

JOSEILME FERNANDES GOUVEIA

**MODELOS VOLUMÉTRICOS MISTOS APLICADOS EM CLONES DE  
*EUCALYPTUS* NO POLO GESSEIRO DO ARARIPE - PE**

**RECIFE – PE  
OUTUBRO – 2013**



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**PRÓ-REITORA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA**  
**APLICADA**

**MODELOS VOLUMÉTRICOS MISTOS APLICADOS EM CLONES DE**  
***EUCALYPTUS* NO POLO GESSEIRO DO ARARIPE - PE**

Tese apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Doutor.

**Área de Concentração: Estatística Aplicada**

Orientador: Prof<sup>o</sup>. PhD. José Antônio Aleixo da Silva  
Coorientadores: Prof<sup>o</sup>. Dr. Rinaldo Luiz Caraciolo Ferreira

**RECIFE – PE**  
**OUTUBRO – 2013**

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA  
APLICADA**

**MODELOS VOLUMÉTRICOS MISTOS APLICADOS EM CLONES DE  
*EUCALYPTUS* NO POLO GESSEIRO DO ARARIPE – PE**

JOSEILME FERNANDES GOUVEIA

Tese julgada adequada para obtenção do título de doutor em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 25/10/2013 pela Comissão Examinadora.

Orientador:

---

Prof. PhD. José Antônio Aleixo da Silva (DCFL – UFRPE)

Banca Examinadora:

---

Prof. Dra. Tatijana Stosic – Membro Interno (DEINFO/UFRPE)

---

Prof. Dra. Claudia Regina Oliveira de Paiva Lima – Membro Externo (DE/UFPE)

---

Prof. Dr. João Agnaldo do Nascimento – Membro Externo (DE/UFPB)

---

Prof. Dr. Luiz Medeiros de Araújo Lima Filho – Membro Externo (DE/UFPB)

*Dedico, a todos que de alguma forma tornaram a realização deste trabalho possível. Em especial a meu filho, Arthur, a minha mãe Josélia e a minha esposa Cidia.*

## **AGRADECIMENTOS**

Primeiramente a Deus, pelo dom da vida.

Ao meu orientador professor PhD. José Antônio Aleixo da Silva por sua dedicação, competência e atenção dedicada no desenvolvimento desta tese.

A todos meus amigos do doutorado pela interação produtiva e harmoniosa durante nosso convívio.

Aos professores e funcionários do Departamento de Estatística e Informática pela convivência agradável durante esse período.

Muitas foram às pessoas que de alguma maneira contribuíram para a realização deste trabalho. A todas elas, meu sincero agradecimento. Todavia, agradeço de forma especial. A minha mãe, pessoa que sempre foi exemplo de coragem, amor, determinação, retidão e perseverança.

A minha esposa Cidia e ao meu filho Arthur que sempre me apoiaram para realização deste trabalho.

Aos colegas de curso pelo companheirismo e amizade, em especial a Patrícia, Macio, Jader, Dennis, Rozilda, Kleber e Rodrigo que foram amigos e companheiros durante o curso.

*Tudo o que um sonho precisa  
para ser realizado é alguém que  
acredite que ele possa ser  
realizado.*

**Roberto Shinyashiki**

## Resumo

A estimativa volumétrica de povoamentos florestais é uma das mais importantes ferramentas no manejo florestal sustentado, pois permite ações que garantam a produção contínua dos povoamentos. O presente estudo teve como objetivo estimar o volume dos clones de *Eucalyptus* utilizando modelos mistos. A base de dados foi proveniente do experimento de clones de *Eucalyptus*, localizado no Município de Araripina, no semiárido de Pernambuco. Foram selecionados os principais modelos volumétricos clássicos, sendo que de Schumacher e Hall (1933) foi utilizado como testemunha para comparação com os modelos mistos. O modelo não linear misto serviu como ferramenta bastante eficaz para modelagem do volume de clones de *Eucalyptus*, devido sua flexibilidade, mostrando-se superior ao ajuste do modelo não linear de Schumacher Hall, de acordo com os critérios estabelecidos, resultando, com isso na recomendação do seu uso. Deste modo, o uso desse modelo tende a contribuir para a redução dos custos do inventário, com maior precisão.

**Palavras-chave:** crescimento de *Eucalyptus*, modelos mistos, Pólo Gesseiro do Araripe.

## **Abstract**

The volumetric estimation of the forest is one the most important tools in the sustainable forest management because allow to plan actions that will guarantee the continuous production of the forest. The present study aimed to estimate the volume of *Eucalyptus* clones by using mixed models. The database was provided from an experiment of Eucalyptus clones, located in the town of Araripina, in the semiarid of Pernambuco. The major classical volumetric models were selected, and the Schumacher e Hall (1933) was used as a reference for comparison with the mixed models. The nonlinear mixed model served as a very effective tool for modeling the volume of Eucalyptus clones, due to its flexibility, being superior to the adjustment of the non-linear model by Schumacher Hall, according to the established criteria, resulting with that, the recommendation of its use. Therefore, the use of this model can contribute to a reduction of inventory costs improving the accuracy.

**Keywords:** *Eucalyptus* growth, mixed models, Plasterer industry of Araripe.



## LISTA DE FIGURAS

Figura 1: Planta baixa da parcela .....	35
Figura 2: Vistas interna e externa do experimento .....	36
Figura 3: Análise de resíduos gráficos para o modelo de Schumacher e Hall; parte (a) resíduos padronizados <i>versus</i> valores ajustados; parte (b) quantis da distribuição normal padrão <i>versus</i> resíduos padronizados; parte (c) erros percentuais <i>versus</i> índices das observações .....	46
Figura 4: Comparação entre os valores médios observados e estimados para os principais modelos volumétricos .....	47
Figura 5: Resíduos padronizados <i>versus</i> valores ajustados para o modelo não linear de efeitos mistos .....	55
Figura 6: Resíduos padronizados <i>versus</i> valores ajustados para o modelo não linear de efeitos mistos, adotando matriz ARH(1), com função de variância .....	57
Figura 7: Resíduos padronizados <i>versus</i> valores ajustados para o modelo não linear de efeitos mistos, adotando matriz ARH(1), com função de variância sem as observações 1094 e 1183 .....	59
Figura 8: Gráfico dos resíduos padronizados <i>versus</i> quantis da distribuição normal padrão .....	60
Figura 9: Gráfico dos quantis da distribuição normal padrão <i>versus</i> efeitos aleatórios .....	61
Figura 10: Comparação do crescimento em volume ao longo do tempo da árvore 1 .....	62

Figura 11: Comparação do crescimento médio em volume ao longo do tempo ..... 64

## LISTA DE TABELAS

Tabela 1 – Estimativas dos parâmetros, erros padrão e p-valores dos principais modelos volumétricos .....	44
Tabela 2 – Critérios de seleção para escolha do modelo clássico volumétrico ....	45
Tabela 3 – Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura diagonal na matriz de variância e covariância .....	49
Tabela 4 – Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura simétrica composta heterogênea na matriz de variância e covariância .....	50
Tabela 5 – Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura não estruturada na matriz de variância e covariância .....	50
Tabela 6 – Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura componente de variância na matriz de variância e covariância .....	50
Tabela 7 – Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura auto regressiva de primeira ordem na matriz de variância e covariância .....	51
Tabela 8 – Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura auto regressiva de primeira ordem heterogênea na matriz de variância e covariância .....	51

Tabela 9 – Seleção das estruturas da matriz de variância e covariância .....	52
Tabela 10 – Critérios de seleção para os efeitos aleatórios .....	54
Tabela 11 – Resultados da análise de variância e dos critérios de seleção do modelo utilizado na estimação volumétrica .....	56
Tabela 12 – Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura auto regressiva de primeira ordem heterogênea na matriz de variância e covariância e função de variância “ <i>Varpower</i> ” .....	56
Tabela 13 – Variação percentual das estimativas dos parâmetros do modelo de Schumacher e Hall com efeitos aleatório, sem as observações 1094 e 1183 .....	58
Tabela 14 – Critério de seleção dos modelos .....	63

## LISTA DE QUADRO

Quadro 1 – Modelo de Schumacher e Hall com efeitos aleatórios .....	53
---	----

# SUMÁRIO

	Página
<b>1. Introdução</b> .....	01
<b>2. Revisão de Literatura</b> .....	05
2.1 Polo Gesseiro do Araripe .....	05
2.2 Gênero <i>Eucalyptus</i> .....	07
2.3 Planejamento florestal .....	09
2.4 Volume da árvore .....	10
2.5 Modelos volumétricos .....	12
2.6 Modelos lineares mistos .....	15
2.6.1 Definição do modelo .....	17
2.7 Modelos não lineares de efeitos mistos .....	18
2.7.1 Estimação de $\beta$ e $\gamma$ .....	21
2.8 Estrutura da matriz de covariância .....	22
2.9 Teste da razão de verossimilhança e intervalo de confiança .....	30
2.10 Análise de resíduos .....	31
<b>3. Material e Métodos</b> .....	34
3.1 Descrição do experimento .....	34
3.2 Estimação dos volumes dos <i>Eucalyptus</i> .....	37
3.3 Escolha do modelo .....	38
3.3.1 Teste de Vuong para modelos distintos .....	39
3.4 Avaliação da qualidade de ajuste dos modelos .....	40
3.5 Modelagem da variância e autocorrelação dos resíduos .....	41

3.6 Teste de Shapiro-Wilk .....	42
3.8 Teste de Breusch-Pagan .....	43
<b>4. Resultados e Discussão .....</b>	<b>44</b>
4.1 Ajuste dos principais modelos volumétricos .....	44
4.2 Ajuste dos modelos mistos não lineares .....	48
4.3 Comparação entre os modelos .....	61
<b>5. Conclusões .....</b>	<b>65</b>
<b>Referências Bibliográficas .....</b>	<b>67</b>
<b>Apêndice .....</b>	<b>80</b>

## 1. Introdução

O Brasil vem sendo explorado em seus recursos florestais desde o início de sua colonização, e como eram abundantes, as populações os usavam sem restrições. Entretanto, com o passar dos tempos, o crescimento populacional e a urbanização os recursos florestais naturais começaram a se tornar escassos. Hoje, há problemas de suprimento energético tendo como consequência o desmatamento, aquecimento global, erosão, processos de desertificação, entre outros.

Deste modo, com o aumento do desenvolvimento dos setores sociais e industriais tem crescido a procura por recursos energéticos, que tem sido suprida principalmente pelas florestas nativas, com a procura da biomassa, que possui a vantagem de ser de fácil aquisição e menor custo. Assim, cresce o número de áreas desertificadas em ritmo acelerado, em razão da exploração irresponsável e sem planejamento dos recursos florestais (OLIVEIRA et al., 2006; BARROS et al., 2010).

O uso racional das florestas é chamado de manejo florestal sustentável, ou seja, é uma forma de atividade florestal que promove a colheita de árvores, com técnicas de mínimo impacto ambiental, pois o retirado da floresta corresponde ao seu ritmo de crescimento, gerando uma exploração contínua.

Pela falta de estudos aprofundados do ritmo de crescimento das florestas nativas, entre elas a vegetação da Caatinga, bem como pelas baixas taxas de incremento médio anual (IMA) dessa vegetação, a demanda da indústria que usa a madeira como fonte energética é, na maioria das vezes, muito superior à oferta de madeira proveniente da vegetação nativa, mesmo sob regime de manejo florestal sustentado. Com isso a taxa de reposição da vegetação é inferior à retirada e com o passar dos anos essa vegetação vai raleando cada vez mais, comprometendo a sustentabilidade ambiental e muitas vezes levando a sua extinção na área.

Por outro lado, em função da necessidade cada vez maior de material energético pela indústria em regiões onde a madeira é a principal fonte energética, o desmatamento ilegal passa a ser uma realidade, comprometendo mais ainda a reposição da vegetação.

Uma opção para atenuar essa devastação da vegetação nativa é o uso de florestas de rápido crescimento em áreas desmatadas, onde a vegetação nativa não



se regenera em função de sua alta biodiversidade de plantas ombrófilas e heliófilas que compartilham o mesmo espaço, quando a floresta se encontra em seu estado de clímax. Essas áreas se não forem ocupadas por alguma forma de vegetação, certamente serão transformadas em focos de pré-desertificação.

A região do Polo Gesseiro do Araripe, que produz cerca de 95% do gesso consumido no país, tem como principal fonte energética a madeira proveniente da Caatinga, conseqüentemente, pode em um futuro não muito distante, passar por uma crise energética, pois com uma demanda bem superior à oferta de madeira pela vegetação nativa, o problema tende a se agravar, caso medidas alternativas não sejam efetivadas em curto espaço de tempo.

Uma opção viável economicamente e ambientalmente é o plantio de florestas de rápido crescimento nessas áreas, pois em curto espaço de tempo podem oferecer material energético para a demanda da indústria, além de atenuarem a devastação da vegetação nativa, recuperarem solos e combaterem os processos erosivos. O que de certa forma diminui o êxodo rural.

Entre os gêneros que são considerados para florestas de rápido crescimento, os *Eucalyptus* estão entre as melhores opções para o reflorestamento, devido a sua alta taxa de crescimento, resistência e elevado nível de melhoramento genético em produtividade e qualidade (LIMA FILHO et al., 2012).

Enquanto que uma vegetação nativa, na região do Araripe, necessita de um período de quinze anos para completar seu ciclo de rotação, os *Eucalyptus* apresentam maiores produtividades em um período compreendido entre 5 e 7 anos de rotação.

O Polo Gesseiro do Araripe, localizado na microrregião de Araripina, semiárido pernambucano, é o principal consumidor de energéticos florestais em forma de lenha. Aproximadamente 94% da madeira da região é retirada sem o devido manejo florestal, ou seja, na maior parte das vezes, de forma ilegal. As indústrias de gesso consomem cerca de 496 mil m<sup>3</sup> de lenha por ano, porém esse consumo vem aumentando com o crescimento da indústria do gesso. Estima-se que a região do Polo consuma anualmente algo em torno de 589 mil m<sup>3</sup> de lenha, que é utilizada para diversos fins como fonte de energia (BARROS, et al., 2010; PNUD, 2012).

Baseado neste contexto, o volume constitui-se uma das melhores opções para avaliar o potencial produtivo disponível em um povoamento florestal. Essa

informação é fundamental para todas as atividades de planejamento, principalmente, para predizer a viabilidade da implantação dessas florestas, para fins energéticos na região do Polo Gesseiro do Araripe, sendo o volume individual da árvore base para se estimar todo o conteúdo lenhoso dos povoamentos florestais (FRANCO et al., 1998; MACHADO et al., 2000; THOMAS et al., 2006).

Normalmente, a estimativa volumétrica dos clones de *Eucalyptus* é baseada na relação diâmetro à altura do peito (DAP) e altura total da árvore (H). Essas informações são obtidas durante certo período nas mesmas árvores, caracterizando um estudo longitudinal.

Dados longitudinais estão presentes em diversas áreas de estudo. Esses dados são medidas repetidas, em uma ou mais variáveis, em ocasiões sucessivas na mesma unidade experimental, com o objetivo de verificar seu comportamento ao longo do tempo.

De acordo com Hauser et al. (2009), na análise desse tipo de dados é comum admitir correlação não nula entre as observações, o que acaba violando a suposição de independência entre as observações, exigindo a modelagem da estrutura para matriz de variância e covariância. Além disso, esses dados apresentam heterogeneidade de variâncias, típicos em ajuste de curvas de crescimento com medidas repetidas, análise de perfis, entre outros.

Uma alternativa para considerar dados correlacionados, e modelar explicitamente a estrutura de covariância é o uso de Modelos Mistos, na análise de dados, com medidas repetidas que permitem uma maior flexibilidade na modelagem da matriz de covariâncias, podendo ser aplicados a dados incompletos, desbalanceados ou irregulares, não obstante inclui-se análise univariada e multivariada (LAIRD e WARE, 1982), e também permite descrever o comportamento dos perfis por meio das curvas.

O modelo possibilita incluir em estrutura um conjunto de variáveis não observáveis, denominado de efeitos aleatórios, incorporando a variabilidade da árvore, de maneira a permitir, junto às variáveis observáveis, denominadas de efeitos fixos, o ajuste de um modelo mais consistente, uma vez que esse modelo é capaz de incorporar o comportamento individual de cada árvore. Geralmente, nas ciências florestais, os modelos mistos aparecem em problemas não lineares como crescimento em altura de *Eucalyptus* e avaliações genéticas (CALEGARIO et al., 2005; COSTA et al., 2005; MORA, 2006).

No ajuste dos modelos mistos é possível realizar a estimação volumétrica para cada árvore, além da estimação do volume médio. Essa estimação individual permitirá verificar melhor a variabilidade do modelo e a tomada de decisões.

A análise de resíduos tem a finalidade de verificar se as suposições de normalidade e homocedasticidade são atendidas, além de identificar possíveis observações que estejam influenciando as estimativas dos efeitos fixos e aleatórios.

Neste contexto, os modelos são de fundamental importância para o manejo sustentado, pois promovem o melhor uso dos recursos florestais e são essenciais na estimação do volume e produção de floresta.

Os dados da pesquisa foram provenientes da Estação Experimental do Instituto Agrônomo de Pernambuco, na Chapada do Araripe - PE, no projeto Módulo de Experimentação Florestal para o Polo Gesseiro do Araripe, que utilizou 15 clones do gênero *Eucalyptus* e teve início em março de 2002.

As análises estatísticas deste estudo foram realizadas no programa R Development Core Team versão 2.15.2.

O presente trabalho teve como objetivo estimar o volume dos clones de *Eucalyptus*, por meio dos modelos não lineares de efeitos mistos, e recomendar o seu emprego, servindo de base para o aprofundamento na escolha de clones de *Eucalyptus* na região.

## 2. Revisão de Literatura

### 2.1 Polo Gesseiro do Araripe

A região do Araripe é composta pelos municípios de Araripina, Bodocó, Cedro, Dormentes, Exu, Granito, Ipubí, Moreilândia, Ouricuri, Parnamirim, Santa Cruz, Santa Filomena, Serrita, Terra Nova e Trindade, e está localizada no semiárido Pernambucano. Essas cidades juntas correspondem a 18% da área total do Estado. É nessa região onde se encontra o Arranjo Produtivo Local (APL) do gesso, denominado de Polo Gesseiro do Araripe (SILVA, 2008/2009).

Segundo Sobrinho et al. (2003), no Brasil, a área territorial que possui as melhores condições para o aproveitamento econômico da gipsita é a região do Araripe, que está localizada na fronteira dos Estados do Piauí, Ceará e Pernambuco.

O grande destaque da Chapada do Araripe, com área de 11.972  $km^2$ , é o Polo Gesseiro do Araripe, cujas principais jazidas de gipsita estão nos municípios de Araripina, Ipubi, Trindade, Bodocó e Ouricuri, de onde saem 95% de todo o gesso consumido no Brasil (ALBUQUERQUE, 2002).

De acordo com o Departamento Nacional de Produção Mineral - DNPM (2011) a produção brasileira de gipsita foi de 2.750.000 toneladas, apresentando um crescimento de 17,1%, em relação a 2009. A produção de gipsita possui uma correlação muito alta com a construção civil, que nos últimos anos vem tendo grande crescimento. Estima-se que o faturamento total do setor tenha chegado a R\$ 1,6 bilhão. Os principais setores da produção nacional são em ordem de importância: calcinação, gesso, cimento e gesso agrícola.

De acordo com Sindusgesso (2012) o Polo Gesseiro do Araripe cresceu 30% em 2010, em função da alta procura por gesso. Apesar de um grande número de empresas existente no setor, apenas seis empresas são responsáveis por cerca de 70% de toda produção nacional. A maior parte da exploração de gipsita, a matéria-prima do gesso, é retirada das jazidas do Polo Gesseiro e tratada em Araripina e Trindade, cidades mais desenvolvidas da microrregião.

No Polo Gesseiro do Araripe são gerados milhares de empregos diretos e indiretos, em decorrência das atividades de 39 minas de gipsita, 139 indústrias de calcinação e cerca de 730 indústrias de pré-moldados, que juntas movimentam cerca de U\$\$ 364 milhões/ano (SINDUSGESSO, 2012).

Em 2008, a produção anual de gesso em Pernambuco foi de 5,5 milhões de toneladas, em que 61% foram destinados à produção de blocos e placas, 35% para revestimento, 3% para moldes cerâmicos e 1% para outros fins. O Polo Gesseiro do Araripe produz, além disso, produziu cerca de 800 mil toneladas de gipsita usada pela indústria de cimento, e 200 mil toneladas de gesso agrícola. Pernambuco possui uma reserva de gipsita estimada em 261.395.640 de toneladas, sendo umas das mais expressivas e importantes do mundo (DNPM, 2012).

Segundo Silva (2008/2009) o processo de calcinação da gipsita é a etapa em que o minério é submetido a altas temperaturas para produção do gesso. A maioria da madeira utilizada no processo de calcinação é proveniente da Caatinga e vem das cidades circunvizinhas.

A matriz energética da indústria do gesso da região do Araripe é bastante diversificada, ela utiliza cerca de 4% de energia elétrica, 5% de óleo diesel, 8% de óleo BPF (baixo poder de fusão), 10% de coque e 73% de lenha (ATECEL, 2006). Porém as indústrias de pequeno porte utilizam, exclusivamente, a lenha como fonte energética. A lenha possui a vantagem de ser menos poluente e bem mais rentável economicamente do que outras fontes energéticas.

Devido à grande demanda de lenha exigida na desidratação da gipsita, o Polo Gesseiro do Araripe é de fundamental importância para microrregião de Araripina. Porém, essa região vem sendo drasticamente degradada, ocasionada pela crescente devastação da vegetação nativa, na maioria das vezes, de forma ilegal.

Há uma grande preocupação por parte dos empresários da indústria do gesso pelo desmatamento da região, provavelmente, não por causa do meio ambiente, mas por receio da escassez de combustível que pode provocar uma crise energética na região (SILVA, 2008/2009).

Visando atenuar o processo de devastação ambiental, o IBAMA, nos anos de 2006 e 2007, realizou um trabalho de identificação de empresas irregulares, que funcionavam sem licença ambiental, das quais apenas 15 empresas se adequaram às exigências do IBAMA. Em 2008 o IBAMA embargou 42 empresas, por não possuírem licença ambiental, e foram aplicados mais de R\$ 9,084 milhões em

multas (DP, 2008). No mesmo ano foi implementado o manejo florestal sustentável, para retirada da madeira de forma ambientalmente correta, visando, em certo prazo, um estoque razoável de lenha que viesse suprir a demanda do Polo. Além disso, garantir a sobrevivência da fauna e da flora. Mas em função do número de planos de manejo florestal sustentado e do baixo incremento médio anual da Caatinga em termos volumétricos, outras fontes energéticas se fazem necessárias para suprir a demanda da indústria do gesso. Entre elas, destacam-se as florestas de rápido crescimento.

