Federal de Lavras. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-77602015000100097&lng=en&nrm=iso&tlng=pt>. Acesso em: 31/12/2015.

CALAMA, R.; MONTERO, G. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. **Silva Fennica**, v. 39, n. 1, p. 37–54, 2005.

CASTRO, R. V. O.; SOARES, C. P. B.; MARTINS, F. B.; LEITE, H. G. Crescimento e produção de plantios comerciais de eucalipto estimados por duas categorias de modelos. **Pesquisa Agropecuária Brasileira**, v. 48, n. 3, p. 287–295, 2013. SciELO Brasil.

CLUTTER, J. L. Compatible growth and yield models for loblolly pine. **Forest Science**, v. 9, n. 3, p. 354–371, 1963. Society of American Foresters.

FRAGA FILHO, C. V.; SIMIQUELI, A. P. .; ALTOE, W. . Fgmutils: Forest Growth Model Utilities. Disponível em: <<https://cran.r-project.org/web/packages/Fgmutils/index.html>>. Acesso em: 31/12/2015.

GONÇALVES, J. C. **Análise de risco aplicada à determinação da rotação econômica**, 2014. 89f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Lavras. Lavras, 2014.

ISLAM, M. N.; KURTTILA, M.; MEHTÄTALO, L.; HAARA, A. Analyzing the effects of inventory errors on holding-level forest plans: the case of measurement error in the basal area of the dominated tree species. **Silva Fennica**, v. 43, n. 1, p. 71–85, 2009.

Finnish Society of Forest Science.

KANKARE, V.; LIANG, X.; VASTARANTA, M.; et al. Diameter distribution estimation with laser scanning based multisource single tree inventory. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 108, p. 161–171, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0924271615001902>>. Acesso em: 27/11/2015.

MARTINS, F. B. **Modelagem de crescimento em nível de árvore individual para plantios comerciais de eucalipto**, 2011. 143f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2011.

MARTINS, F. B.; SOARES, C. P. B.; SILVA, G. F. DA. Individual tree growth models for eucalyptus in northern Brazil. **Scientia Agricola**, v. 71, n. 3, p. 212–225, 2014. SciELO Brasil.

MENDONÇA, A. R. DE E. **Modelagem não linear do crescimento e da produção de plantações florestais**, 2010. 84f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras, 2010.

MÔRA, R. **Funções de afilamento de forma variável e modelagem de efeitos mistos em fustes de *Pinus taeda* e *Eucalyptus saligna***, 2015. 277f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2015.

PIETILÄ, I.; KANGAS, A.; MÄKINEN, A.; MEHTÄTALO, L. Influence of growth prediction errors on the expected losses from forest decisions. **Silva Fennica**, v. 44, n. 5, p. 829–843, 2010.

UZOH, F. C. C. C.; OLIVER, W. W. Individual tree diameter increment model for managed even-aged stands of ponderosa pine throughout the western United States using a multilevel linear mixed effects model. **Forest ecology and management**, v. 256, n. 3, p. 438–445, 2008. Elsevier. Disponível em: <<http://www.treesearch.fs.fed.us/pubs/40984#.VoV-xYIcjD4.mendeley>>. Acesso em: 31/12/2015.

VISMARA, E. DE S.; MEHTÄTALO, L.; BATISTA, J. L. F. Linear mixed-effects models and calibration applied to volume models in two rotations of *Eucalyptus grandis*

plantations. **Canadian Journal of Forest Research**, p. 132–141, 2015. NRC Research Press. Disponível em: <<http://www.nrcresearchpress.com/doi/abs/10.1139/cjfr-2014-0435#.VoWA8-JYWug>>. Acesso em: 31/12/2015.

ZHAO, L.; LI, C.; TANG, S. Individual-tree diameter growth model for fir plantations based on multi-level linear mixed effects models across southeast China. **Journal of Forest Research**, v. 18, n. 4, p. 305–315, 2013.

4. CONCLUSÕES GERAIS

O modelo de Mendonça (2010) foi o melhor modelo dentre os modelos de efeitos fixos avaliados, contudo, superado pelo o modelo misto de Clutter (1963) modificado, com classe de idade atual como grupo de efeitos aleatórios, com altura média da parcela na idade futura e logaritmo do volume na idade atual como inclinação aleatória (alternativa 11).

Os efeitos mistos trouxeram grande ganho de acurácia, especialmente para o nível de distribuição diamétrica.

A modelagem no nível de árvores individuais foi similar à modelagem no nível de povoamento.

Os efeitos mistos, em todos os níveis de modelagem, mostraram ser uma ferramenta estatística com potencial de ganhos significativos em acurácia e ausência de viés na modelagem do crescimento e da produção de povoamentos florestais implantados.