## **2.2 Gênero *Eucalyptus***

A propagação da cultura de *Eucalyptus* ocorreu no início do século passado. Os países pioneiros na América do Sul a plantar tal gênero foram Chile, Argentina e o Uruguai. No Brasil não há registro de uma data exata, mas segundo Navarro de Andrade (1911) provavelmente foi em torno de 1868. Ele foi quem realizou o primeiro plantio experimental *Eucalyptus*, em 1909, no qual comparou diversas espécies nativas e os *Eucalyptus* se destacaram em relação às demais (NAVARRO DE ANDRADE, 1961; HASSE, 2006).

Conforme Lima Filho et al. (2012) os *Eucalyptus* vêm se destacando nos plantios florestais por causa do seu rápido crescimento, facilidade de reprodução, alto nível de adaptação e produtividade. Atualmente, esse gênero pode ser utilizado para diversos fins, tais como: produção de lenha, carvão, postes, celulose, móveis, etc. (FOELKEL, 2007; SOUZA et al. 2009).

Segundo Silva (2008/2009) entre as diversas espécies exóticas plantadas no Brasil, a que mais obteve sucesso foram as do gênero *Eucalyptus*. Existem, aproximadamente, mais de 700 espécies identificadas, pertencentes à família Myrtaceae, subfamília Leptospermoideae. Essas espécies foram originárias da Austrália, Tasmânia e Indonésia, apresentando uma variação latitudinal de ocorrência de 7° Norte à 43°39' Sul. São formadas por densas florestas, variando de arbustos a gigantes árvores, adaptadas, praticamente a todo tipo de solo e clima.

Possuem folhas de disposições alternas ou opostas e, às vezes, cruzadas com estípulas muito pequenas.

Segundo Garcia e Mora (2000), com a criação do IPEF - Instituto de Pesquisas e Estudos Florestais em 1968, os trabalhos de pesquisas envolvendo o melhoramento genético dos *Eucalyptus*, obtiveram grande avanço. Foi instalada uma ampla rede de ensaios, nas áreas de melhoramento genético e silvicultura, o que serviu de base para que o setor florestal brasileiro alcançasse níveis significativos de produtividade nos plantios comerciais. A partir de 1975, com a seleção de áreas para produção de sementes geneticamente melhoradas e aplicação dos resultados obtidos em testes de adubação, os plantios comerciais passaram a produzir florestas com incrementos de 30 a 40 m<sup>3</sup>/ha/ano, representando grande avanço silvicultural. A década de 80 foi marcada pela aplicação de técnicas de propagação vegetativa, cujos plantios clonais de *Eucalyptus* alcançaram incrementos de até 50 m<sup>3</sup>/ha/ano, devido à procedência das sementes que vieram da Austrália. Porém, na atualidade, os incrementos médios são em torno de 45 m<sup>3</sup>/ha/ano (RIPASSI et al., 2008).

O aumento da produtividade das florestas plantadas que são cultivadas atendendo o plano de manejo sustentável, tende a reduzir os impactos ambientais. Conseqüentemente quando essas florestas são bem planejadas, geralmente, são mais eficientes e econômicas (HARRISON et al., 2000).

Segundo Trugilho et al. (2001) e Scarpinella (2002), dentre diversas espécies arbóreas existentes os *Eucalyptus* vêm se destacando, e sendo amplamente utilizados em plantios florestais, por possuir os maiores índices de produtividade do mundo.

No Brasil os plantios de *Eucalyptus* possuem um caráter estratégico, pois sua madeira serve como principal fonte de abastecimento para o setor industrial, que consome matéria prima florestal, principalmente, em função da adaptabilidade desse gênero no Brasil que, conseqüentemente, coloca a indústria florestal brasileira no grupo de investimento de grande competitividade (RODRIGUEZ et al., 1997).

Outro fator que favorece para disseminação dos *Eucalyptus* é a clonagem, pois possibilita a manutenção das características das melhores árvores, obtendo talhões mais uniformes, de rápido crescimento, maior homogeneização da matéria prima a ser utilizada pela indústria, além de proporcionar material genético de melhor qualidade em curto espaço de tempo (FOEKEL e ASSIS, 1995, citados por FLORES et al., 2000; ALFENAS et al., 2004; ASSIS e MAFIA, 2007).

De acordo com o Anuário Estatístico da ABRAF (2012) o Brasil possui 6.515.844 hectares de florestas plantadas de *Eucalyptus* e *Pinus*, sendo 74,8% de *Eucalyptus* e 25,2% de *Pinus*. Em 2011, a área total de florestas plantadas de *Eucalyptus* e *Pinus* no Brasil praticamente não apresentou crescimento, pois o valor foi de apenas 0,1 % em relação ao total de 2010, considerado nula, já que fica dentro da margem de erro do levantamento. No período de 2005 a 2011 o crescimento médio anual de florestas plantadas foi de 27,9%.

### **2.3 Planejamento florestal**

As atividades de reflorestamento exercem importante papel no cenário socioeconômico do país, pois contribuem para produção de bens e serviços, agregação de valor aos produtos florestais, geração de empregos, tributos e rendas. Além de garantir a sustentabilidade em nível regional e nacional, devido ao tratamento responsável, em termos econômicos, ambientais e sociais, combate às mudanças climáticas, evitam ou reduzem a erosão do solo e sua contaminação, reduz o efeito estufa e aumenta o estoque sustentável de madeira legal (EMBRAPA, 2010).

A predição da produção em volumes de madeiras no reflorestamento é essencial no planejamento das atividades florestais, bem como o conhecimento do DAP e H, tornando-se elemento fundamental no manejo florestal (DAVIS e JOHNSON, 1987; GALVÃO, 2000).

De acordo com Berger (2000) a estimativa do crescimento de uma floresta, por meio de curvas de projeção do aumento em diâmetro, altura e volume é premissa indispensável em todo processo de planejamento racional. A determinação da idade ótima e econômica de corte de um povoamento é fundamental para o planejamento florestal, bem como a pesquisa silvicultural, que permite o controle do espaçamento, a intensidade de desbaste e a adubação (CAMPOS, et al., 1986; RESENDE, 2004).

É imprescindível que na obtenção das informações sobre as árvores evitem-se medições imprecisas, uma vez que pode ocasionar sérios erros nas estimativas,



e seja escolhido um profissional habilitado para as mensurações. Pois essas informações são utilizadas para negociação de serviços ou vendas de madeira em pé, e serve como base no inventário florestal (TONINI, 2004).

## **2.4 Volume da árvore**

Um dos principais problemas na realização dos inventários florestais é a estimativa volumétrica das árvores. Os métodos matemáticos são usados rotineiramente, na tentativa de otimizar a estimativa do volume de madeira, em uma unidade de área, geralmente um hectare. Inicialmente, consideram-se as árvores de uma espécie, em seguida podem-se formar subconjuntos com espécies que apresentam similaridades de forma e tamanho, e assim por diante (NETTO, 2004).

Os troncos de diversas espécies de árvores são semelhantes a determinados sólidos geométricos, os chamados protótipos dendrométricos. Dentre esses, os mais comuns são: o cilindro, o neilóide, o parabolóide e o cone (LOETSCH et al, 1973; AVERY e BURKHERT, 1997).

Segundo Machado, Urbano e Conceição (2005) o volume de uma árvore representa uma das informações de maior relevância para se avaliar o potencial produtivo florestal de uma região, sendo o volume individual o ponto de partida para se conhecer potencial que uma floresta guarda. O volume de uma árvore sofre influência de vários fatores, dentre eles, destaca-se o material genético (semente ou clone), características edafo-climáticas do sítio onde a árvore foi plantada e práticas de manejo adotadas.

O comércio de madeira e o fornecimento da grande maioria das indústrias do setor se desenvolvem com base no volume das árvores (AHRENS, 1982). Conforme Scolforo e Mello (1997) o uso de madeira proveniente do reflorestamento está aumentando cada vez mais, colocando-a em cenário de destaque nacional. Além disto, tem se observado outra tendência, que é uma maior preocupação com o planejamento e com métodos que busquem maior precisão nos inventários florestais.

A atividade de cubagem rigorosa do volume é realizada de maneira distinta das medições às unidades amostrais. Nos inventários florestais, o método mais utilizado é o que consiste na medição sucessiva de diâmetros ao longo do tronco, dividindo-o em seções que variam de comprimento, de acordo com o medidor (LEITE e ANDRADE, 2002).

Existem diversas fórmulas para estimação do volume na literatura, porém algumas têm se destacado, entre elas as de Huber, Smalian e Newton dados por:

i. Fórmula de Huber:

$$V = g_1 \cdot l$$

ii. Fórmula de Smalian:

$$V = \left( \frac{g_1 + g_2}{2} \right) \cdot l$$

iii. Fórmula de Newton:

$$V = \frac{1}{6} (g_1 + 4 \cdot g_{1/2} + g_2) \cdot l$$

Em que:

$V$  = volume da seção ( $m^3$ );

$g_1$  = a área seccional da base da tora ( $m^2$ );

$g_{1/2}$  = a área seccional no meio da tora ( $m^2$ );

$g_2$  = a área seccional no final da tora ( $m^2$ );

$l$  = o comprimento da tora ( $m$ ).

Esses métodos foram utilizados por (GOMES, 1957). O mesmo verificou que os processos de Huber e Newton conduzem a resultados diferentes. Já os processos de Huber e Smalian não apresentaram diferenças significativas entre as médias dos volumes calculados.

Machado et al. (2006) compararam os métodos de Huber, Newton, Smalian e Centróide com o volume real das árvores. Todos os tratamentos apresentaram diferenças significativas com relação ao volume real, tendendo a superestimar os volumes quando comparado ao volume verdadeiro. Dentre os métodos de cubagem, o que mais se aproximou ao valor exato foi o método de Huber.

A fórmula de Smalian tem sido a mais utilizada no Brasil nos levantamentos florestais, devido principalmente à facilidade de seu emprego do que por questões de precisão (MACHADO e FIGUEIREDO, 2006).

## **2.5 Modelos volumétricos**

De acordo com Davis e Johnson (1987) os modelos volumétricos, também chamados de modelos de produção, são utilizados rotineiramente nos inventários florestais para estimar a produção volumétrica de determinada área.

Após ajustados, os modelos geram equações. As equações volumétricas, são funções matemáticas que, em geral, relacionam uma variável de difícil obtenção com variáveis de fácil obtenção. Neste caso, o volume, com variáveis mais facilmente mensuráveis e conseqüentemente com menor custo, como o DAP, H total ou alguma medida que expresse a forma (MACHADO et al., 2008).

Na área florestal, normalmente, são utilizados os modelos de regressão para estimação dos volumes das árvores. As aplicações desses modelos, simples ou múltiplos, requerem que sejam satisfeitas as suposições de independência nas observações, normalidade dos erros e homocedasticidade dos erros (CLUTTER et al., 1983; SILVA e BAILEY, 1991).

Basicamente existem três tipos de modelos volumétricos:

- ✓ Os que utilizam apenas uma variável independente, DAP ou circunferência a altura do peito (CAP). Esses modelos podem ser adotados quando existe uma forte correlação entre as variáveis dependente e independente e são denominadas de locais;
- ✓ Existem os que usam duas variáveis independentes, DAP ou CAP com H. Este procedimento é o mais utilizado já que, normalmente, não existe uma forte relação hipsométrica. Esses modelos são denominados de regionais ou de dupla entrada;
- ✓ Os formais que utilizam três ou mais variáveis independentes, ou seja, além do DAP ou CAP e H, incluem também o fator de forma da árvore (FF) como variável independente. Este tipo de modelo é o menos utilizado, já que a forma da árvore é uma variável difícil de ser quantificada, e quando se faz cubagem rigorosa, esta já é controlada (SILVA, 1997).

É importante que a determinação do volume de uma árvore seja feita corretamente, pois as metodologias desenvolvidas para estimação do volume dessa poderão ser utilizadas para outras árvores de igual diâmetro, altura e forma (THIERSCH *et. al.*, 2006).

Os principais modelos utilizados para estimativa de volume de *Eucalyptus* são:

Modelo de Schumacher e Hall (1933)

$$vol_{ij} = \beta_0 \cdot DAP_{ij}^{\beta_1} \cdot H_{ij}^{\beta_2} \cdot \epsilon_{ij}$$

em que:

$vol_{ij}$  = volume da *i*-ésima árvore no *j*-ésimo tempo,  $i = 1, \dots, n$ , e  $j = 1, \dots, t$ ;

$DAP_{ij}$  = diâmetro a altura do peito da *i*-ésima árvore no *j*-ésimo tempo;

$H_{ij}$  = altura da *i*-ésima árvore no *j*-ésimo tempo;

$\epsilon_{ij}$  = é o erro aleatório.

A mesma notação é utilizada para os modelos seguintes com  $i = 1, \dots, n$ , e  $j = 1, \dots, t$ .

Modelo de Spurr (1952)

$$vol_{ij} = \beta_0 (DAP_{ij}^2 \cdot H_{ij})^{\beta_1} \cdot \epsilon_{ij}$$

Modelo da Variável Combinada de Spurr

$$vol_{ij} = \beta_0 + \beta_1 (DAP_{ij}^2 \cdot H_{ij}) + \epsilon_{ij}$$

Modelo de Honner

$$vol_{ij} = \frac{DAP_{ij}^2}{\beta_0 + \beta_1 \cdot \frac{1}{H_{ij}}} + \epsilon_{ij}$$

Modelo de Takata

$$vol_{ij} = \frac{DAP_{ij}^2 \cdot H_{ij}}{\beta_0 + \beta_1 \cdot DAP_{ij}} + \epsilon_{ij}$$

Existem diversos modelos volumétricos na literatura florestal e uso desses modelos nos inventários florestais tem se tornado cada vez mais frequente, tendo em vista que eles são capazes de prognosticar o crescimento e a produção das árvores (SPATHELF e NUTTO, 2000; SOARES et al., 2001).

Os modelos clássicos que apresentam estimativas mais precisas dos volumes são o de Schumacher e Hall (1933) e o de Spurr (1952) quando comparados com outros modelos volumétricos (NETO, 1977; CAMPOS et al., 1985).

Em outro estudo, (ROLIM et al., 2006), compararam os modelos volumétricos de Kopezky-Gehrhardt, Hohenadl-Krenn, Husch, Spurr, Variável Combinado de

Spurr, Stoate, Naslund, Meyer, Meyer modificada, Logarítmico de Spurr, Logarítmico de Prodan e Logarítmico de Schumacher e Hall. Para selecionar o melhor modelo foram usados o coeficiente de determinação, erro padrão da estimativa e distribuição dos resíduos. Entre os modelos, o que apresentou maior precisão na estimativa de volume de árvore foi o Logarítmico de Schumacher e Hall.

Em sua pesquisa (OLIVEIRA et al., 2009) quantificaram o volume de árvores em plantios clonais de *Eucalyptus* não desbastados pelo método da similaridade de perfis. Esse método apresentou resultado semelhante ao modelo de Schumacher e Hall.

Segundo McTague et al. (1989), muitas vezes, essas equações são específicas para uma determinada espécie, idade, ciclo e local e, obviamente, os custos envolvidos na obtenção de dados de cubagem rigorosa para construir equações são altos. Outra característica comum entre as equações de volume é que podem ser usadas somente para estimar o volume até um diâmetro fixo no topo. Em certos casos, porém, uma empresa pode mudar os limites de diâmetro superior de volume comercial, e isso obriga a elaboração de novas equações de volume comercial até o novo diâmetro mínimo.

## **2.6 Modelos lineares mistos**

Uma das preocupações estatísticas, ao analisar dados, é construir modelos que explicitem estruturas do fenômeno em observação, as quais frequentemente estão misturadas com variações aleatórias. A identificação dessas estruturas permite conhecer melhor o fenômeno, bem como fazer afirmações sobre possíveis comportamentos das mesmas.

A seleção adequada de modelos que sejam o mais simples possível, e que descrevem bem os dados observados que surgem em diversas áreas, é de fundamental importância em Agronomia, Biologia, Ecologia, Engenharia Florestal, entre outras.

Segundo Cordeiro et al. (2009) os principais modelos utilizados até o início da década de 1970 eram os modelos lineares e não lineares que se baseiam na

suposição de normalidade para variável resposta. Em 1972, surgem os modelos lineares generalizados que estendem a distribuição da variável resposta para a família exponencial propostos por Nelder e Wedderburn. Porém, esses modelos sempre utilizam a suposição de que as observações são independentes. Contudo é frequente a presença de situações em que os dados apresentam comportamentos dependentes entre si, ou seja, caso em que uma observação depende de outra. De maneira geral, são chamados de dados agrupados que incluem dados longitudinais, medidas repetidas e outras. Todavia, é foco de determinados estudos a análise da correlação entre grupos e intragrupos que está associado com a presença de estruturas de dados agrupados. Neste sentido, os modelos usuais, tanto lineares como não lineares, tornam-se inadequados, uma vez que a suposição de independência dos dados não pode ser verificada.

Na análise de dados longitudinais a variável resposta pode ser contínua (peso, ganho de peso, volume de árvore, consumo, idade entre outras), assim como, discreta (contagem de algum evento, presença ou ausência de alguma substância, etc.). As unidades experimentais, como árvores, plantas, animais, indivíduos e etc., podem ser classificadas em diferentes grupos, segundo um ou mais fatores (ou tratamento) como tipo de ração consumida, sexo, densidade de plantio, espaçamento entre linhas de plantio, etc. (BARBOSA, 2009).

As técnicas de ajuste de modelos têm permitido cada vez mais flexibilidade aos modelos permitindo a presença de dados agrupados, levando em consideração a possível correlação existente entre as observações, sendo possível modelar as correlações entre e intragrupos. Tais modelos são denominados mistos.

Os modelos mistos levam em consideração parâmetros de efeitos fixos ( $\beta$ ) e parâmetros de efeito aleatório ( $\gamma$ ). Incorporando a parte aleatória por meio da inclusão de uma matriz de variância e covariância (LITTELL et al., 1996). Diversas metodologias para ajuste de crescimento em dados longitudinais são discutidas em em Rao (1965); Grizzle e Allen (1969); Fearn (1975); Hui (1984); Harville (1997); Laird e Ware (1982).

Os modelos lineares mistos se fundamentam em três aspectos: estimação e teste de hipóteses sobre os parâmetros de efeito fixo; predição dos parâmetros de efeito aleatório; e estimação dos componentes de variância. A vantagem desse método está fortemente associada ao exame dos efeitos aleatórios e à modelagem da estrutura de variância e covariância (LAIRD e WARE, 1982; CAMARINHA FILHO,

2003; CUNHA et al., 2013). Segundo Marcelino e lemma (2000) a estimação dos parâmetros de efeitos fixos e a predição dos efeitos aleatórios dependem da estimação dos componentes de variância. Dessa forma, se a estimação dos componentes de variância for realizada sobre dados desbalanceados a estimativa encontrada será função da técnica utilizada.

### 2.6.1 Definição do modelo

Laird e Ware (1982) desenvolveram os modelos mistos para o  $i$ -ésimo grupo, abrangendo apenas um nível de agrupamento que é dado por:

$$Y_i = X_i\beta + Z_i\gamma_i + \varepsilon_i, \quad i = 1, \dots, n \quad (1.1)$$

$$\gamma_i \sim N(0, \Psi), \quad \varepsilon_i \sim N(0, \Sigma)$$

em que  $Y_i$  é o vetor ( $n \times 1$ ) de observações referente à variável resposta;  $\beta$  é vetor ( $p \times 1$ ) de parâmetros de efeitos fixos desconhecidos;  $X_i$  é a matriz ( $n \times p$ ) de covariáveis observadas, relacionada com os efeitos fixos;  $\gamma_i$  é o vetor ( $q \times 1$ ) de efeitos aleatórios desconhecidos;  $Z_i$  é a matriz ( $n \times q$ ) associada aos efeitos aleatórios; e  $\varepsilon_i$  é um vetor ( $n \times 1$ ) de erros aleatórios. Admite-se que  $\gamma_i$  e  $\varepsilon_i$  são não correlacionados.

De (1.1) se tem que:

$$E(Y) = X\beta$$

$$Var(Y) = V = Z\Psi Z' + \Sigma$$

em que  $\Psi$  é a matriz de covariância e variância dos efeitos aleatórios presente no vetor  $\gamma$  e  $\Sigma$  a matriz de variância e covariância residual. Assim, podem-se modelar



diversas estruturas de covariância para  $\Sigma$  (COSTA et al.,2012). Observe que, quando  $\Sigma = \sigma^2 I$  e  $Z = 0$ , o modelo misto reduz-se ao modelo linear padrão.

## 2.7 Modelos não lineares de efeitos mistos

Os modelos não lineares de efeitos mistos têm recebido grande atenção na literatura nos últimos anos devido à sua flexibilidade em trabalhar com dados que possuem medidas repetidas. Esses modelos permitem que alguns parâmetros sejam fixos e outros variando pelos grupos por meio dos efeitos aleatórios. Além disso, conseguem modelar a correlação existente entre as observações, dentro de um mesmo grupo, que é frequentemente observada nesse tipo de dados. Os modelos clássicos assumem independência das observações, logo não são adequados para dados agrupados, além de estimarem incorretamente as estimativas dos erros padrão nesse caso.

A modelagem dos modelos não lineares de efeitos mistos envolve questões semelhantes os modelos lineares mistos como:

- ✓ Escolher os efeitos fixos do modelo;
- ✓ Especificar estratégias a serem utilizadas na seleção dos efeitos aleatórios do modelo e de suas estruturas de covariância;
- ✓ Determinar qual estrutura de covariância que melhor representa a variabilidade dos dados;
- ✓ Verificar, por meio da análise de resíduos, possíveis afastamentos das suposições envolvidas na definição do modelo.

Em geral, o modelo não linear de efeitos mistos (LINDSTROM e BATES, 1990) para  $i$ -ésima árvore no  $j$ -ésimo tempo é dado por:

$$y_{ij} = f(\phi_{ij}, x_{ij}) + \varepsilon_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, t$$

em que  $y_{ij}$  é o vetor ( $n \times 1$ ) de volume observado da  $i$ -ésima árvore no  $j$ -ésimo tempo;  $x_{ij}$  é vetor ( $n \times p$ ) de covariáveis da  $i$ -ésima árvore no  $j$ -ésimo tempo;  $f$  é uma função diferenciável e não linear em pelo menos um dos componentes do vetor de parâmetros  $\phi_{ij}$ ; e  $\varepsilon_{ij}$  é o vetor ( $n \times 1$ ) de erros aleatórios das árvores, com distribuição normal, média zero e variância  $\sigma^2 \Lambda_i$ . O vetor de parâmetros varia de indivíduo para indivíduo. Em um segundo estágio, o vetor de parâmetros  $\phi_{ij}$  é modelado da seguinte forma:

$$\phi_{ij} = A_{ij}\beta + B_{ij}\gamma_i, \quad \gamma_i \sim N(0, \sigma^2 D),$$

em que  $\beta$  é vetor ( $p \times 1$ ) de parâmetros populacional de efeitos fixos;  $\gamma_i$  é o vetor ( $q \times 1$ ) de efeitos aleatórios desconhecidos;  $A_{ij}$  ( $n \times p$ ) e  $B_{ij}$  ( $n \times q$ ) são as matrizes de incidências dos efeitos fixos e aleatórios, respectivamente;  $\sigma^2 D$  é a matriz de variância e covariância positiva-definida (PINHEIRO, BATES e LINDSTROM, 1993).

Em geral, o modelo não linear de efeitos mistos pode ser expresso da seguinte forma  $f(A_i\beta + B_i\gamma_i, x_{ij})$ . Admite-se que  $\gamma_i$  e  $\varepsilon_i$  são independentes e que os  $\varepsilon_{ij} \sim N(0, \sigma^2 \Lambda_i)$ .

As matrizes  $A_i$  e  $B_i$  são usadas para simplificar o modelo. Os dados estão em dois grupos,  $A_i$  pode ser escolhido por permitir diferentes efeitos fixos em diferentes grupos (isto é,  $A_i = [I \ 0]$  para indivíduos no grupo 1 e  $A_i = [0 \ I]$  para indivíduos no grupo 2). Se os  $B_i$  são definidos todos iguais, ambos os grupos têm os mesmos efeitos aleatórios com mesma distribuição. Em outras situações, apenas alguns parâmetros são aleatórios.

Frequentemente, o modelo pode ser escrito da seguinte forma:

$$y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{in} \end{bmatrix}; \quad \eta_i(\phi_i) = \begin{bmatrix} f(\phi_i, x_{i1}) \\ \vdots \\ f(\phi_i, x_{in}) \end{bmatrix}; \quad \epsilon_i = \begin{bmatrix} \epsilon_{i1} \\ \vdots \\ \epsilon_{in} \end{bmatrix}.$$

Portanto,

$$y_i = \eta_i(\phi_i) + \epsilon_i,$$

em que  $\epsilon_i \sim N(0, \sigma^2 \Lambda_i)$ . Em algumas situações  $\sigma^2 \Lambda_i = \sigma^2 I_n$ , em que  $I_n$  é uma matriz identidade de ordem  $n$ , porém  $\Lambda_i$  pode assumir diversas estruturas de covariância marginal espacial, como por exemplo, a estrutura autorregressiva de primeira ordem. Usualmente o número de parâmetros em  $\Lambda_i$  é pequeno (LINDSTROM e BATES, 1990).

O modelo pode ser escrito individualmente, da seguinte forma:

$$y_i = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}; \quad \phi_i = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_n \end{bmatrix}; \quad \text{e} \quad \eta_i(\phi_i) = \begin{bmatrix} \eta_i(\phi_i) \\ \vdots \\ \eta_M(\phi_n) \end{bmatrix}$$

Sendo  $\tilde{D} = \text{diag}(D, D, \dots, D)$  e  $\Lambda = \text{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_n)$ . Logo, o modelo geral é expresso por:

$$y | \gamma \sim N(\eta(\phi), \sigma^2 \Lambda), \quad \phi = A\beta + B\gamma,$$

$$\gamma \sim N(0, \sigma^2 \tilde{D}),$$

em que,

$$B = \text{diag}(B_1, \dots, B_n), \quad \gamma = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix}, \quad A = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}.$$

Berkey (1982), Johansen (1984) e Racine-Poon (1985) propuseram alguns modelos de efeitos aleatórios, que são um caso especial do modelo anterior, em que  $A_i = B_i = I$ . Isto é, o valor esperado do modelo do  $i$ -ésimo indivíduo do vetor de resposta é  $\eta(\phi_i)$ , em que  $\phi_i \sim N(\phi, D)$ . Essa restrição entre os efeitos fixos e aleatórios é menos flexível para modelagem de dados agrupados.

### 2.7.1 Estimação de $\beta$ e $\gamma$ .

Os parâmetros do modelo são estimados pelos métodos da máxima verossimilhança, com base na densidade marginal de  $Y$ . Em geral, a densidade marginal do vetor de respostas  $y_i$  não possui forma fechada, quando se trata de modelo de função não linear em  $\gamma$ , exceto quando  $f$  é linear em  $\gamma$ . Diferentes aproximações têm sido propostas na literatura, para estimação dos parâmetros. Pinheiro e Bates (1994) analisaram diversas aproximações para os modelos não lineares de efeitos mistos e concluíram que as aproximações Lindstrom e Bates e da máxima verossimilhança, geralmente, apresentam resultados mais precisos.

Quando os componentes de variância das matrizes  $D$  e  $\Lambda$  são conhecidos e  $\eta$  é uma função linear de  $\beta$  e  $\gamma$ :

$$\eta(A_i\beta + B_i\gamma_i) = X_i\beta + Z_i\gamma_i.$$

Logo, pode-se definir

$$X = \begin{bmatrix} X_1^T \\ \vdots \\ X_n^T \end{bmatrix} \quad \text{e} \quad Z = \text{diag}(Z_1, Z_2, \dots, Z_n).$$

O estimador linear de mínimos quadrados generalizado de  $\beta$  é dado por:

$$\hat{\beta}_{lin} = \hat{\beta}_{lin}(\theta) = (X^T V^{-1} X)^{-1} X^T V^{-1} y$$

e o preditor linear de  $\gamma$  é expresso por:

$$\hat{\gamma}_{lin} = \hat{\gamma}_{lin}(\theta) = \tilde{D} Z^T \Sigma^{-1} [y - X \hat{\beta}_{lin}(\theta)],$$

em que  $\Sigma = \Lambda + Z \tilde{D} Z^T$  e  $\theta$  contém os elementos das matrizes de covariância de  $D$  e  $\Lambda$ . As estimativas  $\hat{\beta}_{lin}$  e  $\hat{\gamma}_{lin}$  maximizam conjuntamente a função:

$$g_{lin}(\beta, \gamma|y) = -\frac{1}{2\sigma^2} [y - (X\beta + Z\gamma)]^T \Lambda^{-1} [y - (X\beta + Z\gamma)] - \frac{1}{2\sigma^2} \gamma^T \tilde{D}^{-1} \gamma$$

que para  $\beta$  fixo é o algoritmo de densidade a posteriori de  $\gamma$  e para  $\gamma$  fixo é o logaritmo da verossimilhança para  $\beta$  (LINDSTROM e BATES, 1990).

No caso em que  $\eta$  é uma função não linear, a estimação é realizada de forma semelhante. O estimador de máxima verossimilhança do modelo não linear de efeitos mistos  $\hat{\beta}(\theta)$  e a moda a posteriori  $\hat{\gamma}(\theta)$  maximizam a função:

$$g(\beta, \gamma|y) = -\frac{1}{2\sigma^2} [y - \eta(A\beta + B\gamma)]^T \Lambda^{-1} [y - \eta(A\beta + B\gamma)] - \frac{1}{2\sigma^2} \gamma^T \tilde{D}^{-1} \gamma.$$

Para  $\beta$  fixo,  $g$  é uma constante mais o logaritmo da função de densidade a posteriori de  $\gamma$ . Assim, o parâmetro  $\gamma$  que maximiza  $g$  para um dado valor de  $\beta$ , será a moda a posteriori. Então,  $\hat{\beta}$  é um estimador de máxima verossimilhança relativo para uma distribuição marginal aproximada de  $y$  (LINDSTROM e BATES, 1990). Semelhante ao caso linear, esses estimadores podem ser obtidos resolvendo um problema de mínimos quadrados não linear que é obtido acrescentando-se o vetor de dados com “pseudo-dados”, como:

$$\tilde{y} = \tilde{\eta}(A\beta + B\gamma) + \tilde{\varepsilon},$$

em que

$$\tilde{y} = \begin{bmatrix} \Lambda^{1/2} y \\ 0 \end{bmatrix}, \quad \tilde{\eta}(A\beta + B\gamma) = \begin{bmatrix} \Lambda^{1/2} \eta(A\beta + B\gamma) \\ \tilde{D}^{-1/2} \gamma \end{bmatrix}, \quad e \quad \tilde{\varepsilon} \sim N(0, \sigma^2 I).$$

## 2.8 Estrutura da matriz de covariância

Uma etapa importante na modelagem dos modelos mistos é a escolha da estrutura de covariância, pois ela visa obter uma estrutura parcimoniosa, que

explique bem a variabilidade dos dados e a correlação entre as medidas com um número pequeno de parâmetros. Essa escolha pode afetar diretamente as estimativas dos parâmetros, os erros padrão de efeitos fixos e aleatórios, diagnósticos e inferências. Sua seleção irá depender das estruturas dos dados, de informações empíricas e da disponibilidade computacional.

As modelagens dessas estruturas para o modelo linear misto e não linear misto são basicamente idênticas. Mais detalhes podem ser encontrados em (PINHEIROS e BATES, 2000; DAVIDIAN e GILTINAN, 1995).

A seguir, apresentam-se quatorze estruturas que serão consideradas. Com  $n_i = 4$ , tem-se:

### 1. Componente da variância (CV)

$$\psi = \begin{vmatrix} \sigma_A^2 & 0 & 0 & 0 \\ 0 & \sigma_A^2 & 0 & 0 \\ 0 & 0 & \sigma_B^2 & 0 \\ 0 & 0 & 0 & \sigma_B^2 \end{vmatrix}$$

Nesta estrutura envolve apenas dois parâmetros, com duas variâncias distintas.

### 2. Não estruturada (NE)

$$\psi = \begin{vmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{vmatrix}$$

Atribui-se variâncias e covariâncias distintas para cada uma das  $n_i$  ocasiões. Sendo assim, com  $n_i(n_i + 1)/2$  parâmetros.

### 3. Simétrica composta (SC)

$$\Psi = \begin{vmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 \end{vmatrix}$$

Nesta estrutura as variâncias são iguais em todas as ocasiões de medidas e possuem mesmas covariâncias entre medidas feitas em ocasiões distintas e envolve dois parâmetros.

### 4. Auto regressiva de primeira ordem (AR(1))

$$\Psi = \sigma^2 \begin{vmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{vmatrix}$$

Esta estrutura possui dois parâmetros e aplicam-se variâncias iguais nas diversas situações e correlação decrescente com o acréscimo do intervalo entre as ocasiões.

### 5. Auto regressiva de primeira ordem heterogênea (ARH(1))

$$\Psi = \begin{vmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{vmatrix}$$

É uma generalização da estrutura AR(1), com variâncias e covariâncias desiguais. A estrutura ARH(1) somente é viável se os dados são igualmente espaçados no tempo e envolve  $n_i + 1$  parâmetros.

## 6. Simetria composta heterogênea (CSH)

$$\Psi = \begin{vmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{vmatrix}$$

Aplicam-se diferentes variâncias para cada elemento da diagonal principal, algumas covariâncias desiguais e coeficiente de correlação satisfazendo  $|\rho| < 1$ . Esta estrutura envolve  $n_i + 1$  parâmetros.

As estruturas CSH e ARH(1) permitem variâncias distintas ao longo da diagonal principal e são generalizações das estruturas básicas SC e AR(1). Para a estrutura CSH, o coeficiente de correlação  $\rho$  permanece constante, e para a estrutura ARH(1),  $\rho$  é um coeficiente auto-regressivo, o que significa que as correlações vão decrescendo exponencialmente (CASTRO, 1999).

## 7. Toeplitz (TOEP)

$$\Psi = \begin{vmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{vmatrix}$$

Nesta estrutura as covariâncias seguem um processo estocástico de médias móveis. Envolvendo  $n_i$  parâmetros.



## 8. Auto regressiva de primeira ordem médias móveis (ARMA(1,1))

$$\Psi = \sigma^2 \begin{vmatrix} 1 & \gamma & \gamma\rho & \lambda\rho^2 \\ \gamma & 1 & \gamma & \gamma\rho \\ \gamma\rho & \gamma & 1 & \gamma \\ \lambda\rho^2 & \gamma\rho & \gamma & 1 \end{vmatrix}$$

É uma estrutura relacionada aos dados de séries temporais com parâmetro auto regressivo  $\rho$ , componente de médias móveis  $\gamma$ , sendo  $\sigma^2$  a variância residual. De acordo com Castro (1999) esta estrutura é utilizada quando existem poucas unidades experimentais, número relativamente grande de observações por unidade experimental e essas observações são igualmente espaçadas no tempo, correspondendo a estrutura ARMA(1,1) é um caso particular dos modelos Auto Regressivos de Média Móveis. O número de parâmetros envolvidos é igual a quatro.

## 9. Toeplitz heterogênea (TOEPH)

$$\Psi = \begin{vmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \sigma_1\sigma_4\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \sigma_2\sigma_4\rho_2 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \sigma_3\sigma_4\rho_1 \\ \sigma_4\sigma_1\rho_2 & \sigma_4\sigma_2\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{vmatrix}$$

Nesta estrutura os dados são associados a séries temporais igualmente espaçados, com variâncias diferentes para cada elemento da diagonal, sendo as covariâncias funções de variâncias e do k-ésimo parâmetro da autocorrelação ( $|\rho| < 1$ ) para cada defasagem  $q - 1$ , e zeros para as últimas defasagens. Envolvem-se  $2n_i - 1$  parâmetros.

## 10. Ante-dependência de ordem 1 (ANTE(1))

$$\Psi = \begin{vmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \sigma_1\sigma_4\rho_1\rho_2\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \sigma_2\sigma_4\rho_2\rho_3 \\ \sigma_3\sigma_1\rho_2\rho_1 & \sigma_3\sigma_2\rho_2 & \sigma_3^2 & \sigma_3\sigma_4\rho_1 \\ \sigma_4\sigma_1\rho_3\rho_2\rho_1 & \sigma_4\sigma_2\rho_3\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{vmatrix}$$

Aplicam-se diferentes variâncias para cada elemento da diagonal, sendo os elementos fora da diagonal principal, funções de variâncias e do k-ésimo parâmetro de autocorrelação, satisfazendo ( $|\rho| < 1$ ). Envolvem  $2n_i - 1$  parâmetros.

Segundo Barbosa (2009) esta estrutura é aplicável a dados longitudinais em que as condições de avaliação não são igualmente espaçadas, além de apresentar heterogeneidade de variâncias e correlação serial.

## 11. Correlação sem estrutura (CSE)

$$\Psi = \begin{vmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} & \sigma_1\sigma_4\rho_{41} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & \sigma_2\sigma_4\rho_{42} \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ \sigma_4\sigma_1\rho_{41} & \sigma_4\sigma_2\rho_{42} & \sigma_4\sigma_3\rho_{43} & \sigma_4^2 \end{vmatrix}$$

Nesta estrutura a matriz de variância e covariância é especificada em termos da variância e correlações. Esta estrutura possui 10 parâmetros e se ajusta a modelos semelhantes ao tipo sem estrutura, mas com diferente parametrização. Sendo  $\sigma_i^2$  o i-ésimo parâmetro de variância e  $\rho_{jk}$  é a correlação entre a j-ésima e a k-ésima medida satisfazendo ( $|\rho| < 1$ ).

## 12. Espacial (SP(POW)(c-list))

$$\Psi = \sigma^2 \begin{vmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} \\ \rho^{d_{12}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} \\ \rho^{d_{13}} & \rho^{d_{23}} & 1 & \rho^{d_{34}} \\ \rho^{d_{14}} & \rho^{d_{24}} & \rho^{d_{34}} & 1 \end{vmatrix}$$

As covariâncias são funções da distância Euclidiana entre os vetores especificados pelas coordenadas. O parâmetro c-list associado a esta estrutura espacial, corresponde aos nomes das variáveis numéricas no conjunto de dados. Essas variáveis são usadas como coordenadas das observações no espaço. A estrutura de covariância espacial tem grande utilidade quando os dados longitudinais estão tomados no espaço e não no tempo, pois considera as distâncias euclidianas entre medidas repetidas, essa estrutura possui 2 parâmetros (CASTRO, 1999).

## 13. Diagonal principal (UN(1))

$$\Psi = \begin{vmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{vmatrix}$$

Esta estrutura é caracterizada por variâncias diferentes e covariâncias nulas possuindo nesse caso 4 parâmetros.

#### 14. Huynh-Feldt (H-F)

$$\Psi = \begin{vmatrix} \sigma_1^2 & \frac{(\sigma_1^2 + \sigma_2^2)}{2} - \lambda & \frac{(\sigma_1^2 + \sigma_3^2)}{2} - \lambda & \frac{(\sigma_1^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_2^2 + \sigma_1^2)}{2} - \lambda & \sigma_2^2 & \frac{(\sigma_2^2 + \sigma_3^2)}{2} - \lambda & \frac{(\sigma_2^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_3^2 + \sigma_1^2)}{2} - \lambda & \frac{(\sigma_3^2 + \sigma_2^2)}{2} - \lambda & \sigma_3^2 & \frac{(\sigma_3^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_4^2 + \sigma_1^2)}{2} - \lambda & \frac{(\sigma_4^2 + \sigma_2^2)}{2} - \lambda & \frac{(\sigma_4^2 + \sigma_3^2)}{2} - \lambda & \sigma_4^2 \end{vmatrix}$$

Esta estrutura é semelhante à CSH, pois cada elemento da diagonal principal possui variância diferente com mesmo número de parâmetros desconhecidos. Além disso, na estrutura H-F, cada elemento fora da diagonal é obtido pela média aritmética de duas variâncias, menos o parâmetro desconhecido. Essa estrutura envolve-se  $n_i + 1$  parâmetros.

Existem outras estruturas de variâncias e covariâncias. Aqui foram listadas algumas disponibilizadas nos softwares estatísticos. Diante disso, surge o problema: Qual melhor estrutura a ser testada em  $\Lambda_i$ ? Wald (2000) sugeriu que a estrutura geral de variância e covariância é a melhor opção para o modelo linear misto. Lima (1996) indicou o uso das estruturas NE e AR(1) para matriz dispersão  $\Psi$ . Vonesh e Chinchilli (1997) sugeriram que as estruturas mais comuns para matriz de covariância são a de componente de variância, simétrica composta e autorregressiva de primeira ordem.

Littel et al. (2000) usaram outras estruturas além dessas mencionadas neste estudo. Os mesmos sugerem que sejam utilizadas diversas estruturas de variância e covariância, e haja a escolha da estrutura mais parcimoniosa, por meio do critério de informação de Akaike (AIC). Em geral, as estruturas mais parcimoniosas são a simetria composta, não estruturada e autorregressiva.

## 2.9 Teste da razão de verossimilhança e intervalo de confiança

Para comparação dos modelos ajustados será utilizado o método da máxima verossimilhança proposto por Neyman e Pearson (1928). Esse método é baseado na razão de verossimilhança maximizada segundo ambas as hipóteses. Seja então, uma partição de vetor de parâmetros  $\theta$  dada por  $\theta = (\theta'_1, \theta'_2)'$ , no teste de  $H_0: \theta_1 = \theta_1^{(0)}$  versus  $H_A: \theta_1 \neq \theta_1^{(0)}$ . O teste da razão de verossimilhança é baseado na comparação dos logaritmos da função de verossimilhança maximizada  $l(\hat{\theta})$  e  $l(\tilde{\theta})$  sob a  $H_A$  (hipótese alternativa) e  $H_0$  (hipótese nula), respectivamente. De acordo com Pinheiro e Bates (2000) a estatística para esse teste é especificada por:

$$w = 2[l(\hat{\theta}) - l(\tilde{\theta})],$$

que tem, assintoticamente e sob certas condições de regularidade, distribuição  $\chi^2_q$  em que  $q$  é o número de graus de liberdade obtidos subtraindo o número de parâmetros do modelo sob  $H_A$  do número de parâmetros do modelo sob  $H_0$ . Rejeita-se  $H_0$ , se  $w \geq \chi^2_{q,1-\alpha}$ , em que  $\chi^2_q$  é o ponto crítico da distribuição Qui-quadrado correspondente ao nível de significância  $\alpha$  (NADARAJAH e KOTZ, 2005). De acordo com Pinheiro e Bates (2000) dado que os parâmetros foram estimados pelo método da máxima verossimilhança restrita, o teste da razão de verossimilhança só deve ser aplicado se ambos modelos forem ajustados pela mesma metodologia e se os efeitos fixos têm a mesma estrutura.

O intervalo de confiança para os parâmetros de efeitos fixos é dado por:

$$\hat{\beta}_j \pm t_{(1-\alpha/2)} [\hat{\Omega}_{ii}]^{1/2}$$

em que  $\hat{\beta}_j$  o estimador do  $j$ -ésimo efeito fixo,  $t_{(1-\alpha/2)}$  é o quantil da distribuição t-Student com  $(np - r)$  graus de liberdade, ao nível de  $100(1 - \alpha/2)\%$  de confiança; sendo  $r = \text{posto}(\hat{\Omega})$ , em que  $\hat{\Omega} = \widehat{\text{cov}}(\hat{\beta})$  é o estimador da matriz de covariâncias dos estimadores  $\hat{\beta}$ , que é dada por:

$$\hat{\Omega} = \left[ \sum_{i=1}^n J_i^T(\hat{\beta}) \hat{\Sigma}_i^{-1} J_i(\hat{\beta}) \right]^{-1}$$

em que  $J(\hat{\beta})$  é o jacobiano da função e  $\hat{\Sigma} = \widehat{cov}(y_i)$ ;  $\hat{\Omega}_{ii}$  é o i-ésimo elemento da matriz diagonal de  $\hat{\Omega}$  (BRANDÃO, 1996).

## 2.10 Análise de resíduos

Uma etapa importante no ajuste de qualquer modelo é análise de resíduos e diagnósticos, pois permite verificar possíveis afastamentos das suposições feitas para o modelo e auxilia na identificação de observações extremas, com alguma interferência desproporcional na estimativa dos parâmetros. Se as pressuposições do modelo são violadas, a análise resultante pode conduzir a resultados duvidosos.

Segundo Cordeiro e Demétrio (2008), na prática, porém, pode acontecer que, após uma escolha cuidadosa de um modelo e subsequente ajuste ao conjunto de observações, o resultado obtido seja insatisfatório. Isso pode ocorrer em função de algum desvio sistemático entre as observações e os valores ajustados ou, então, porque uma ou mais observações são discrepantes em relação às demais.