APÊNDICES

Apêndice A – Valores σ^2 , σ_ε^2 , função de máxima verossimilhança restrita (REML) (21) obtido para o exemplo 1

Iteração	σ^2	σ_ε^2	REML	Iteração	σ^2	σ_ε^2	REML
1	1,0000	1,0000	-15,43890352	58	3,6126	0,7159	-14,47735
2	1,0000	1,0000	-15,43890352	59	3,6106	0,7159	-14,47735
3	1,0010	1,0000	-15,43748236	60	3,6116	0,7169	-14,47735
4	0,9990	1,0000	-15,44032736	61	3,6116	0,7149	-14,47735
5	1,0000	1,0010	-15,4391772	62	3,6127	0,7159	-14,47735
6	1,0000	0,9990	-15,43863121	63	3,6137	0,7159	-14,47735
7	1,9821	0,8115	-14,67074411	64	3,6117	0,7159	-14,47735
8	1,9831	0,8115	-14,67042848	65	3,6127	0,7169	-14,47735
9	1,9811	0,8115	-14,67106022	66	3,6127	0,7149	-14,47735
10	1,9821	0,8125	-14,67101878	67	3,6127	0,7159	-14,47735
11	1,9821	0,8105	-14,67047242	68	3,6137	0,7159	-14,47735
12	2,3090	0,4974	-14,77868068	69	3,6117	0,7159	-14,47735
13	2,3100	0,4974	-14,77845855	70	3,6127	0,7169	-14,47735
14	2,3080	0,4974	-14,77890315	71	3,6127	0,7149	-14,47735
15	2,3090	0,4984	-14,77639902	72	3,6127	0,7159	-14,47735
16	2,3090	0,4964	-14,7809815	73	3,6127	0,7159	-14,47735
17	2,1061	0,6924	-14,62684674	74	3,9740	0,7159	-14,48125
18	2,1071	0,6924	-14,62657018	75	3,2514	0,7159	-14,48236
19	2,1051	0,6924	-14,62712371	76	3,7933	0,7159	-14,47838
20	2,1061	0,6934	-14,62664434	77	3,4321	0,7159	-14,47852
21	2,1061	0,6914	-14,6270548	78	3,7030	0,7159	-14,47761
22	2,3427	0,6911	-14,57211394	79	3,5224	0,7159	-14,47763
23	2,3437	0,6911	-14,57192144	80	3,6579	0,7159	-14,47741
24	2,3417	0,6911	-14,57230675	81	3,5675	0,7159	-14,47742
25	2,3427	0,6921	-14,57192997	82	3,6127	0,7875	-14,48847
26	2,3427	0,6901	-14,5723036	83	3,6127	0,6443	-14,49187
27	2,8745	0,7245	-14,50165847	84	3,6127	0,7517	-14,48031
28	2,8755	0,7245	-14,50158282	85	3,6127	0,6801	-14,48073
29	2,8735	0,7245	-14,50173428	86	3,6127	0,7338	-14,47811
30	2,8745	0,7255	-14,50168084	87	3,6127	0,6980	-14,47816
31	2,8745	0,7235	-14,50164081	88	3,6127	0,7248	-14,47754
32	3,2367	0,7112	-14,48289813	89	3,6127	0,7069	-14,47755
33	3,2377	0,7112	-14,48286694	90	3,9740	0,7875	-14,49275
34	3,2357	0,7112	-14,48292942	91	3,2514	0,6443	-14,49747
35	3,2367	0,7122	-14,48286862	92	3,7933	0,7517	-14,48145
36	3,2367	0,7102	-14,4829327	93	3,4321	0,6801	-14,48203
37	3,4889	0,7175	-14,47788797	94	3,7030	0,7338	-14,47840
38	3,4899	0,7175	-14,47787914	95	3,5224	0,6980	-14,47848

39	3,4879	0,7175	-14,47789688	96	3,6579	0,7248	-14,47762
40	3,4889	0,7185	-14,47789603	97	3,5675	0,7069	-14,47762
41	3,4889	0,7165	-14,4778848	98	3,6127	0,7159	-14,47735
42	3,5910	0,7153	-14,47736308	99	3,6131	0,7159	-14,47735
43	3,5920	0,7153	-14,47736162	100	3,6123	0,7159	-14,47735
44	3,5900	0,7153	-14,4773646	101	3,6127	0,7159	-14,47735
45	3,5910	0,7163	-14,4773623	102	3,6127	0,7158	-14,47735
46	3,5910	0,7143	-14,4773688	103	3,6129	0,7159	-14,47735
47	3,6287	0,7303	-14,47786023	104	3,6125	0,7159	-14,47735
48	3,6297	0,7303	-14,47786159	105	3,6127	0,7159	-14,47735
49	3,6277	0,7303	-14,47785894	106	3,6127	0,7158	-14,47735
50	3,6287	0,7313	-14,47793127	107	3,6128	0,7159	-14,47735
51	3,6287	0,7293	-14,4779374	108	3,6126	0,7159	-14,47735
52	3,5942	0,7166	-14,47735862	109	3,6127	0,7159	-14,47735
53	3,5952	0,7166	-14,47735741	110	3,6127	0,7158	-14,47735
54	3,5932	0,7166	-14,4773599	111	3,6127	0,7159	-14,47735
55	3,5942	0,7176	-14,47736419	112	3,6127	0,7159	-14,47735
56	3,5942	0,7156	-14,47735796	113	3,6127	0,7159	-14,47735
57	3,6116	0,7159	-14,47734613	114	3,6127	0,7159	-14,47735