A análise de resíduos nos modelos mistos deve ser cuidadosa, pois cada modelo tem uma determinada estrutura de covariância. Assim, a estrutura dos resíduos pode variar de modelo para modelo. Nobre e Singer (2007) definem três tipos de resíduos, expressos a seguir:

Resíduos condicionais

$$\varepsilon = Y - \hat{Y};$$

Efeitos aleatórios

$$Z\gamma = E[Y|\gamma] - E[Y];$$

## Resíduos marginais

$$\xi = Y - X\beta = Z\gamma + \varepsilon.$$

Portanto, os valores estimados de  $\varepsilon$  e  $\xi$  são expressos por  $\hat{\varepsilon}$  e  $\hat{\xi}$  e definidos por:  $\hat{\varepsilon} = Y - X\hat{\beta} - Z\hat{\gamma}$  e  $\hat{\xi} = Y - X\hat{\beta}$ , em que  $\hat{\beta}$  é o melhor estimador não viesado de  $\beta$  e  $\hat{\gamma}$  é o melhor preditor não viesado de  $\gamma$ .

Segundo Pinheiro e Bates (2000) antes de qualquer inferência sobre os modelos mistos, as seguintes pressuposições devem ser verificadas: se os erros intra-grupos seguem aproximadamente distribuição normal com variância  $\sigma^2$  e se os erros são não correlacionados. Além disto, os efeitos aleatórios devem seguir uma distribuição normal e serem independentes para diferentes grupos.

Para verificação das pressuposições de homocedasticidade e normalidade do modelo podem ser realizados os seguintes procedimentos:

- a) Gráfico de probabilidade normal dos resíduos condicionais.

Esse gráfico é uma ferramenta de diagnóstico muito útil para avaliar a suposição de normalidade dos resíduos. A ocorrência de pontos próximos dispostos em uma reta indica que o modelo satisfaz tal suposição.

- b) Gráfico dos resíduos condicionais *versus* os valores ajustados.

Esse gráfico revela a homocedasticidade, ou seja, variância constante. O padrão para esse gráfico é uma distribuição aleatória de média zero e amplamente constante. Esse procedimento também pode ser utilizado para identificação de pontos discrepantes.

Nobre (2004) propôs uma padronização dos resíduos condicionais para o caso da violação de homocedasticidade definido por:

$$\hat{\varepsilon}_i^* = \frac{\hat{\varepsilon}_i}{\sigma\sqrt{q_{ii}}},$$

em que  $\hat{\varepsilon}_i$  é o  $i$ -ésimo elemento de  $\hat{\varepsilon}$  e  $q_{ii}$  o  $i$ -ésimo elemento da matriz  $Q$ , sendo esta definida por  $Q = M - MX(X'MX)^{-1}X'M$  e  $M = \sigma^2V^{-1}$ . Esse resíduo também pode ser utilizado para verificar a presença de heterocedasticidade, por meio do gráfico de resíduos padronizado condicionais versus valores ajustados.

c) Gráfico dos resíduos percentuais

O gráfico dos resíduos percentuais é utilizado para verificar as magnitudes dos desvios em termos percentuais, ou seja, o quanto as estimativas do modelo adotado se distanciam do valor real. Os resíduos percentuais são obtidos a partir da seguinte expressão

$$\hat{\varepsilon}_i(\%) = \left( \frac{Y_i - \hat{Y}_i}{\hat{Y}_i} \right) 100,$$

em que  $Y_i$  é o valor observado e  $\hat{Y}_i$  é o valor estimado.

d) Gráfico dos quantis da distribuição normal padrão *versus* efeitos aleatórios

O gráfico dos quantis da distribuição normal padrão *versus* efeitos aleatórios revela a normalidade dos efeitos aleatórios. Espera-se que os pontos estejam dispostos sobre uma reta para que a suposição de normalidade seja satisfeita.



### 3. Material e Métodos

#### 3.1 Descrição do experimento

A base de dados foi proveniente de um experimento clonal de *Eucalyptus*, localizado na estação experimental do Instituto Agronômico de Pernambuco (IPA) no Município de Araripina, no semiárido de Pernambuco. Possui como coordenadas geográficas de posição 07°29'00' S e 40°36'00' W e altitude de 816 metros. Segundo Thornthwaite (1955) apud Encarnação (1980), apresenta clima Dd'B'4<sup>a</sup> semiárido mesotérmico. A precipitação média anual é de, aproximadamente, 735 mm. A temperatura média anual é de 24<sup>o</sup> C com muita variação entre o dia e a noite. Solo é do tipo Latossolo vermelho-amarelo (LIMA FILHO et al. 2012).

O experimento foi implementado no início do mês de março de 2002, em uma área de 2,352 ha, com espaçamento de 3 x 2 m. Foram utilizados 15 tratamentos (clones) com 04 repetições, sendo utilizados 15 híbridos de *Eucalyptus*, estabelecidos em um delineamento multivariado de medidas repetidas.

As parcelas foram compostas de 49 árvores, com 25 árvores na área útil (Figura 1). Cada parcela foi de 14x21 metros com área de 294 m<sup>2</sup>. As mensurações foram realizadas a cada 6 meses, iniciadas para volume aos 12 meses de idade e a última aos 90 meses, data em que as árvores foram cortadas. Todas as árvores foram cortadas 15 centímetros acima do solo.

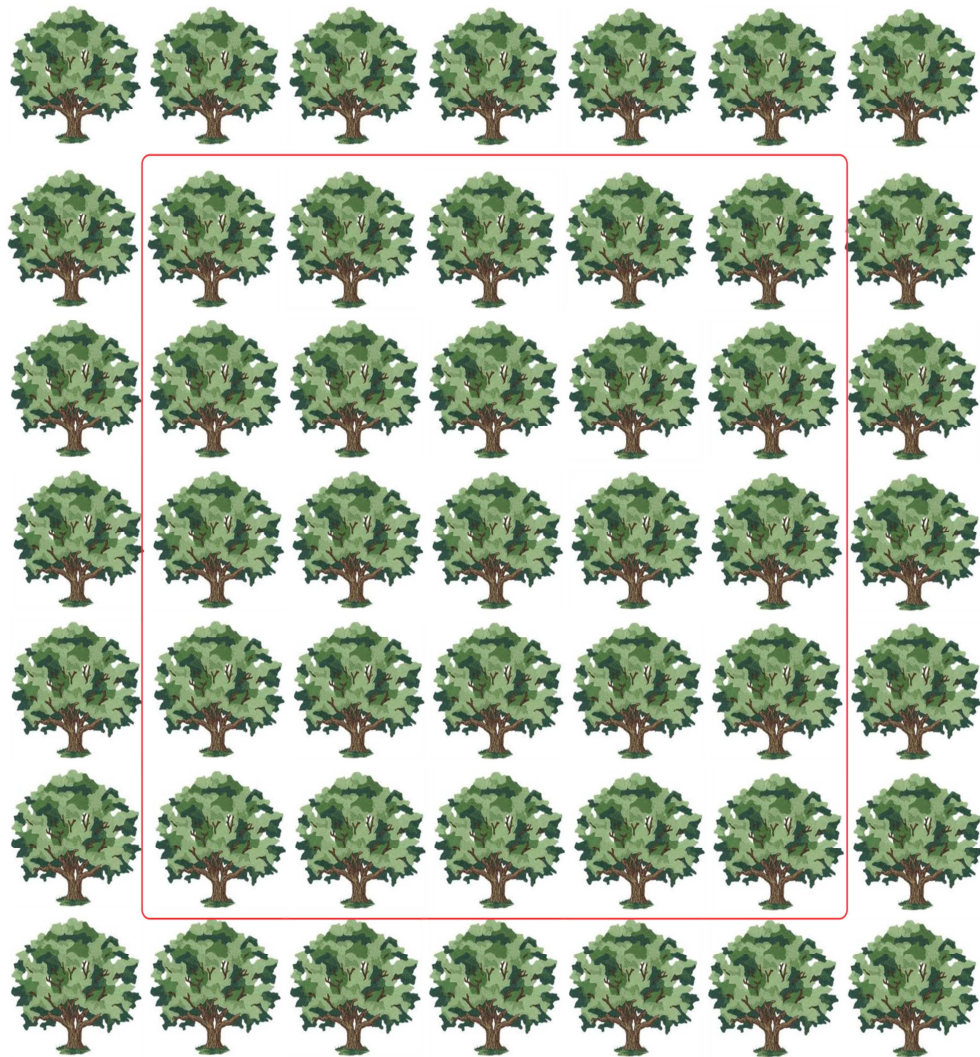


Figura 1 – Planta baixa da parcela.

Para realização desta pesquisa foi considerado o clone Híbrido de *Eucalyptus. urophylla* (cruzamento natural), devido o mesmo apresentar a maior produtividade (ROCHA, 2012). Nas árvores derrubadas foram medidas a altura e a circunferência a altura do peito. As alturas foram medidas com auxílio da trena graduada em centímetros. As CAPs foram tomadas com fita métrica, graduada em centímetro com aproximação em milímetro, conforme o procedimento descrito por SILVA (1977). A fórmula de Smalian (LOETSCH et al., 1973) foi empregada para calcular o volume rigoroso total e o volume das seções.

Na Figura 2 é possível observar as vistas interna e externa do experimento intitulado Módulo de experimentação florestal para a região do Araripe.





Figura 2: Vistas interna e externa do experimento.

### 3.2 Estimação do volume dos *Eucalyptus*

Os modelos mais clássicos para estimação volumétrica são o de Schumacher e Hall (1933) e o de Spurr (1952). Esses modelos se baseiam em funções que têm como variáveis independentes o DAP e H. De acordo com Scolforo et al. (2004) e Thomas et al. (2006), o modelo de Schumacher e Hall tem se destacado na estimativa do volume de árvores e de povoamento florestais e será usado como controle para comparação com modelos não lineares mistos.

O modelo de Schumacher-Hall é dado por:

$$vol_{ij} = \beta_0 DAP_{ij}^{\beta_1} H_{ij}^{\beta_2} \cdot \epsilon_{ij}$$

em que  $\beta_0$ ,  $\beta_1$  e  $\beta_2$  são os parâmetros a serem estimados,  $vol$  representa o volume em  $m^3$  e  $\epsilon_{ij}$  é o erro aleatório.

Para melhor estimação do volume dos clones de *Eucalyptus* será aplicado modelo não linear de efeitos mistos, que é expresso por:

$$f(A_i\beta + B_i\gamma_i, x_{ij}) + \epsilon_{ij}, \quad (2.2)$$

$$\gamma_i \sim N(0, \sigma^2 D) \text{ e } \epsilon_{ij} \sim N(0, \sigma^2 \Lambda_i).$$

em que  $\beta$  é o vetor de parâmetros populacional de efeitos fixos,  $x_{ij}$  é vetor de covariáveis da j-ésima resposta do i-ésimo indivíduo,  $\gamma_i$  é o vetor de efeitos aleatórios desconhecidos,  $A_i$  e  $B_i$  são as matrizes de incidências dos efeitos fixos e aleatórios, respectivamente,  $\sigma^2 D$  é a matriz de variância e covariância positiva-definida e  $\sigma^2 \Lambda_i$  a matriz de covariância de erros aleatórios. Admite-se que  $\gamma_i$  e  $\epsilon_i$  são independentes.

### 3.3 Escolha do modelo

A escolha do melhor modelo é obtida por meio do Critério de Informação de Akaike (AIC), teste da razão de verossimilhança e Índice de Ajuste Corrigido (IAC). De acordo com Akaike (1974) o AIC é baseado no valor da verossimilhança e no número de parâmetro e sua expressão é dada por:

$$AIC = -2\hat{l} + 2p,$$

em que  $\hat{l}$  é a log-verossimilhança maximizada,  $p$  é o número de parâmetros. Segundo este critério, o melhor modelo será o que apresentar o menor *AIC*.

A escolha do melhor modelo passa também por outras etapas, como selecionar a melhor estrutura de covariâncias, melhor estrutura para as médias, estimação dos parâmetros e comparação entre modelos. O programa R permite especificar a estrutura da matriz de covariância associada aos efeitos aleatórios, por meio do comando *random*.

Segundo Pinheiro e Bates (2000) a estatística para o teste da razão de verossimilhança é especificada por:

$$w = 2[l(\hat{\theta}) - l(\tilde{\theta})],$$

que tem, assintoticamente e sob certas condições de regularidade, distribuição  $\chi_q^2$ , em que  $q$  é a diferença entre a quantidade de parâmetros do modelo sob a hipótese alternativa e do modelo sob hipótese nula. Rejeita-se a hipótese nula, se  $w \geq \chi_{q,1-\alpha}^2$ , em que  $\chi_q^2$  é o ponto crítico da distribuição correspondente ao nível de significância  $\alpha$  (NADARAJAH e KOTZ, 2005).

Outra medida que também vem sendo utilizada na literatura para selecionar modelos é o Índice de Ajuste de Schlaegel (IA%) (SCHLAEGEL, 1981; MARTINS et al. 2007), expressa por:

$$IA = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2},$$

em que:

$Y_i$  = Volume observado;

$\hat{Y}_i$  = Volume estimado;

$\bar{Y}$  = Média dos volumes.

De acordo com Miguel et al. (2010) este índice indica a quantidade de variação que a variável resposta é explicada pelas variáveis independentes. Esse coeficiente varia entre 0 e 1, e quanto mais próximo de 1 (um) for o valor do  $IA$  em relação a unidade, melhor será o ajuste.

Entretanto, conforme aumentamos o número de variáveis explicativas o  $IA$  nunca diminui, ele sempre aumenta, ou permanece igual. Logo, é necessário reajustar esses coeficientes, para que seja possível comparar as equações, já que o número de parâmetros difere entre os modelos testados. Este reajuste recebe o nome de Índice de Ajuste Corrigido ( $IAC$ ) e sua expressão é dada por:

$$IAC = IA - (1 - IA) \left( \frac{n - 1}{n - p} \right)$$

em que,

$IA$  = índice de ajuste;

$n$  = número de observações;

$p$  = número de parâmetros.

### 3.3.1 Teste de Vuong para modelos distintos

Para comparação entre os modelos não aninhados foi utilizado o teste de razão de verossimilhança de Vuong (1989). Esse teste baseia-se em  $F_\theta$  e  $G_\theta$  sendo dois modelos diferentes, e  $f(y_i|x_i, \theta)$  e  $g(y_i|x_i, \beta)$  as respectivas densidades, a estatística desse teste é obtida por:

$$LR(\hat{\theta}, \hat{\beta}) \equiv l_f(\hat{\theta}) - l_g(\hat{\beta}) = \sum_{i=1}^n \log \frac{f(y_i|x_i, \hat{\theta})}{g(y_i|x_i, \hat{\beta})}.$$

Vuong (1989) propôs uma alternativa para trabalhar com esses modelos com base na divergência de Kullback-Leibler (KULLBACK-LEIBLER, 1951). O método baseia-se na distância entre cada modelo e o verdadeiro processo que gera os dados, ou seja,  $h_0(y_i, X_i)$ , a estatística do teste é dada por:

$$T_{RLNN} = \frac{1}{\sqrt{n}} \frac{LR(\hat{\theta}, \hat{\beta})}{\hat{\omega}^2}$$

em que

$$\hat{\omega}^2 = \frac{1}{n} \sum_{i=1}^n \left( \log \frac{f(y_i|x_i, \hat{\theta})}{g(y_i|x_i, \hat{\beta})} \right)^2 - \left[ \frac{1}{n} \sum_{i=1}^n \left( \log \frac{f(y_i|x_i, \hat{\theta})}{g(y_i|x_i, \hat{\beta})} \right) \right]^2$$

é um estimador para variância de  $\frac{1}{\sqrt{n}} LR(\hat{\theta}, \hat{\beta})$ .

A estatística tem assintoticamente

$$T_{LRNN} \xrightarrow{d} N(0,1)$$

sob

$$H_0: E \left[ \log \frac{f(y_i|x_i, \hat{\theta})}{g(y_i|x_i, \hat{\beta})} \right] = 0,$$

ou seja, os modelos são equivalentes. Ao nível de  $\alpha$  de significância, sendo  $Z_{\alpha/2}$  o valor crítico da distribuição normal padrão, rejeita-se a hipótese nula se  $|T_{RLNN}| > Z_{1-\alpha/2}$ .

### 3.4 Avaliação da qualidade dos ajustes dos modelos

Os modelos tiveram sua qualidade de ajuste avaliada pelo Erro Percentual Absoluto Médio (*EPAM*) definido por:

$$EPAM(\%) = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i},$$

em que,

$y_i$  = Valores observados;

$\hat{y}_i$  = Valores ajustados;

$n$  = número de observações.

O modelo que apresentar o menor valor *EPAM* é o que melhor explica os dados (LIMA FILHO et al., 2012).

### 3.5 Modelagem da variância e autocorrelação dos resíduos

Com o objetivo de retratar a realidade dos dados, foi utilizada a função de variância de “poder das covariantes”, expressa por “*varPower*” no software R, para modelar a heterocedasticidade existente nos dados. De acordo com Pinheiro e Bates (2009) o modelo de variância é dado por:

$$Var(\varepsilon_{ij}) = \sigma^2 |v_{ij}|^{2\delta},$$

com a seguinte função de variância correspondente:

$$g(v_{ij}, \delta) = |v_{ij}|^{\delta},$$

em que  $v_{ij}$  é um poder do valor absoluto da variância da covariante. Segundo Carvalho et al. (2011) o parâmetro  $\delta$  é irrestrito, ou seja, pode assumir qualquer valor real e assim pode modelar casos cuja variância aumenta ou diminui com o valor absoluto da variância da covariante.



### 3.6 Teste de Shapiro-Wilk

O teste de Shapiro-Wilk foi proposto em 1965, e tem se revelado um dos melhores testes para verificação da normalidade. A estatística do teste é expressa por:

$$W = \frac{b^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2},$$

em que  $x_{(i)}$  são os valores da amostra ordenados. Valores baixos de  $W$  indicam que os dados seguem uma distribuição normal (SHAPIRO e WILK, 1965). A constante  $b$  é dada por:

$$b = \begin{cases} \sum_{i=1}^{n/2} a_{n-i+1} \cdot (x_{(n-i+1)} - x_{(i)}), & \text{se } n \text{ for par,} \\ \sum_{i=1}^{(n+1)/2} a_{n-i+1} \cdot (x_{(n-i+1)} - x_{(i)}), & \text{se } n \text{ for ímpar,} \end{cases}$$

em que  $a_{n-i+1}$  são constantes geradas pelas médias, variâncias e covariâncias das estatísticas de ordem de uma amostra de tamanho  $n$  de uma distribuição Normal. As hipóteses do teste são as seguintes

$H_0$ : Os dados seguem distribuição normal

$H_1$ : Os dados não seguem distribuição normal

Rejeita-se  $H_0$  se  $W > W_\alpha$ , em que  $W_\alpha$  é o valor crítico da estatística de  $W$ .

### 3.7 Teste de Breusch-Pagan

O teste de Breusch-Pagan (GUJARATI, 1995) é um dos testes mais utilizados para verificação da suposição de homocedasticidade dos resíduos. A hipótese nula é de que as variâncias dos erros são iguais versus a hipótese alternativa de que nem todas as variâncias são iguais. A estatística desse teste é obtida por:

$$u_i = \frac{e_i^2}{(\sum_{i=1}^n e_i^2)/n}, \quad i = 1, \dots, n,$$

em que  $e_i^2$  são os resíduos. Após o cálculo de  $u_i$  é realizada a regressão entre  $u_i$  (variável resposta) e  $\hat{y}_i$  (variável explicativa). Posteriormente, a estatística do teste  $\chi_{BP}^2$  é obtida calculando a soma de quadrados da regressão de  $u_i$  sobre  $\hat{y}_i$  e dividindo o resultado por 2. Sob a hipótese nula, a estatística possui uma distribuição  $\chi^2$  com  $k - 1$  graus de liberdade, em que  $k$  é o número de parâmetro (SOUZA, 1998). Assim, rejeita-se  $H_0$  ao nível de significância  $\alpha$  quando:

$$\chi_{BP}^2 > \chi_{\alpha, k-1}^2.$$

No ajuste dos modelos não lineares foi utilizado o pacote nls (Non-linear Least Squares) do programa R, e no ajuste de modelos não lineares mistos o pacote (Non-linear Mixed-Effects Models), que necessita ser instalado.

## 4. Resultados e Discussão

Inicialmente, os resultados desta pesquisa serão apresentados em três seções: A primeira seção será realizado o ajuste com os principais modelos volumétricos, em que o melhor modelo será utilizado como testemunha para a comparação com os modelos não lineares mistos. Na próxima seção, ajusta-se o modelo não linear misto. E por último, será realizada a comparação entre todos os modelos.

### 4.1 Ajustes dos Principais Modelos Volumétricos.

Na Tabela 1 são apresentadas as estimativas dos parâmetros dos modelos analisados e seus respectivos erros padrão. Observa-se que todos os parâmetros foram significativos, com pequenos erros padrão.

Tabela 1: Estimativas dos parâmetros, erros padrão e p-valores dos principais modelos volumétricos.

Modelo	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
Schumacher e Hall	0,2986	2,0241	1,0811
	(0,0156)	(0,0045)	(0,0101)
	< 0,001	< 0,001	< 0,001
Spurr	0,3634	1,0201	----
	(0,0007)	(0,0018)	----
	< 0,001	< 0,001	----
Variável Combinada de Spurr	-0,0017	0,3632	----
	(0,0002)	(0,0007)	----
	< 0,001	< 0,001	----
Takata	2,868	-0,3840	----
	(0,038)	(0,245)	----
	< 0,001	< 0,001	----
Honner	-0,0193	3,1017	----
	(0,0017)	(0,0286)	----
	< 0,001	< 0,001	----

De acordo com tais critérios da Tabela 2, o modelo que obteve melhor resultado foi o de Schumacher e Hall, pois apresentou menores valores de AIC, *EPAM*(%) e maior *IAC*. Para confirmação do melhor modelo foi realizado o Teste de Vuong para modelos distintos tomando o modelo de Schumacher e Hall como hipótese alternativa em relação aos modelos restantes. Verificou-se que praticamente todas as hipóteses nulas foram rejeitadas, com a exceção do modelo Honner que não apresentou diferença significativa ao nível de 5% de significância. Porém, decidiu-se utilizar o modelo Schumacher e Hall, devido aos critérios de AIC, *EPAM*(%) e *IAC*.

Tabela 2 – Critérios de seleção para escolha do modelo clássico volumétrico.

Modelo	AIC	<i>EPAM</i> (%)	<i>IAC</i>	Teste	Vuong	p-valor
1 - Schumacher e Hall	-9740,199	4,4609	0,9948	----	----	----
2 - Spurr	-9705,137	4,8513	0,9944	2 vs 1	2,1643	< 0,027
3 - Variável Combinada de Spurr	-9663,271	10,062	0,9943	3 vs 1	4,6593	< 0,001
4 - Takata	-9676,910	4,7663	0,9943	4 vs 1	3,8128	< 0,001
5 - Honner	-9711,873	4,9149	0,9945	5 vs 1	1,7903	< 0,073

Na Figura 3 é possível observar a análise dos resíduos do modelo de Schumacher e Hall. No gráfico (A) (resíduos padronizados versus valores ajustados), observa-se que os pontos se encontram distribuídos em torno de zero, porém, com variância não constante, mas crescente do tipo funil. Além disso, existem algumas observações discrepantes.

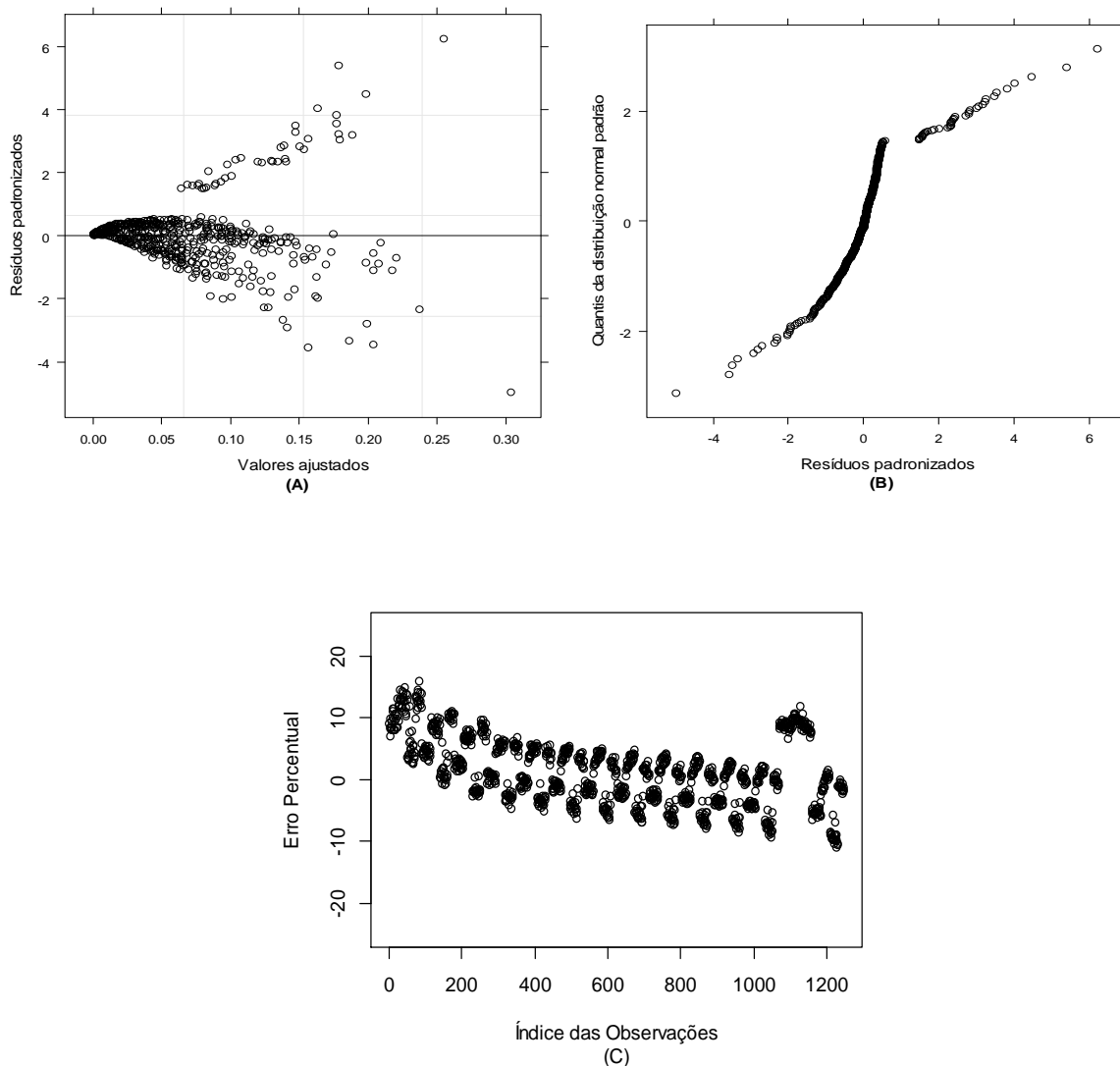


Figura 3: Análise de resíduos gráficos para o modelo de Schumacher e Hall; parte (a) resíduos padronizados *versus* valores ajustados; parte (b) quantis da distribuição normal padrão *versus* resíduos padronizados; parte (c) erros percentuais *versus* índices das observações.

Na parte (B) (quantis da distribuição normal padrão *versus* resíduos padronizado) da Figura 2, verifica-se que os dados não possuem aproximadamente distribuição normal. Foi realizado o teste de Shapiro-Wilk para confirmar se os dados possuem uma distribuição normal, o valor  $p$  foi  $< 0,001$ , ou seja, rejeita-se a hipótese nula de que os dados seguem uma distribuição normal. No terceiro gráfico, parte (C) (erros percentuais *versus* índices das observações) da Figura 2, percebe-

se que os erros percentuais estão próximos de zero e melhores distribuídos, além de mostrar quão discrepantes (ou não) são as estimativas.

Na Figura 4 é possível observar a comparação entre os valores médios observados e estimados para os principais modelos volumétricos, nota-se que os valores observados e estimados são muito próximos até o 60º mês, a partir desse mês em diante, até 90º surgem às discrepâncias. Mas, no geral, os modelos se ajustaram razoavelmente aos dados, com uma pequena vantagem para o modelo de Schumacher e Hall. Algumas curvas não aparecem, por estarem sobrepostas sobre as outras.

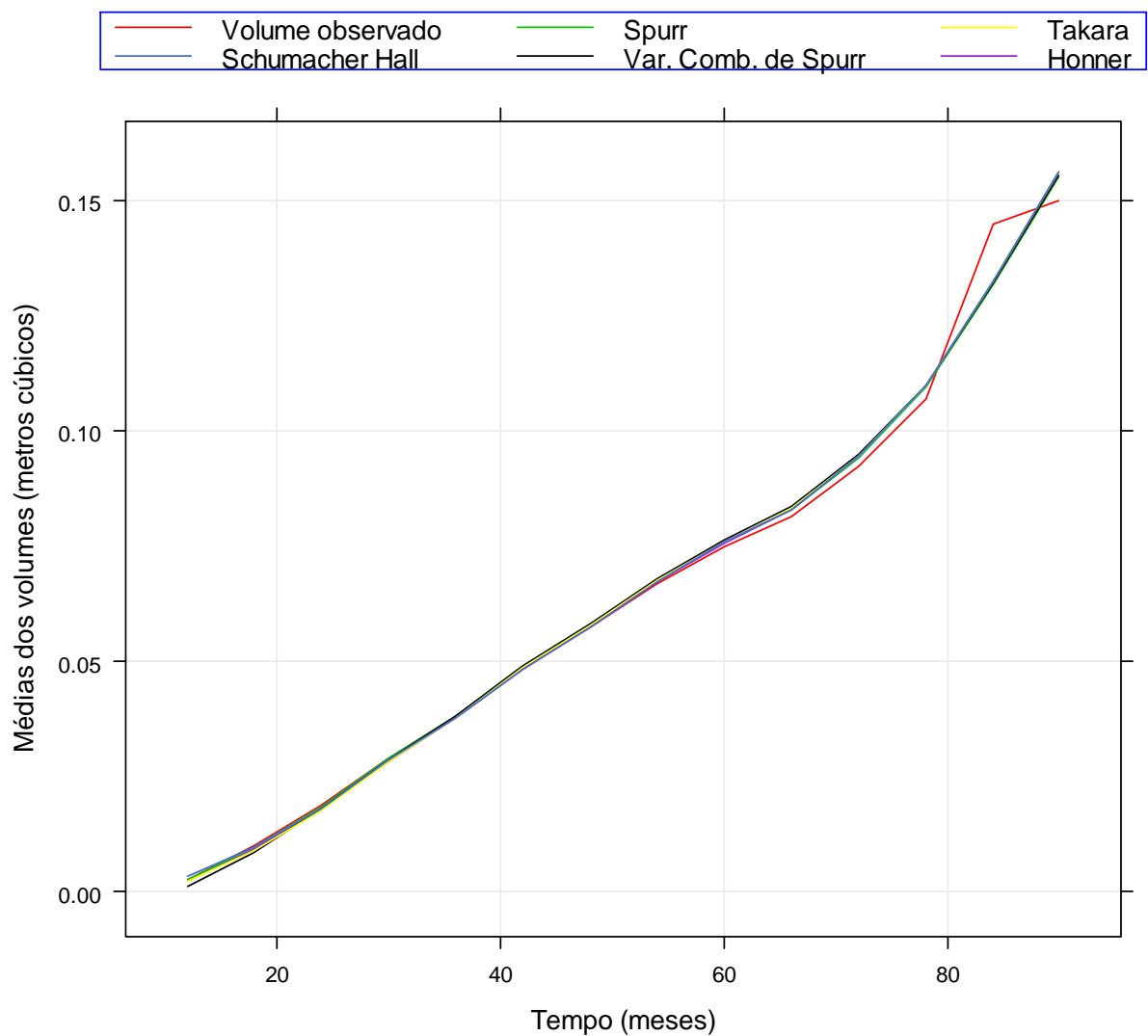


Figura 4: Comparação entre os valores médios observados e estimados para os principais modelos volumétricos.

Dessa forma, a equação do modelo de Schumacher e Hall ajustado à estimativa volumétrica de *Eucalyptus* é dada por:

$$\widehat{Vol}_i = 0,2986 \cdot DAP_i^{2,0241} \cdot H_i^{1,0811}.$$

Nota-se que tal modelo se ajustou razoavelmente aos dados, mesmo não satisfazendo as suposições de homocedasticidade e normalidade.

Silva et al. (2011) também utilizaram alguns desses modelos clássicos volumétricas na estimação volumétrica e o modelo de Schumacher e Hall foi o que apresentou melhor resultado, sendo a equação selecionada e recomendada para o povoamento de floresta Ombrófila Densa localizada na Fazenda Rio Capim, pertencente à Cikel Brasil Verde Madeiras Ltda., no município de Paragominas, estado do Pará. Guimarães e Leite (1996) concluíram que o modelo de Schumacher e Hall também foi o mais adequado para prever os volumes de *Eucalyptus* no Vale do Rio Doce, em Minas Gerais.

Entretanto, o modelo Schumacher e Hall não deve ser aplicado em inventários florestais que possuam medidas repetidas, por não satisfazer as pressuposições de normalidade e homocedasticidade.

#### **4.2 Ajuste dos modelos mistos não lineares**

Com o objetivo de estimar a produção volumétrica foi incorporado no modelo de Schumacher e Hall, a variabilidade existente em cada árvore. Com isso, temos o modelo não linear de efeitos mistos com parâmetros fixos e aleatórios.

Inicialmente, foi ajustado o modelo não linear de Schumacher e Hall misto com efeitos aleatórios nos três parâmetros pelo método da máxima verossimilhança e considerando diversas estruturas para a matriz de variâncias e covariância. A seguir serão apresentadas as estimativas dos parâmetros dos efeitos fixos e aleatórios pelo método da máxima verossimilhança para seis estruturas de variância e covariância. Outras estruturas também foram testadas, porém ora elas não

convergiram, ora não apresentaram matriz de variância positiva definida. Por esta razão, não foram apresentadas.

Ajustando o modelo não linear misto com todos os parâmetros aleatórios pelo método da máxima verossimilhança e considerando a matriz de variâncias e covariância diagonal na Tabela 3, verifica-se que todos os parâmetros fixos foram significativos e apresentando erros padrão pequenos. Os erros padrão dos efeitos aleatórios, em geral, foram pequenos e o valor para o parâmetro  $\gamma_2$  foi bem menor do que  $\gamma_0$  e  $\gamma_1$ , o que sugere que tal parâmetro pode ser considerado fixo. Isto será verificado posteriormente. Os erros padrão residuais dos efeitos fixos e aleatórios, em geral, foram baixos.

Tabela 3 - Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura diagonal na matriz de variância e covariância.

Parâmetros Estimados	Efeitos fixos				Efeitos aleatórios	
	Estimativa	Erro Padrão	Valor-t	Valor-p	Estimativa	Erro Padrão
$\beta_0$	0,2770	0,0073	37,8821	< 0,001	0,0057	0,0002
$\beta_1$	2,0225	0,0046	440,6183	< 0,001	0,0133	0,0051
$\beta_2$	1,1119	0,0092	120,6345	< 0,001	$2,7 \cdot 10^{-6}$	$3,210^{-9}$
Residual		0,0580			0,0022	

Em seguida, foi ajustado o modelo não linear misto pelo método da máxima verossimilhança considerando as estruturas de variância e covariância simétrica composta heterogênea, não estruturada, componente de variância, auto regressiva de primeira ordem e auto regressiva de primeira ordem heterogênea nas Tabelas 4, 5, 6, 7 e 8.



Tabela 4 - Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura simétrica composta heterogênea na matriz de variância e covariância.

Parâmetros Estimados	Efeitos fixos				Efeitos Aleatórios	
	Estimativa	Erro Padrão	Valor-t	Valor-p	Estimativa	Erro Padrão
$\beta_0$	0,2775	0,0200	13,8439	< 0,001	0,0633	0,0276
$\beta_1$	2,0492	0,0087	233,6054	< 0,001	0,0241	0,0130
$\beta_2$	1,1292	0,0285	39,6685	< 0,001	0,0875	0,0408
Residual		0,3009			0,0032	

Tabela 5 - Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura não estruturada na matriz de variância e covariância.

Parâmetros Estimados	Efeitos fixos				Efeitos Aleatórios	
	Estimativa	Erro Padrão	Valor-t	Valor-p	Estimativa	Erro Padrão
$\beta_0$	0,2753	0,0101	27,0970	< 0,001	0,0357	0,0091
$\beta_1$	2,0238	0,0033	609,4046	< 0,001	0,0089	0,0079
$\beta_2$	1,1188	0,0103	108,4122	< 0,001	0,0364	0,0128
Residual		0,0415			0,0023	

Tabela 6 - Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura componente de variância na matriz de variância e covariância.

Parâmetros Estimados	Efeitos fixos				Efeitos Aleatórios	
	Estimativa	Erro Padrão	Valor-t	Valor-p	Estimativa	Erro Padrão
$\beta_0$	0,2897	0,0087	33,2958	< 0,001	0,0348	0,0093
$\beta_1$	1,9870	0,0151	131,5845	< 0,001	0,0044	0,0014
$\beta_2$	1,1876	0,0423	28,0789	< 0,001	0,0121	0,0097
Residual		0,0415			0,0022	

Tabela 7 - Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura auto regressiva de primeira ordem na matriz de variância e covariância.

Parâmetros Estimados	Efeitos fixos				Efeitos Aleatórios	
	Estimativa	Erro Padrão	Valor-t	Valor-p	Estimativa	Erro Padrão
$\beta_0$	0,3359	0,0107	31,2025	< 0,001	0,0324	0,0116
$\beta_1$	2,0555	0,0057	358,8336	< 0,001	0,0082	0,0039
$\beta_2$	1,0542	0,0123	85,8080	< 0,001	0,0357	0,0132
Residual		0,2352			0,0021	

Tabela 8 - Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura auto regressiva de primeira ordem heterogênea na matriz de variância e covariância.

Parâmetros Estimados	Efeitos fixos				Efeitos Aleatórios	
	Estimativa	Erro Padrão	Valor-t	Valor-p	Estimativa	Erro Padrão
$\beta_0$	0,3456	0,0099	34,9090	< 0,001	0,0301	0,0098
$\beta_1$	2,0487	0,0047	435,8937	< 0,001	0,0073	0,0014
$\beta_2$	1,0425	0,0155	67,2580	< 0,001	0,0321	0,0105
Residual		0,2314			0,0021	

Em geral, todos os parâmetros fixos foram significativos e erros padrão dos efeitos fixos e aleatórios foram considerados baixos. As estruturas diagonal, simétrica composta, não estruturada e componente de variância apresentaram parâmetros fixos próximos. Já as estruturas auto regressiva de primeira ordem e auto regressiva de primeira ordem heterogênea apresentaram comportamento diferente em relação às demais, porém semelhante entre si. Em relação aos efeitos aleatórios, as estimativas dos parâmetros foram consideradas diferentes, com exceção da auto regressiva de primeira ordem e auto regressiva de primeira ordem heterogênea.

A Tabela 9 apresenta os critérios de seleção para diversas estruturas para matriz de variância e covariância, dentre essas a que melhor se ajustou à estimação

volumétrica dos *Eucalyptus* foi a autorregressiva de primeira ordem heterogênea, pois apresentou menor valor de AIC. Além disso, realizando o teste da razão de verossimilhanças e tomando tal estrutura como hipótese alternativa, comparou-se com os modelos restantes. Da Tabela 9 é possível observar que as hipóteses nulas foram praticamente todas rejeitadas, com exceção para estrutura autorregressiva de primeira ordem que não apresentou diferença significativa em relação à autorregressiva de primeira ordem heterogênea. Porém, decidiu-se utilizar a ARH(1) devido ao critério do AIC.

Tabela 9 – Seleção das estruturas da matriz de variância e covariância.

Modelo	AIC	LogLik	Teste	TRMV	p-valor
1. Diagonal	-11608,09	5811,044	1 vs 6	1555,58	< 0,001
2. Simétrica C. H. (CSH)	-11652,75	5813,135	2 vs 6	1502,92	< 0,001
3. Não Estruturada	-11639,34	5811,375	3 vs 6	1518,33	< 0,001
4. Comp. Var. (CV)	-11654,28	5815,254	4 vs 6	773,582	< 0,001
5. Auto Reg. AR(1)	-13153,91	6587,958	5 vs 6	1,175	<0,714
6. Auto. Reg. ARH(1)	-13155,67	6588,836	---	---	---

Após a escolha da melhor estrutura para matriz de covariância surge a seguinte questão: A variabilidade causada pela árvore está influenciando quais parâmetros? Tal resposta só será conhecida com os ajustes dos modelos combinando os efeitos aleatórios de acordo com Quadro 1. O modelo de Schumacher e Hall com três efeitos aleatórios, adotando estrutura auto regressiva de primeira ordem heterogênea na matriz de variância e covariância será tomado como referência e para efeito de análise os modelos foram definidos de 1 a 7.

Quadro 1 – Modelo de Schumacher e Hall com efeitos aleatórios.

Nº	Modelo
1	$vol_i = (\beta_0 + \gamma_{0i}).DAP_i^{(\beta_1+\gamma_{1i})}.H_i^{(\beta_2+\gamma_{2i})}.\epsilon_i$
2	$vol_i = (\beta_0 + \gamma_{0i}).DAP_i^{(\beta_1)}.H_i^{(\beta_2)}.\epsilon_i$
3	$vol_i = (\beta_0).DAP_i^{(\beta_1+\gamma_{1i})}.H_i^{(\beta_2)}.\epsilon_i$
4	$vol_i = (\beta_0).DAP_i^{(\beta_1)}.H_i^{(\beta_2+\gamma_{2i})}.\epsilon_i$
5	$vol_i = (\beta_0 + \gamma_{0i}).DAP_i^{(\beta_1+\gamma_{1i})}.H_i^{(\beta_2)}.\epsilon_i$
6	$vol_i = (\beta_0).DAP_i^{(\beta_1+\gamma_{1i})}.H_i^{(\beta_2+\gamma_{2i})}.\epsilon_i$
7	$vol_i = (\beta_0 + \gamma_{0i}).DAP_i^{(\beta_1)}.H_i^{(\beta_2+\gamma_{2i})}.\epsilon_i$

A Tabela 10 mostra os resultados dos ajustes dos 7 modelos por meio do método da máxima verossimilhança adotando estrutura auto regressiva de primeira ordem heterogênea na matriz de variância e covariância, tomando como referência os critérios de seleção dos modelos, ou seja, AIC e *EPAM* que indicam qual modelo deve ser selecionado. De acordo com tais critérios, o modelo que melhor se ajustou aos dados foi o 7 com efeitos aleatórios  $\gamma_0$  e  $\gamma_2$ , pois apresentou menor valor de AIC e *EPAM*. Além disso, foi realizado o teste da razão de verossimilhança tomando tal modelo com hipótese alternativa em comparação aos modelos restantes. Verificou-se que todas as hipóteses nulas foram rejeitadas, comprovando que o modelo de Schumacher e Hall com efeito aleatório  $\gamma_0$  e  $\gamma_2$  adotando estrutura auto regressiva de primeira ordem heterogênea na matriz de covariância foi o que melhor se ajustou aos dados.

Tabela 10 – Critérios de seleção para os efeitos aleatórios.

Modelo	AIC	EPAM(%)	Teste	TRMV	p-valor
1	-13155.67	3,8474	1 vs 7	8,092	< 0,015
2	-13003.48	4,1060	2 vs 7	170,28	< 0,001
3	-13121.56	3,9020	3 vs 7	52,204	< 0,001
4	-12985.17	4,2236	4 vs 7	188,598	< 0,001
5	-13132.18	3,8360	5 vs 7	37,58	< 0,001
6	-13128.16	3,8547	6 vs 7	41,602	< 0,001
7	-13175,76	3,8137	----	---	----

Uma das grandes vantagens dos modelos que incluem os efeitos aleatórios, com relação aos modelos de efeitos fixos é a redução do erro-padrão residual. Em (CALEGARIO et al., 2005) a redução foi de aproximadamente 15 vezes. Neste trabalho, a redução do erro padrão residual foi de 0,0591 para 0,0023, ou seja, uma diminuição de aproximadamente 26 vezes.

Dando sequência à análise do modelo, o próximo passo foi realizar análise de resíduos para verificar se as pressuposições de normalidade e homocedasticidade do modelo foram satisfeitas. Na Figura 5, observa-se a distribuição dos resíduos que indicam que eles estão aproximadamente distribuídos em torno de zero, porém não possuem variância constante, mas crescente do tipo funil, situação semelhante encontrada em Carvalho et al. (2011). Além disso, verifica-se algumas observações discrepantes.