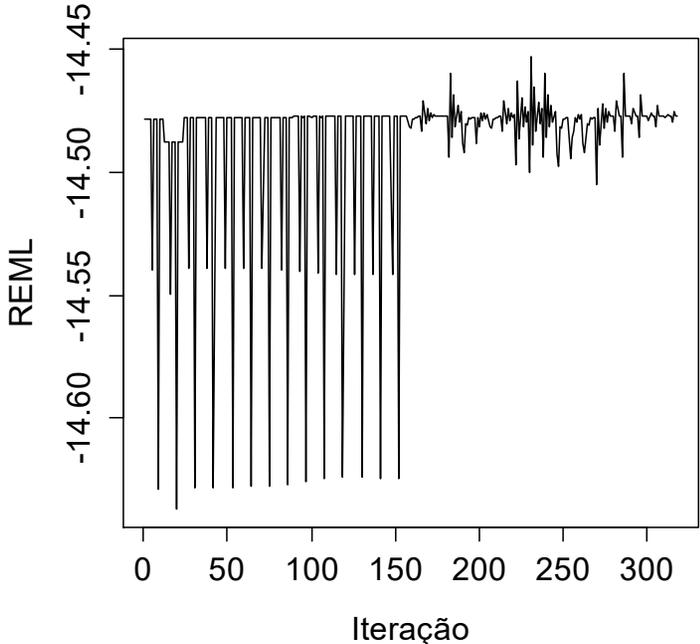
$$\hat{\mathbf{V}} = \begin{bmatrix} 4,328565 & 3,61271 & 3,61271 & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 0 & 0 & 0 \\ 3,61271 & 4,328565 & 3,61271 & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 0 & 0 & 0 \\ 3,61271 & 3,61271 & 4,328565 & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 0 & 0 & 0 \\ 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 4,328565 & 3,61271 & 3,61271 & 0 & 0 & 0 \\ 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 3,61271 & 4,328565 & 3,61271 & 0 & 0 & 0 \\ 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 6,06 \cdot 10^{-6} & 3,61271 & 3,61271 & 4,328565 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4,328565 & 3,61271 & 3,61271 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3,61271 & 4,328565 & 3,61271 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3,61271 & 3,61271 & 4,328565 \end{bmatrix}$$

Apêndice C – Matrizes \hat{R} e \hat{V} do exemplo 4 – Multinível com intercepto e inclinação

$$\mathbf{Z} = \begin{bmatrix} 0 & 0 & 1 & 23,5 & 0 & 0 & 1 & 23,5 & 0 & 0 \\ 0 & 0 & 1 & 23,5 & 0 & 0 & 1 & 23,5 & 0 & 0 \\ 0 & 0 & 1 & 23,5 & 0 & 0 & 1 & 23,5 & 0 & 0 \\ 1 & 30,5 & 0 & 0 & 0 & 0 & 1 & 30,5 & 0 & 0 \\ 1 & 30,5 & 0 & 0 & 0 & 0 & 1 & 30,5 & 0 & 0 \\ 1 & 30,5 & 0 & 0 & 0 & 0 & 1 & 30,5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 37,5 & 0 & 0 & 1 & 37,5 \\ 0 & 0 & 0 & 0 & 1 & 37,5 & 0 & 0 & 1 & 37,5 \\ 0 & 0 & 0 & 0 & 1 & 37,5 & 0 & 0 & 1 & 37,5 \end{bmatrix}$$

$$\hat{\mathbf{V}} = \begin{bmatrix} 4,328595 & 3,612735 & 3,612735 & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 0 & 0 & 0 \\ 3,612735 & 4,328595 & 3,612735 & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 0 & 0 & 0 \\ 3,612735 & 3,612735 & 4,328595 & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 0 & 0 & 0 \\ 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 4,328595 & 3,612735 & 3,612735 & 0 & 0 & 0 \\ 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 3,612735 & 4,328595 & 3,612735 & 0 & 0 & 0 \\ 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 1,59 \cdot 10^{-13} & 3,612735 & 3,612735 & 4,328595 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4,328595 & 3,612735 & 3,612735 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3,612735 & 4,328595 & 3,612735 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3,612735 & 3,612735 & 4,328595 \end{bmatrix}$$

Figura 1 – Valores da função de máxima verossimilhança restrita para o exemplo 4.



Fonte: O autor