Para confirmação da violação de homocedasticidade, foi realizado o teste de Breusch-Pagan, gerando um valor não significativo (valor de  $p < 0,001$ ), ou seja, rejeitamos a hipótese nula de que os resíduos possuem variâncias constantes.

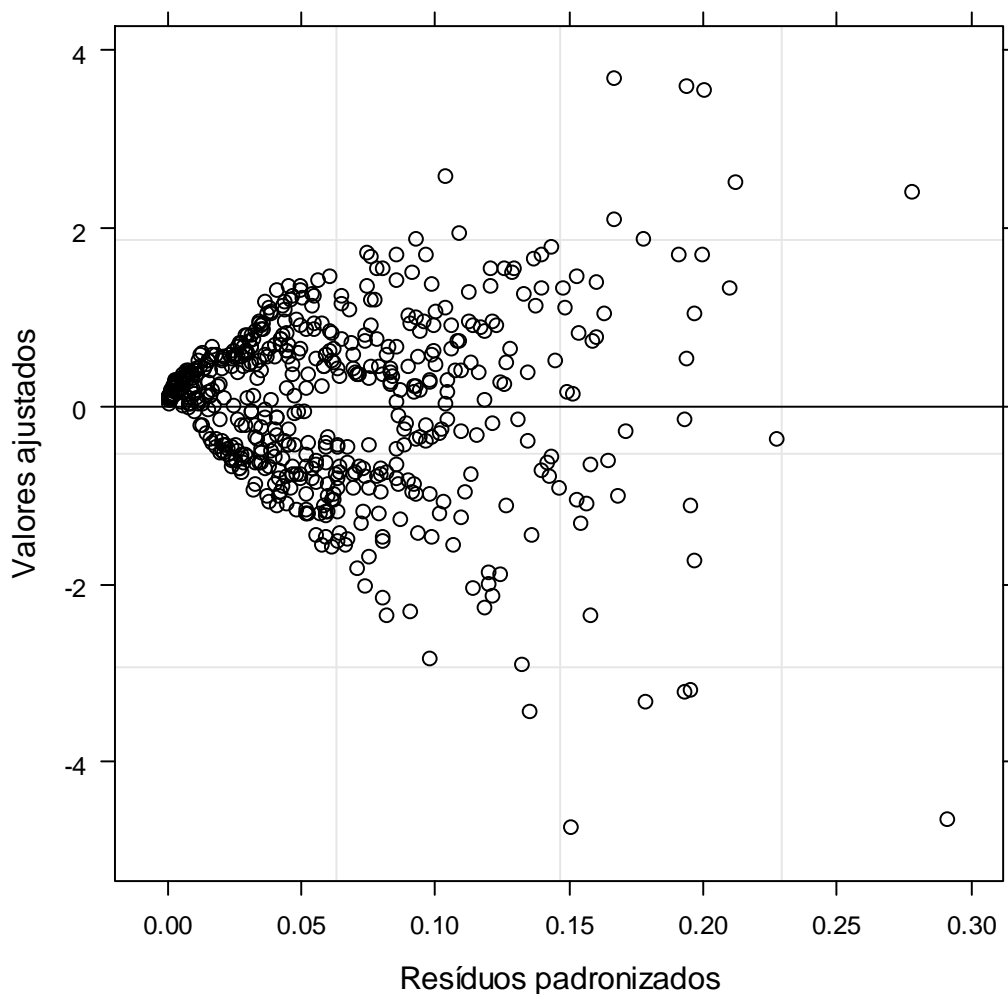


Figura 5 – Resíduos padronizados versus valores ajustados para o modelo não linear de efeitos mistos.

Para resolver o problema da heteroscedasticidade existente nos dados, foi utilizada a função poder de covariâncias que, no programa R, é expressa por “*varPower*”, com o objetivo de modelar a variância. A inclusão de tal função reduziu sensivelmente o erro padrão da estimativa do modelo de 0,243 para 0,021, bem como um aumento significativo na log-verossimilhança de 6592,882 para 7625,697, e redução no valor do AIC, como se pode observar na Tabela 11.

Carvalho et al. (2011) também fizeram uso de tais funções nos processos de modelagem, que tem a vantagem de trabalhar com dados originais, sem a necessidade de transformações, para que possam ser atendidas as pressuposições do modelo.

Tabela 11 – Resultados da análise de variância e dos critérios de seleção do modelo utilizado na estimação volumétrica.

Modelo	AIC	LogLik*	Teste	TRMV	p-valor
7*	-13175,76	6592,882	-	-	-
7**	-15233,39	7625,697	1 vs 2	2065,629	< 0,001

7\* = Modelo 7, adotando matriz ARH(1); 7\*\* Modelo 7, adotando matriz ARH(1) com função de variância.

Foi possível observar uma melhora significativa na precisão do modelo, que foi confirmado com o teste da razão de verossimilhanças, com valor de  $p$  ( $< 0,001$ ), ou seja, o modelo com função de variância ajustou-se melhor aos dados. Na Tabela 12 é possível observar as estimativas dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura autorregressiva de primeira ordem heterogênea, na matriz de variância e covariância e função variância “*Varpower*”. Percebe-se que todas as estimativas dos parâmetros fixos foram significativas, e para os parâmetros dos efeitos aleatórios  $\gamma_0$  e  $\gamma_2$  foram bastante pequenos, evidenciando que o modelo está sendo bem ajustado.

Tabela 12 - Estimativa dos parâmetros dos efeitos fixos e efeitos aleatórios pelo modelo não linear misto, adotando estrutura auto regressiva de primeira ordem heterogênea na matriz de variância e covariância e função de variância “*Varpower*”.

Parâmetros Estimados	Efeitos fixos				Efeitos Aleatórios	
	Estimativa	Erro Padrão	Valor-t	Valor-p	Estimativa	Erro Padrão
$\beta_0$	0,3641	0,0057	64,0469	< 0,0001	0,0113	0,0042
$\beta_1$	2,0114	0,0021	966,9601	< 0,0001	----	----
$\beta_2$	0,9944	0,0061	163,2323	< 0,0001	0,0153	0,0090
Residual		0,0154			0,0023	

Na Figura 6, observa-se que, com uso da função de variância, os resíduos apresentaram-se bem distribuídos com média em torno de zero, e variância constante, ou seja, não há afastamento da suposição de homogeneidade em relação às observações.

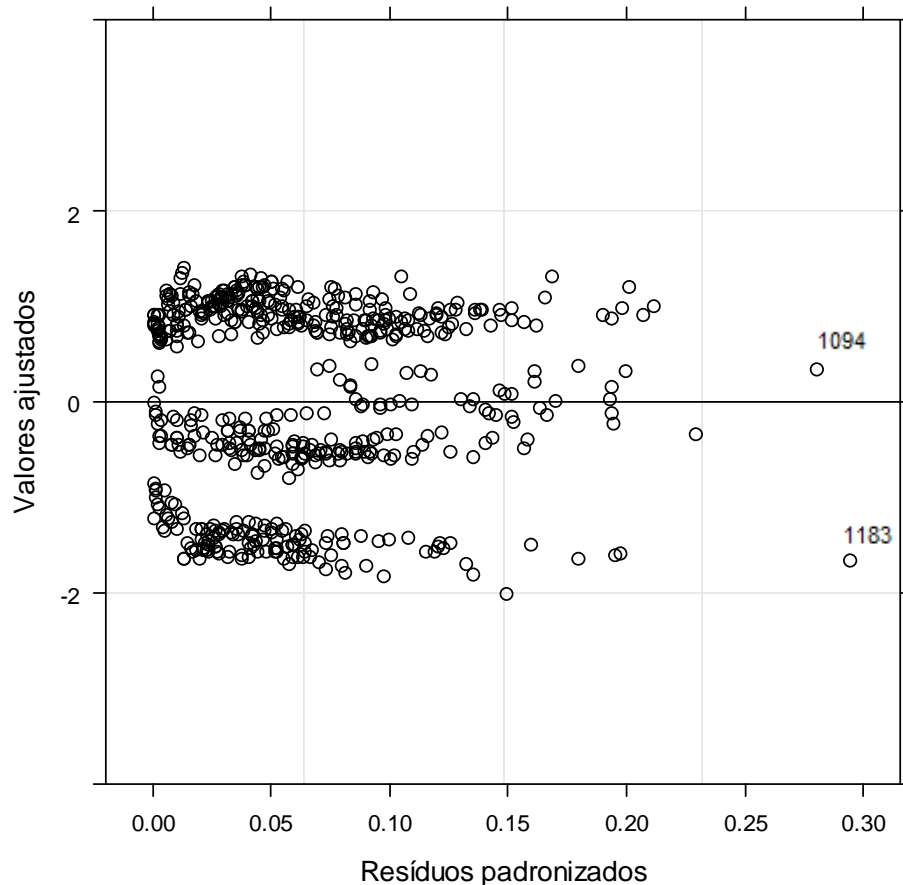


Figura 6 – Resíduos padronizados *versus* valores ajustados para o modelo não linear de efeitos mistos, adotando matriz ARH(1), com função de variância.

Decidiu-se ainda verificar a influência das duas observações que estão um pouco mais afastadas em relação às demais, ao lado direito na Figura 5. Elas correspondem às observações 1094 e 1183, para verificação de tal influência que as observações exercem sobre as estimativas dos parâmetros foi utilizada a seguinte medida

$$\text{Variação percentual} = \left( \frac{\hat{\beta}_* - \hat{\beta}}{\hat{\beta}} \right) \cdot 100$$



em que  $\hat{\beta}_*$  é a estimativa do parâmetro  $\beta$  sem a observação e  $\hat{\beta}$  é a estimativa do parâmetro  $\beta$  com todas as observações no modelo. Esta medida verifica, descritivamente, a variação percentual nas estimativas dos parâmetros na ausência de uma ou mais observações no ajuste do modelo.

A Tabela 13 apresenta as estimativas e as variações percentuais dos parâmetros do modelo de Schumacher e Hall com efeitos aleatório em  $\gamma_0$  e  $\gamma_2$ , sem as observações 1094 e 1183. Percebe-se que, em geral, a variação percentual entre as estimativas não foram grandes, a maior variação ocorreu no parâmetro de efeito aleatório  $\gamma_0$  com apenas 1,97%, ou seja, excluindo as observações 1094 e 1183 não trazem mudanças significativas nas estimativas dos parâmetros. Assim, o modelo de Schumacher e Hall com efeitos aleatório em  $\gamma_0$  e  $\gamma_2$  com todas as observações foi considerado adequado aos dados.

Tabela 13 – Variação percentual das estimativas dos parâmetros do modelo de Schumacher e Hall com efeitos aleatório, sem as observações 1094 e 1183.

Parâmetros	Todas as observações	Sem as observações 1094 e 1183	Variação Percentual
	Estimativa	Estimativa	
$\beta_0$	0,3641	0,3635	-0,18%
$\beta_1$	2,0114	2,0109	-0,02%
$\beta_2$	0,9944	0,9948	0,04%
$\gamma_0$	0,0113	0,0115	1,97%
$\gamma_2$	0,0153	0,0153	0,26%

Na Figura 7, nota-se que retirando as observações não altera a interpretação dos resíduos em relação à Figura 6, ou seja, eles continuam com média em torno de zero e variância constante.

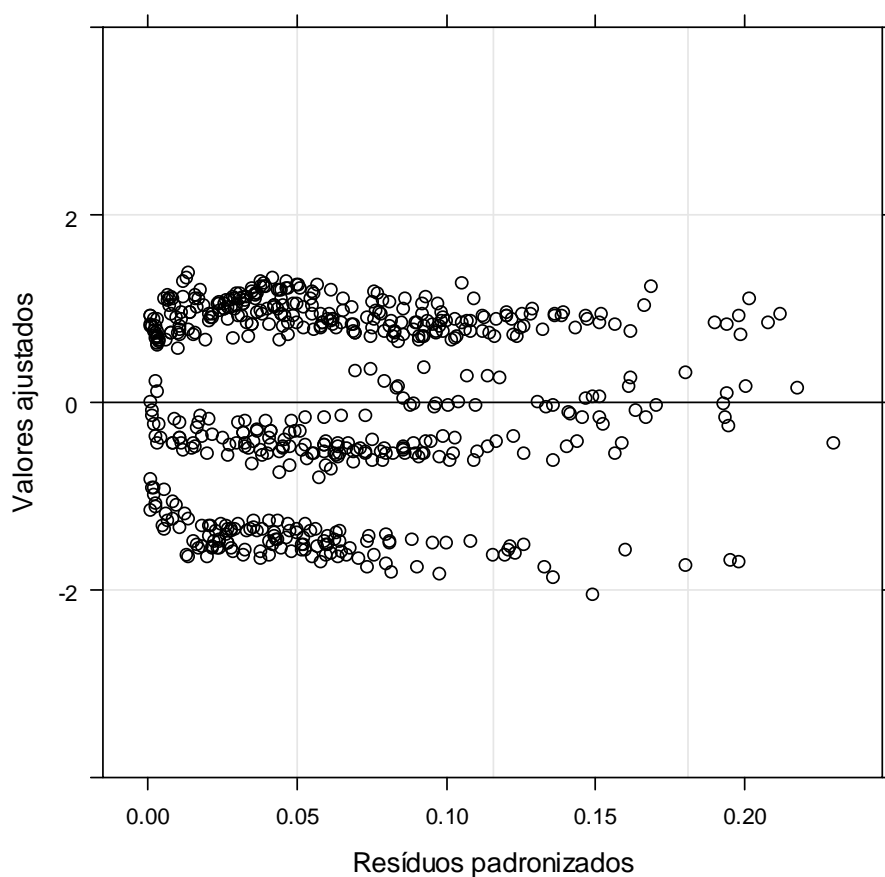


Figura 7 – Resíduos padronizados *versus* valores ajustados para o modelo não linear de efeitos mistos, adotando matriz ARH(1), com função de variância sem as observações 1094 e 1183.

Quanto à normalidade dos resíduos utilizou-se o gráfico dos resíduos padronizados *versus* quantis da distribuição normal padrão apresentado na Figura 8. Como os pontos estão dispostos aproximadamente sobre uma reta, é razoável supor que os dados seguem uma distribuição normal. Para tal confirmação, foi utilizado o teste de Shapiro-Wilk que revelou um valor de  $p$  de 0,6912, ou seja, não se rejeita a hipótese nula ao nível 5% de significância, de que os dados seguem uma distribuição normal.

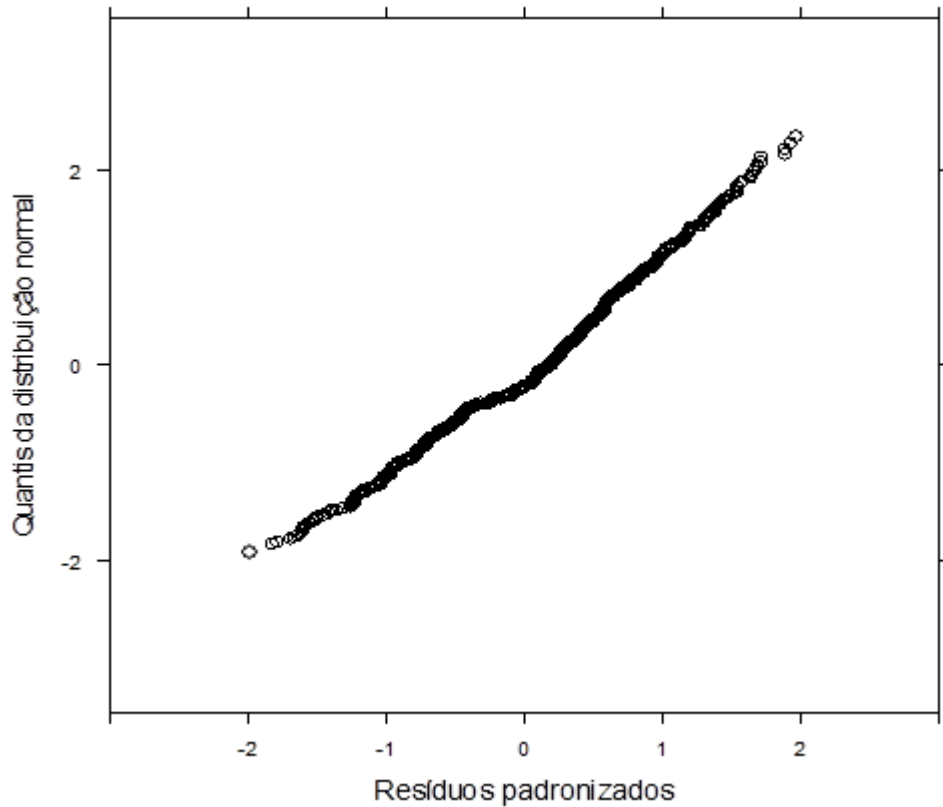


Figura 8 - Gráfico dos resíduos padronizados *versus* quantis da distribuição normal padrão.

Na Figura 9 é possível observar o gráfico dos quantis da distribuição normal padrão *versus* efeitos aleatórios indicado para verificação da normalidade dos efeitos aleatórios. Verifica-se que os efeitos estão razoavelmente dispostos sobre uma reta, o que acaba satisfazendo aproximadamente a suposição de normalidade.

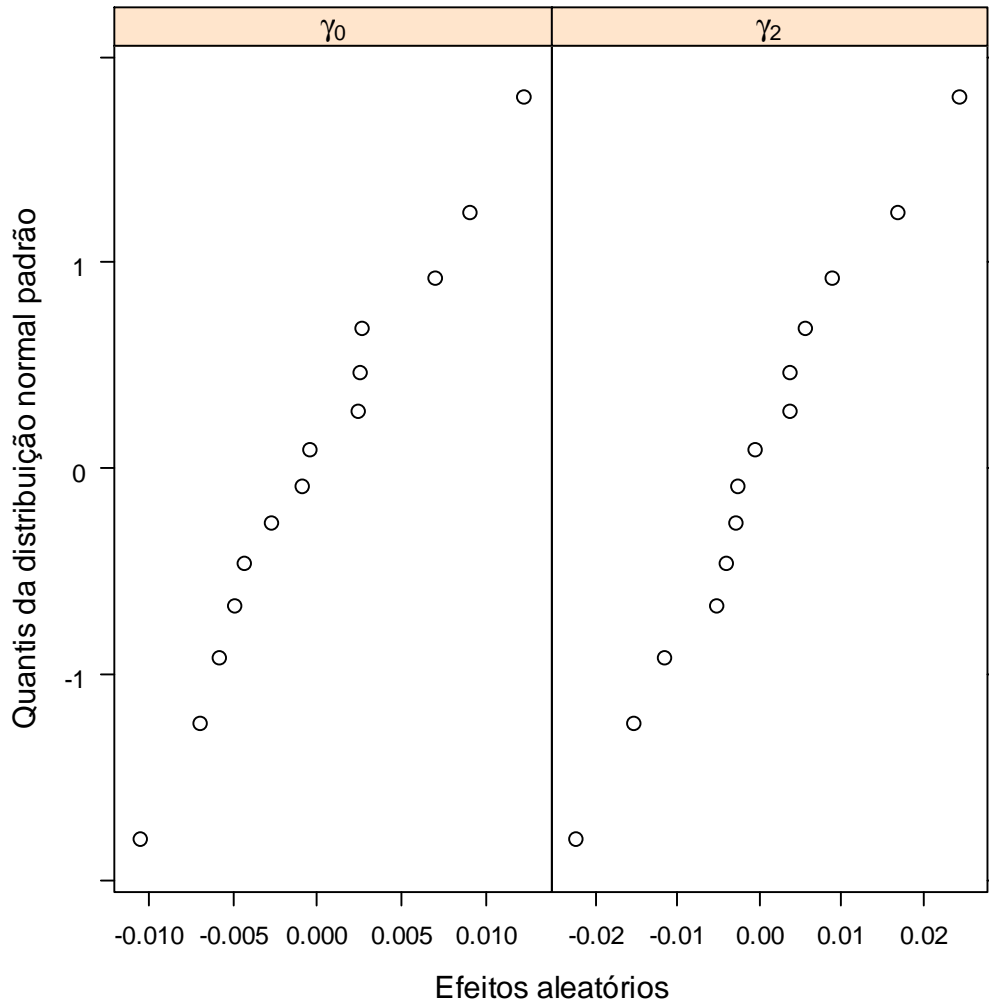


Figura 9 - Gráfico dos quantis da distribuição normal padrão *versus* efeitos aleatórios.

### 4.3 Comparação entre os modelos

Nesta seção foi realizada a comparação entre os modelos de Schumacher Hall (MSH) e Modelos Não Lineares Mistos (MNLM). Os valores estimados dos respectivos modelos foram comparados com os volumes observados do clone de *Eucalyptus* os quais são mostrados nos gráficos em seguida.

Por motivo de brevidade foi apresentado, a seguir, o ajuste de uma árvore, de um total de 89 árvores de *Eucalyptus*, com 1246 observações. Entretanto, os ajustes das árvores restantes podem ser verificados no apêndice.

Na Figura 10 é possível observar que o modelo que se ajustou melhor à estimativa volumétrica dos *Eucalyptus* foi o MNLM, pois em todos os tempos, os valores estimados foram bastante próximos dos observados, algumas vezes praticamente igualando aos valores observados. O MSH apresentou desempenho inferior ao MNLM. As principais discrepâncias, entre valor observado e estimado, surgiram a partir do 60º mês em diante no MSH.

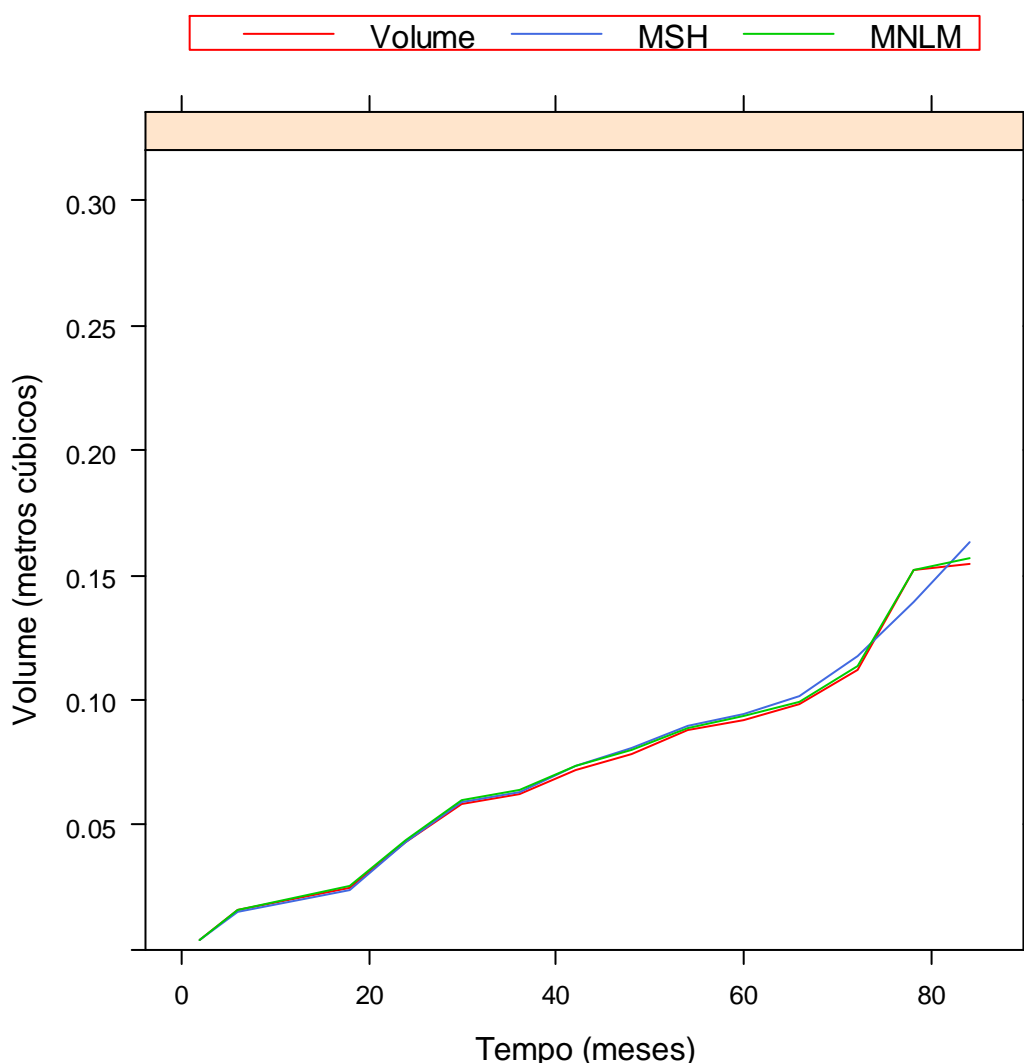


Figura 10: Comparação do crescimento em volume ao longo do tempo da árvore 1.

A Tabela 14 mostra a comparação entre os modelos MSH e MNLM. Como, neste estudo, nem todos os modelos são encaixados, não se podem comparar tais modelos pelo teste da razão verossimilhanças, logo a comparação será realizada pelo Teste de Vuong para modelos distintos. Também serão considerados os critérios seleção de AIC, IAC e EPAM, para escolha do melhor modelo. Rupšys e Petrauskas (2010) também fizeram uso de tais critérios para comparação de modelos na estimação volumétricas de *Pinus sylvestris*. De acordo com os critérios estabelecidos, o que já tinha sido observado pela ilustração gráfica é confirmado pelos critérios de seleção, o modelo não linear misto, com efeitos aleatórios  $\gamma_0$  e  $\gamma_2$ , foi o que melhor se ajustou a estimativa volumétricas dos *Eucalyptus*, pois apresentou menor AIC e EPAM, assim como maior IAC. O Teste de Vuong também confirmou que tal modelo se ajustou melhor aos dados.

Tabela 14 – Critério de seleção dos modelos.

Modelos	AIC	EPAM(%)	IAC	Teste	Vuong	p-valor
MSH	-9740,199	4,609	0,994	----	----	----
MNLM	-15229,410	3,231	0,998	MS vs MNLM	8,982	< 0,001

Na Figura 11 foi gerado o gráfico dos volumes médios observados versus os volumes médios estimados para os modelos MSH e MNLM. Percebe-se, que o MNLM se ajustou muito bem aos dados, praticamente não houve diferenças entre os valores observados e estimados, com os valores médios estimados se sobrepondo sobre os observados. Em relação ao modelo MSH ele não se ajustou bem aos dados, principalmente nos últimos meses.

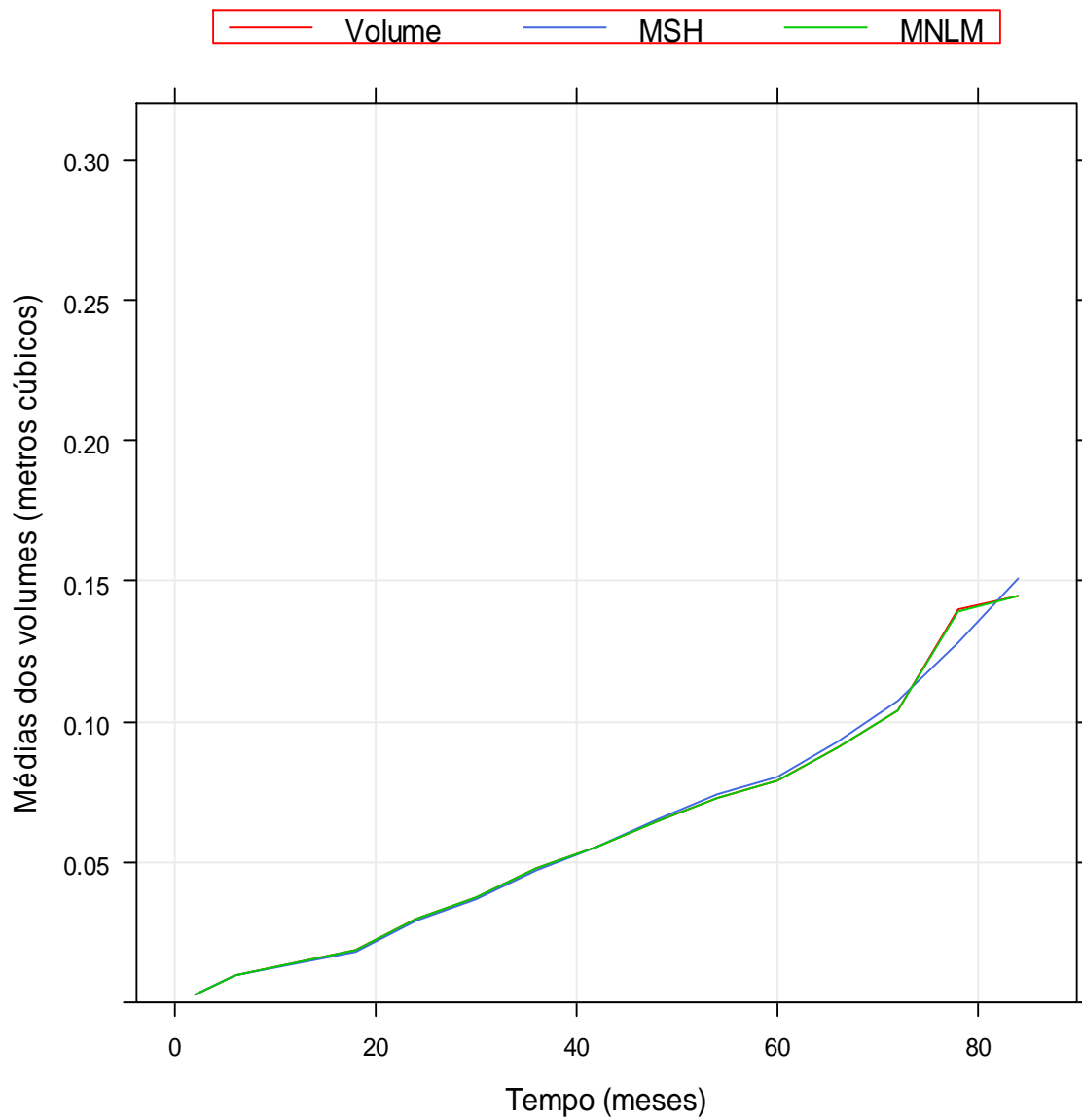


Figura 11: Comparação do crescimento médio em volume ao longo do tempo.

## 5. Conclusões

Os inventários florestais proveem informações fundamentais para a tomada de decisão no manejo sustentado de recursos florestais, fato que o qualifica como essencial para a administração e planejamento de uso atual e futuro das áreas de florestas. Desta forma, o volume é representado como um dos melhores mecanismos para se conhecer o potencial produtivo dessas regiões. Por sua vez, o seu conhecimento se torna imprescindível em todas as atividades de planejamento, principalmente, quando se objetiva prever a viabilidade da implantação dessas árvores para fins energéticos na região do Polo Gesseiro do Araripe.

Sendo assim, a necessidade de informações confiáveis sobre o estoque e o crescimento dos plantios florestais, em contrapartida com os custos envolvidos no inventário florestal, impulsiona ao uso de ferramentas alternativas que permitam melhorar a relação custo/benefício das estimativas. Nesta perspectiva, o modelo não linear misto serviu como ferramenta bastante eficaz para modelagem do volume do clone de *Eucalyptus urophylla*, devido a sua flexibilidade, mostrando-se superior ao ajuste do modelo de Schumacher Hall, quando comparado com o valor observado, resultando, com isso na recomendação de seu uso.

Neste estudo foi traçada a estratégia de ajustar diversas estruturas para matriz de variância e covariância, em que algumas, ora não convergiram ora não apresentaram matriz de covariância definida positiva. A estrutura escolhida, de acordo com os critérios de seleção foi a autorregressiva de primeira ordem heterogênea. Após essa análise foram ajustados 7 modelos combinando os efeitos aleatórios, para depois selecionar quais efeitos seriam incluídos no modelo. O modelo não linear misto com intercepto e o parâmetros associado a altura da árvore aleatórios.

A próxima etapa foi à realização da análise de resíduos em que inicialmente não foi satisfeita a suposição de homocedasticidade. Para contornar o problema, utilizou-se a função poder de covariância, que acomodou bem a dispersão dos resíduos, satisfazendo tal suposição e conseqüentemente diminuindo sensivelmente o erro padrão do modelo, bem com o aumento da verossimilhança e redução do AIC, com a vantagem de não haver transformações sobre os dados originais.



Por fim, os dois modelos analisados foram comparados em que ficou evidente o fato de o modelo não linear misto se ajustar melhor à estimativa volumétrica do clone de *Eucalyptus urophylla*, apresentado a seguinte equação:

$$\widehat{Vol}_i = 0,3754 \cdot DAP_i^{2,0114} \cdot H_i^{1,0097}.$$

Dessa forma, o uso do modelo não linear misto como ferramenta para o inventário florestal, pode contribuir para a diminuição dos custos do inventário, com aumento de precisão.

## Referências Bibliográficas

AHRENS, S. Funções de forma: sua conceituação e utilidade. In: V Seminário sobre atualidade e perspectivas, **Anais**. Curitiba: EMBRAPA-URPFCS, p.7-13., 1982.

ALBUQUERQUE, J. L. **Diagnóstico ambiental e questões estratégicas: uma análise considerando o Polo Gesseiro do Sertão do Araripe** - Estado de Pernambuco. Tese (Doutorado em Engenharia Florestal) - Universidade Federal do Paraná, Curitiba, 185p, 2002.

ALFENAS, A. C. et al. **Clonagem e doença do eucalipto**. Viçosa, Editora UFV, 442p, 2004.

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Boston, v. 19, n. 6, p. 716-723, 1974.

**Anuário Estatístico da ABRAF** (Associação Brasileira dos Produtos de Florestas Plantadas). <http://www.abraflor.org.br/estatisticas>. Acessado em 04/05/2012.

ASSIS, T. F; MAFIA, R. G. Hibridação e clonagem. Viçosa, Editor Aluizio Borém, In: **Biotecnologia Florestal**, p. 93-121, 2007.

ATECEL – Associação Técnica Científica Ernesto Luiz de Oliveira. **Diagnóstico energético do setor industrial do Polo gesseiro da mesorregião de Araripina-PE**. Campina Grande, 126p, 2006.

AVERY, T. E; BURKHART, H. E. **Forest measurements**. 4<sup>o</sup> ed. New York: McGraw-Hill, (McGraw-Hill series in forest resources). 408p, 1997.

BARBOSA, M. **Uma abordagem para análise de dados com medidas repetidas utilizando modelos lineares mistos**. Dissertação de mestrado. USP, 119p, 2009.

BARROS, B. C. et al. Volumetria e sobrevivência de espécies nativas e exóticas no Polo Gesseiro do Araripe, PE. **Ciência Florestal**, Santa Maria, v. 20, n. 4, p. 641-647, out.-dez., 2010.

BERGER, R. **Crescimento e qualidade da madeira de um clone de *Eucalyptus saligna* Smith sob o efeito do espaçamento e da fertilização**. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Santa Maria, Santa Maria, 110 p, 2000.

BERKEY, C. S. Bayesian approach for a nonlinear growth model. **Biometrics** 38,p. 953-961, 1982.

BRANDÃO, A. L. S. **Modelos não lineares para análise de dados longitudinais**. Dissertação (Mestrado em Estatística). Instituto de Matemática, estatística e computação científica, UNICAMP, 1996.

CALEGARIO, N. et al.. **Estimativa do crescimento de povoamentos de *Eucalyptus* baseada na teoria dos modelos não lineares em multinível de efeitos mistos**. Ciências Florestais, Santa Maria, v. 15,n. 3, p 285-292, 2005.

CAMARINHA FILHO, J. A. Utilização de Modelos Lineares Mistos em Agronomia. **Revista de Matemática e Estatística**. v. 21, p. 34-43, 2003.

CAMPOS, J. C. C. et al. Aplicação de um modelo compatível de crescimento e produção de densidade variável em plantações de *Eucalyptus grandis*. **Revista Árvore**, v. 10, n. 2, p. 121-134, 1986.

CAMPOS, J. C. C; TREVISOL J. T.L; NETO P. F. Ainda sobre a seleção de equações de volume. **Revista Árvore**, Viçosa, v.9, n.2, p.115-126, jul./dez. 1985.

CARVALHO, S. P. C. et al. Modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* Clonal. **Cerne**, Lavras, v. 17, n. 4, p. 541-548, out./dez. 2011.

CASTRO, S. M. J. **A metodologia de análise de dados longitudinais em plantas forrageiras no contexto de estruturas de covariância**. Dissertação (Mestrado em Zootecnia) – Universidade Federal do Rio Grande do Sul, 156p, 1999.

CLUTTER, J. L. et al. **Timber management: a quantitative approach**. New York: John, 333p, 1983.

CORDEIRO, M. G; DEMÉTRIO, C. G. B. **Modelos Lineares Generalizados e Extensões**. 475p, 2008.

CORDEIRO, G. M. et al. Uma revisão dos modelos normais não lineares. **Revista Brasileira Biometria**, São Paulo, v.27, n.3, p. 360-393, 2009.

COSTA, B. R. et al.. Avaliação genética de indivíduos de erva-mate (*Ilex paraguariensis* St. Hil) na região de Caarapó, ms, pelo procedimento REML/BLUP. **Revista Ciência Florestal**, Santa Maria, v. 16, n. 2, p. 145-151, 2005.

COSTA, T. R. et al. Modelos lineares mistos: Uma aplicação na curva de lactação de vacas da raça Sindi. **Revista Brasileira de Biometria**, São Paulo, v.30, n.1, p.23-49, 2012.

COUTO, H. T. Z; BASTOS, N. L. M. Erros de medição de altura em florestas de *Eucalyptus* em região plana. **IPEF**, n.39, p.21-31, ago.1988.

CUNHA, T. A; FINGER, C. A. G.; SCHNEIDER, P. R. Linear mixed model to describe the basal area increment for individual cedro (*Cedrela odorata* L.) trees in occidental Amazon, Brazil. **Ciência Florestal**, Santa Maria, v. 23, n. 3, p. 461-470, jul.-set., 2013.

DAVIDIAN, M; GILTINAN, D. M. **Nonlinear Models for Repeated Measurement Data**. New York: Chapman and Hall. 360p, 1995.

DAVIS, L. S; JOHNSON, K. N. **Forest management**. 3º ed. New York: McGraw-Hill inc, 789p. 1987.

DNPM – **Departamento Nacional de Produção Mineral**. 2011, Disponível em: <http://www.dnpm-pe.gov.br>. Acesso em: 12.10.2012.

DP (Diário de Pernambuco). **IBAMA fecha 42 empresas do Polo gesseiro**. Edição de 17.04.2008.

**EMBRAPA**. <http://www.embrapa.br/> Agosto, 2010. Acessado em 13/03/2012.

ENCARNAÇÃO, C. R. F. Observações meteorológicas stipes climáticas das Unidades e Campos Experimentais da Empresa IPA. Recife: **Empresa Pernambucana de Pesquisa Agropecuária**, 110p, 1980.

FEARN, T. A bayesian approach to the growth curves. **Biometrika**, Cambridge, v. 62, p. 89-100, 1975.

FLORES et al. Considerações acerca do tamanho da amostra e número de repetições para avaliação de dados dendrométricos em povoamento clonal de *Eucalyptus saligna*. In: **ABTCPTAPPI 2000**, Congresso Internacional de Celulose e Papel, São Paulo, 15p, 2000.

FOELKEL, C. Gestão ecoeficiente dos resíduos florestais lenhosos da Eucaliptocultura. **Eucalyptus Online Book e Newsletter**. 49p, 2007. Acessado em 14/06/2012.

FRANCO, E. J. et al. Eficiência na estimativa do peso seco para árvores individuais e definição do ponto ótimo de amostragem para determinação da densidade básica *Eucalyptus camaldulensis*. **Ciência Florestal**, Santa Maria, v. 8, n. 1, p. 77-99, nov, 1998.

GALVÃO, A. P. M. **Reflorestamento de propriedade rurais para fins produtivos e ambientais: um guia para ações municipais e regionais**. Brasília: EMBRAPA-CTT, 351p, 2000.

GARCIA, C. H; MORA, A. L. **A cultura do eucalipto no Brasil**. São Paulo, SP: Sociedade Brasileira de Silvicultura. 103p, 2000.

GOMES, A. M. A. **Medição dos arvoredos**. Lisboa: Livraria Sá da Costa, 413p, 1957.

GRIZZLE, J. L; ALLEN, D. M. Analysis of dose and response curves. **Biometrics**, Washington, DC, v. 25, p. 357-382, 1969.

GUIMARÃES, D. P; LEITE, H. G. Influência do número de árvores na determinação de equação volumétrica para *Eucalyptus grandis*. **Revista Scientia Florestalis**, Piracicaba, SP, n. 50, p. 37-42, dez. 1996.

GUJARATI, D. N. **Basic Econometrics**. 3<sup>o</sup> Ed. New York: McGraw-Hill. 838p, 1995.

HARRISON, R. B. et al. Effect of spacing and age on nitrogen and phosphorus distribution in biomass of *Eucalyptus camaldulensis*, *Eucalyptus pellita* and *Eucalyptus urophylla* plantations in southeastern Brazil. **Forest Ecology and Management**, n.133, p. 167-177, 2000.

HARVILLE, T. A. Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, Boston, v. 72, p. 320-338, 1997.

HASSE, G. **Eucalipto: historias de um imigrante vegetal**. Porto Alegre: JÁ Editores, 127p, 2006.

HAUSER, L. et al. Curva de crescimento usando modelo misto: Uma aplicação na progressão da doença de Machado-Joseph. **Revista HCPA**, 29 (1), p. 05-17, 2009.

HUI, S. L. Curve fitting for repeated measurements made at irregular time points. **Biometrics**, Washington, DC, v. 40, p. 691-697, 1984.

JOHANSEN, S. **Functional relations, random coefficients, and nonlinear regression with application to kinetic data**. New York: Springer-Verlag. 546p, 1984.

KULLBACK, S. e LEIBER, R. A. On information and sufficiency. **Annals of Mathematical Statistics**, Vol. 22, p. 78-86, 1951.

LAIRD, N. M; WARE, J. H. Random effects models for longitudinal data. **Biometrics**, Washington, v.38, p. 963-974, 1982.

LEITE, H. G; ANDRADE V. C. L. Um método para condução de inventários florestais sem o uso de equações volumétricas. **Revista Árvore**, viçosa-MG, v.26, n.3, p. 321-328, 2002.

LIMA, C. G. **Análise de dados longitudinais provenientes de experimento em blocos casualizados**. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz” – Universidade de São Paulo, 126p, 1996.

LIMA FILHO, L. M. A. et al. Modelagem do crescimento de clones de *Eucalyptus* usando o modelo de Chapman-Richards com diferentes distribuições simétricas dos erros. **Ciência Florestal**, Santa Maria, v. 22, n. 4, p. 777-785, out.-dez, 2012.

LINDSTROM, M. J e BATES, M. D. Nonlinear Mixed Effects Models for Repeated Measures Data. **Biometrics**, vol. 46, p. 673-687, sep. 1990.

LITTELL, R. C. et al. **SAS System for Mixed Models**. Cary: Statistical Analysis System Institute, 663p. 1996.

LITTELL, R. C. et al. **SAS System for Mixed Models**. 2º ed. North Carolina: SAS Institute Inc., Cary, 615p, 2000.

LOETSCH, F; ZÖHRER, F; HALLER, K. E. **Forest inventory**. München: BLV, v.2. 469p, 1973.

MACHADO, S. A. et al. Comparação de métodos de cubagem absolutos com o volume obtido pelo Xilômetro para Bracatinga (*Mimosa scabrella* Bentham). **Cerne**, Lavras, v. 12, n. 3, p. 2339-253, 2006.

MACHADO, S. A. et al. Modelagem volumétrica para bragatinga (*Mimosa scabrella*) em povoamentos da Região Metropolitana de Curitiba. **Pesquisa Florestal Brasileira**, Colombo, n. 56, p. 17-29, 2008.

MACHADO, S. A; FIGUEIREDO F. A. **Dendrometria**. Guarapuava: UNICENTRO. 2ª Ed, 316p, 2006.

MACHADO, S. A; MELLO, J. A; BARROS, D. A. Comparação entre métodos para avaliação de volume total de madeira por unidade de área, para o pinheiro do Paraná, na região sul do Brasil. **Cerne**, Lavras, v. 6, n. 2, p. 55-66, dez, 2000.

MACHADO, S. A; URBANO, E; CONCEIÇÃO M. B. Comparação de Métodos de Estimativas de volume para *Pinus oocarpa* em diferentes idades e diferentes regimes de desbastes. **Bol. Pesq. Fl.**, Colombo. n. 50, p.81-98, 2005.

MARCELINO, S. D. R; IEMMA, A. F. Métodos de estimação de componentes de variância em modelos mistos desbalanceados. **Scientia Agrícola**, v. 57, n. 4, p. 643-652, 2000.

MARTINS, E. F. P. et al. Curvas de índice de sítio para Leucena (*Leucaena leucocephala* (Lam.) de Wit) no agreste de Pernambuco. **Ciência Florestal**, Santa Maria, v. 17, n. 4, p. 365-376, out-dez, 2007.

McTAGUE, J. P; BATISTA, J. L. F; STEINER, L. H. Equações de volume total, volume comercial e forma do tronco para plantações de *Eucalyptus* nos estados de São Paulo e Rio de Janeiro. **IPEF**, n. 41/42, p. 56-63, 1989.



MIGUEL et al. Ajuste de modelo volumétrico e desenvolvimento de fator de forma para plantios de *Eucalyptus grandis* localizado no município de Rio Verde GO. **Enciclopédia Biosfera**, Centro Científico Conhecer - Goiânia, vol.6, N.11, p. 16 2010.

MORA, F. Herebabilidade y valor genético (REML/BLUP) en genotipos de un eucalipto tolerante a la sequía en el Norte de Chile. **Revista Ciência Florestal**, Santa Maria, v. 16, n. 2, p. 145-151, 2006.

NADARAJAH S; KOTZ S. The beta exponential distribution. **Reliability Engineering and System Safety**, 91, p. 689-697, 2005.

NAVARRO DE ANDRADE. E. **Manual do plantador de Eucalyptos**. São Paulo, Typographia Brazil de Rothschild & Comp. 343p, 1911.

NAVARRO DE ANDRADE, E. **O eucalipto**. 2º edição. Jundiaí: Companhia Paulista de Estradas de Ferro. 170p, 1961.

NETO, P. F. Tabelas volumétricas com e sem casca para *Eucalyptus grandis*. **Revista Árvore**, Viçosa, v.1, n.1, p.31-54, jan./jun. 1977.

NETTO, P. S. Equivalência volumétrica: uma nova metodologia para estimativa do volume de árvores. **Revista Acadêmica: ciências agrárias e ambientais**, Curitiba, v.2, n.1, p. 17-30, 2004.

NEYMAN, J; PEARSON, E. S. On the use and interpretation of certain test criteria for purposes of statistical inference. **Biometrika**, 20A, p. 175-240, 1928.

NOBRE, J. S. **Métodos de diagnostico para modelos lineares mistos**. Dissertação de mestrado, IME/USP, São Paulo, 99p, 2004.

NOBRE, J. S; SINGER, J. M. Residuals analysis for linear mixed models. **Biometrical Journal**, Vol. 49, p. 863-875, 2007.

OLIVEIRA, E. et al. Estrutura anatômica da madeira e qualidade do carvão de *Mimosa tenuiflora* (willd.) Poir. **Revista Árvore**, Viçosa, v. 30, n. 2, p. 311-318, mar/abr, 2006.

OLIVEIRA, M. L. R. et al. Estimação do volume de Árvores de clones de eucalipto Pelo Método da Similaridade de PERFIS. **Revista Árvore** [online]. vol. 33, n.1 p. 133-141, 2009.

PINHEIRO, J. C; BATES, D. M; LINDSTROM, M. J. **Nonlinear mixed effects classes and methods for S**. *Technical Report 906*, Department of Statistics, University of Wisconsin – Madison. 7p, 1993.

PINHEIRO, J. C; BATES, D. M. **Approximations to the loglikelihood function in the nonlinear mixed effects model**. *Technical Report 922*, Department of Statistics, University of Wisconsin – Madison. p. 12-35, 1994.

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

PINHEIRO, J. C; BATES, D. M. **Mixed-effects models in S and S-PLUS**. 2<sup>o</sup> ed. New York: Springer-Verlag, 530p, 2009.

PNUD - **Programa das Nações Unidas para o Desenvolvimento**. Disponível em: <<http://www.pnud.org.br>>, Acesso em: 28 fev, 2012.

RACINE-POON, A. Bayesian approach to nonlinear random effects models. **Biometrics**, p. 1015-1023, 1985.

RAO, C. R. The theory of least squares when the parameters are stochastic and its applications to the analysis of growth curves. **Biometrika**, Cambridge, v. 52, p. 447-458, 1965.

RESENDE, R. R. et al. Emprego de um modelo de crescimento e produção para determinação da rotação em povoamentos de Eucalipto. **Revista Árvore**, Viçosa-MG, v.28, n.2, p.219-225, 2004.

RIPASSI, R. M. A. et al. Cultura do eucalipto na região de Suzanápolis. Estado de São Paulo: Análise econômica. **Informações Econômicas**, v.38, p.7-13, 2008.

ROCHA, K. D. **Produtividade Volumétrica de Clones de Eucalyptus spp. na Região do Polo Gesseiro do Araripe**. Tese (Doutorado em Ciências Florestais) - Universidade Federal Rural de Pernambuco, 110p. 2012.

RODRIGUEZ, L. C. E; BUENO, A. R. S; RODRIGUES, F. Rotações de *eucaliptos* mais longas: Análise volumétrica e econômica. Piracicaba, **Scientia Forestalis**, n.51, p. 15-28, 1997.

ROLIM et al. Modelos volumétricos para a Floresta Nacional do Tapirapé-Aquirí, Serra dos Carajás (PA). **Revista ACTA Amazonica**. Vol 36, p. 107-114, 2006.

RUPSYS, P; PETRAUSKAS, E. Development of q-exponential models for tree height, volume and stem profile. **International Journal of the Physical Sciences**. Vol. 5(15), pp. 2369-2378, 18 November, 2010.

SCARPINELLA, G. D. A., **Reflorestamento no Brasil e o protocolo de Kyoto**. Dissertação (Mestrado em Energia) – Universidade de São Paulo (USP), 182f, 2002.

SCHLAEGEL, B. E. **Testing, reporting, and using biomass estimation models**. In: **Southern forest biomass workshop**. Georgetown. Proceedings Clemson: Clemson University, v. 127, p. 95-112, 1981.

SCHUMACHER, F. X; HALL, F. S. Logarithmic expression of timber-tree volume. **Journal of Agricultural Research**, Lahore, v. 47, n.9, p. 719-734, 1933.

SCOLFORO, J. R. S; MELLO, J. M. **Inventário florestal**. Lavras: UFLA/FAEPE, 310p, 1997.

SCOLFORO, J. R. S. et al. Estimativa de volume, peso seco, peso de óleo e quantidade de moirões para a candeia (*Eremanthus erythropappus* (DC). Macleish). Lavras, **Cerne**, V.10 n.1, p. 87-102, 2004.

SHAPIRO, S. S; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika** 52 (3-4): p. 591–611, 1965.

SILVA, J. A. A. **Análise de equações volumétricas para construção de tabelas de volume comercial para *Eucalyptus* spp., Segundo a Espécie, Região e Método de Regeneração**. Dissertação (Mestrado em Manejo Florestal), UFV, Viçosa MG, 96f, 1997.

SILVA, E. N. et al. Estimação de equações volumétricas para árvores de valor comercial em paragominas, Estado do Pará. **Amazônia: Ci. & Desenv.**, Belém, v. 7, n. 13, jul./dez. p. 7-18, 2011.

SILVA, J. A. A. **Potencialidades de florestas energéticas de *Eucalyptus* no Polo Gesseiro do Araripe-Pernambuco**. Anais da Academia Pernambucana de Ciência Agrônômica, Recife, Vol, 5 e 6, p. 301-319, 2008/2009.

SILVA, J. A. A; BAILEY, R. L. O uso de transformações normalizadoras no ajuste de modelos volumétricos. **Revista Árvore**, Viçosa, MG, v.15, n.2, p. 199-206, 1991.

SINDUSGESSO (Sindicato das Indústrias do Gesso do Estado de Pernambuco), disponível em: [http://www.sindusgesso.org.br/polo\\_gesseiro.asp](http://www.sindusgesso.org.br/polo_gesseiro.asp). Acesso em: 10.02.2012.

SOARES, C. P. B. et al. Estimativas de múltiplos volumes utilizando modelos tipo povoamento total In: **SIMPÓSIO LATINO-AMERICANO SOBRE MANEJO FLORESTAL**. Santa Maria. Anais. Santa Maria: UFSM, p. 279 -286, 2001.

SOBRINHO, A. C. P. L.; AMARAL, A. J. R.; DANTAS, J. O. C. **Gipsita**. Departamento Nacional de Produção Mineral, 2003. Disponível em: <http://www.dnpm-pe.gov.br>. Acesso em: 11.02.2012.

SOUZA, G. S. **Introdução aos modelos de regressão linear e não linear**. Brasília, EMBRAPA, 505p, 1998.

SOUZA, M. O. A. et al. Avaliação da madeira de *Eucalyptus camaldulensis* Dehnh e *Eucalyptus urophylla* ST Blake los Ensaio de usinagem, visando à Produção moveleira. **Revista Árvore**, Viçosa, v 33, n. 4, p. 751-758, 2009.

SPATHELF, P; NUTTO, L. **Modelagem aplicada ao crescimento florestal**. Santa Maria: UFSM, 70p, 2000.

SPURR, S. H. **Forest inventory**. New York: Ronald Press, 476 p, 1952.

THIERSCH, C. R. et al. Acurácia dos métodos para estimativa do volume comercial de clones de *Eucalyptus* sp. **Cerne**, Lavras, v.12, n.2, p.167-181, 2006.

THOMAS, C. et al. Comparação de equações volumétricas ajustadas com dados de cubagem e análise do tronco. **Ciência Florestal**, Santa Maria, v. 16, n. 3, p.319-327, jul./set, 2006.

THORNTHWAITE, C. W; MATHER, J. R. **The water balance**. Centerton, NJ: Drexel Institute of Technology - Laboratory of Climatology, 104 p, 1955.

TONINI, H. Emprego de diferentes metodologias para o cálculo do volume comercial de *Eucalyptus saligna* Smith. In: **SIMPÓSIO LATINO-AMERICANO SOBRE MANEJO FLORESTAL**, Santa Maria. Anais. Santa Maria: UFSM, p. 28-34, 2004.

TRUGILHO, P. F. et al. Avaliação de clones de *Eucalyptus* para produção de carvão vegetal. **Revista Cerne**, v.7, n.2, p. 104-114, 2001.

VONESH, F. E; CHINCHILLI, V. M. **Linear and nonlinear models for the analysis of repeated measurements**. New York: Marcel Dekker, 560p, 1997.

VUONG, Q. Likelihood ratio tests for model selection and nonested hypotheses. **Econometrica**, 57, p. 307-333, 1989.

WALD, V. B. **A metodologia de modelos mistos não lineares aplicados a análise de dados longitudinais em plantas forrageiras**. Dissertação (Mestrado em Zootecnia) – Universidade Federal do Rio Grande do Sul, 106p, 2000.

## Apêndice

```
# _____ Ajustes dos principais modelos volumétricos _____ #

require(nls2)
Volume <- read.table("C:/dados.txt", header=TRUE, sep="", na.strings="NA", dec=".",
strip.white=TRUE)
attach(Volume)

# _____ Modelo de Schumacher e Hall _____ #

MSH<- nls(Vol~b0*(DAP^b1)*(H^b2),data=Volume, start=list(b0=0.01188,b1=1.32860, b2=1.77894))
summary(MSH)
AIC(MSH)

# _____ Modelo de Spurr _____ #

MS<-nls(Vol~b0*((DAP^2)*H)^b1,data=Volume,start=list(b0=0.2831,b1=2.1610))
summary(MS)
AIC(MS)

# _____ Modelo da Variável Combinada de Spurr _____ #
McS<-nls(Vol~b0+b1*((DAP^2)*H),data=Volume,start=list(b0=0.2831,b1=2.1610))
summary(McS)
AIC(McS)

# _____ Modelo Takata _____ #

TKa<-nls(Vol~((DAP^2*H)/(b0+(b1*DAP))),data=Volume,start=list(b0=0.2831,b1=2.1610))
summary(TKa)
AIC(TKa)

# _____ Modelo de Honner _____ #

Hon<-nls(Vol~ (DAP^2)/(b0 + b1*(1/H)),data=Volume,start=list(b0=0.2831,b1=2.1610))
summary(Hon)
AIC(Hon)

AIC(MSH,MS,McS,TKa,Hon)

# _____ Análise de resíduos _____ #

plot(MSH, xlab=" Resíduos padronizados ", ylab="Valores ajustados ")
qqnorm(MSH, xlab=" Resíduos padronizados", ylab=" Quantis da distribuição normal ")

# _____ Gráficos dos resíduos percentuais _____ #

par(mfrow=c(1,2))
X = predict(MSH, data.frame(T = T))
Y = (Vol - X)/Vol
plot(Y, ylab="Erro Percentual", xlab="Índice das Observações",sub="(C)")

# _____ Teste de Shapiro Wilk _____ #

r.pad <- residuals(MSH, type="pearson")
shapiro.test(r.pad)

# _____ Teste de Breusch-Pagan _____ #

bp <- bptest(MSH)
```

```
# _____ Teste de Vuong _____ #
```

```
vuong(MS,MSH)
```

```
# _____ Gráficos _____ #
```

```
Pred1 = predict(MSH, data.frame(T = T))  
Pred2 = predict(MS, data.frame(T = T))  
Pred3 = predict(McS, data.frame(T = T))  
Pred7 = predict(TKa, data.frame(T = T))  
Pred9 = predict(Hon, data.frame(T = T))  
col1 = "red1"  
col2 = "royalblue"  
col3 = "green3"  
col4 = "black"  
col5 = "yellow"  
col6 = "purple"  
cores = c(col1,col2,col3,col4,col5,col6)
```

```
# _____ Vetor de Médias _____ #
```

```
# Médias dos volumes para os valores de T  
# T = 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84, 90  
# Criando os vetores de Médias
```

```
fatorT = as.factor(T)  
nivel = as.numeric( levels(fatorT) )  
  
med.vol = med.pred1 = vector( length = length(nivel) )  
cont = 1  
for(i in nivel){  
  med.vol[cont] = mean( Vol[ T==i ] )  
  med.pred1[cont] = mean( Pred1[ T==i ] )  
  cont = cont + 1  
}  
med.pred2 = vector(length = length(nivel) )  
cont = 1  
for(i in nivel){  
  med.vol[cont] = mean( Vol[ T==i ] )  
  med.pred2[cont] = mean( Pred2[ T==i ] )  
  cont = cont + 1  
}  
med.pred3 = vector(length = length(nivel) )  
cont = 1  
for(i in nivel){  
  med.vol[cont] = mean( Vol[ T==i ] )  
  med.pred3[cont] = mean( Pred3[ T==i ] )  
  cont = cont + 1  
}  
med.pred7 = vector(length = length(nivel) )  
cont = 1  
for(i in nivel){  
  med.vol[cont] = mean( Vol[ T==i ] )  
  med.pred7[cont] = mean( Pred7[ T==i ] )  
  cont = cont + 1  
}  
  
med.pred9 = vector(length = length(nivel) )  
cont = 1
```



```

for(i in nivel){
  med.vol[cont] = mean( Vol[ T==i ] )
  med.pred9[cont] = mean( Pred9[ T==i ] )
  cont = cont + 1
}
medias = as.data.frame( cbind(med.vol, med.pred1, med.pred2, med.pred3, med.pred7,
med.pred9,nivel ) )
colnames(medias) =
c("Med.Vol","Med.Pred1","Med.Pred2","Med.Pred3","Med.Pred7","Med.Pred9","Tempo")
attach(medias)

```

# \_\_\_\_\_ Médias dos volumes \_\_\_\_\_

```

xyplot(Med.Vol + Med.Pred1 + Med.Pred2 + Med.Pred3 + Med.Pred7 + Med.Pred9 ~ Tempo ,
data=medias, type="a",
  xlab="Tempo", ylab="Médias dos volumes",
  strip = strip.custom(strip.names=F, strip.levels=F),
  col=c(col1, col2, col3, col4,col5,col6),
  key =
  list(text = list("", font = 2),
  lines = list(col=cores[1]), text = list("Volume"),
  lines = list(col=cores[2]), text = list("MSH"),
  lines = list(col=cores[3]), text = list("MS"),
  lines = list(col=cores[4]), text = list("McS"),
  lines = list(col=cores[8]), text = list("TKa"),
  lines = list(col=cores[10]), text = list("Hon"),
  border=T),
  panel=function(...) {
    panel.abline(h=(0:3)*0.05,v=(0:4)*20, col="gray92")
    panel.xyplot(...)
  }
)

```

# \_\_\_\_\_ Colocando as Legendas \_\_\_\_\_ #

```

legenda = c("Volume","Schumacher Hall","Spurr","Var. Comb. de Spurr","Takata", "Honner")
cores
xyplot(Med.Vol + Med.Pred1 + Med.Pred2 + Med.Pred3 + Med.Pred7 + Med.Pred9 ~ Tempo ,
data=medias, type="a",
  xlab="Tempo", ylab="Médias dos volumes",sub=" (D)",
  strip = strip.custom(strip.names=F, strip.levels=F),
  col=c(col1, col2, col3, col4,col5,col6),
  key = list( space="top", columns=3,
  lines = list( col=cores ),
  text = list( legenda ), border=T ),
  panel=function(...) {
    panel.abline(h=(0:3)*0.05,v=(0:4)*20, col="gray92")
    panel.xyplot(...)
  }
)

```

# \_\_\_\_\_ Ajuste do modelo não linear misto \_\_\_\_\_ #

```

require(stats)
require(graphics)
library(nlme)
require(nls2)
library(lattice)

```

```

Volume <- read.table("C:/dados.txt", header=TRUE, sep="", na.strings="NA", dec=".",
strip.white=TRUE)
Volume1 <- groupedData(Vol ~ T | ident, data=Volume, order=TRUE,FUN = mean)
Volume = Volume1
attach(Volume)
names(Volume)
plot(Volume)

# _____ Modelo de Schumacher e Hall _____ #

MSH<-nlm(Vol~b0*(DAP^b1)*(H^b2),data=Volume,start=list(b0=0.2241,b1=1.9614,b2=1.1410))

# _____ Ajuste do modelo não linear para cada tempo _____ #

MSH.lis<-nlmList(Vol~b0*(DAP^b1)*(H^b2)|T, data=Volume, start=list(b0=0.2241, b1=1.9614,
b2=1.1410))

# _____ Ajuste do MNLM _____ #

MSH.nlme <- nlme(MSH.lis)

MSH.nlme113A <- nlme(Vol ~ b0 * (DAP^b1)* (H^b2), Volume,
fixed = b0 + b1 + b2 ~ 1,
random = b0 + b1 + b2 ~ 1|T,
start = list(fixed = fixef(MSH.lis)),
method = "ML",
verbose = T)

summary(MSH.nlme113A)
anova(MSH.nlme,MSH.nlme113A)

# _____ Ajuste da Matriz de Covariância _____ #

MSH.nlme1111 <- update(MSH.nlme113A,method = "ML", corr = corAR1() )
MSH.nlme2222 <- update(MSH.nlme113A, corr = corCAR1() )
MSH.nlme3333 <- update(MSH.nlme113A, corr = corCompSymm() )
....
anova(MSH.nlme113A,MSH.nlme1111)
anova(MSH.nlme2222,MSH.nlme1111)
....
# _____ Escolha dos efeitos aleatórios _____ #

MSH.nlme113A1 <- nlme(Vol ~ b0 * (DAP^b1)* (H^b2), Volume,
fixed = b0 + b1 + b2 ~ 1,
random = b0 + b1 + b2 ~ 1|T,
start = list(fixed = fixef(MSH.lis)),
method = "ML",corr = corAR1() ,
verbose = T)
MSH.nlme113A2 <- nlme(Vol ~ b0 * (DAP^b1)* (H^b2), Volume,
fixed = b0 + b1 + b2 ~ 1,
random = b0 ~ 1|T,
start = list(fixed = fixef(MSH.lis)),
method = "ML",corr = corAR1() ,
verbose = T)

....
anova(MSH.nlme113A2,MSH.nlme113A1)
....

```

```
# _____ Inclusão da função de variância _____ #
```

```
MSH.nlme113A77, <- nlme(Vol ~ b0 * (DAP^b1)* (H^b2), Volume,  
  fixed = b0 + b1 + b2 ~ 1,  
  random = b0 + b2 ~ 1|T,  
  start = list(fixed = fixef(MSH.lis)),  
  method = "ML",corr = corAR1() ,weights = varPower(),  
  verbose = T)
```

```
# _____ Retirando as duas observações _____ #
```

```
MSH.nlme113A777 <- nlme(Vol ~ b0 * (DAP^b1)* (H^b2), Volume,  
  fixed = b0 + b1 + b2 ~ 1,  
  random = b0 + b2 ~ 1|T,  
  start = list(fixed = fixef(MSH.lis)),  
  method = "ML",corr = corAR1() ,weights = varPower(),sub=-c(1094,1183),  
  verbose = T)
```

```
# _____ Análise de resíduos _____ #
```

```
plot(MSH.nlme113A77,xlab=" Resíduos padronizados ", ylab="Valores ajustados ")  
qqnorm(MSH.nlme113A77, xlab=" Resíduos padronizados", ylab=" Quantis da distribuição normal ")
```

```
# _____ Gráfico dos efeitos aleatórios _____ #
```

```
titulos = c( expression(gamma[0]), expression(gamma[2]) )  
qqnorm(MSH.nlme113A77, ~ranef(.), col=1, strip=strip.custom(factor.levels=titulos),ylab="Quantis da  
distribuição normal padrão", xlab="Efeitos aleatórios")  
qqnorm(MSH.nlme113A77, ~ ranef(.), titulos = c( expression(gamma[0]), expression(gamma[1]) ))
```

```
# _____ Comparação entre os modelos _____
```

```
MSH.pred = predict(MSH, level = 0:1)  
MSH.nlme4.pred = predict(MSH.nlme113A77, data.frame(T = T))
```

```
# aqui você inclui as variáveis que você quer no dataframe "new"  
new = data.frame( cbind( MSH.pred, MSH.nlme4.pred, Volume))  
names(new)  
col1 = "red1"  
col2 = "royalblue"  
col3 = "green3"  
col4 = "gray"  
col5 = "yellow"  
cores = c(col1,col2,col3)
```

```
# _____ Gráfico individual das árvores _____ #
```

```
fator = which(ident==66) #  
xyplot(Vol[fator] + MSH.pred[fator] + MSH.nlme4.pred[fator] ~ T[fator]  
  | ident[fator], data=new, type="l", xlab="Tempo (meses)",ylim=c(0,0.32),  
  col=cores, pch=c(1,1,1), lty=c(1,1,1), ylab="Volume (metros cúbicos)",  
  key =  
  list(text = list("", font = 2),  
  lines = list(lty = 1, col=cores[1]), text = list("Volume"),  
  lines = list(lty = 1, col=cores[2]), text = list("MSH"),  
  lines = list(lty = 1, col=cores[3]), text = list("MNLM"),  
  border=T)  
)
```

```

#_____ Gráfico da média dos volumes _____#

Pred1 = predict(MSH, data.frame(T = T))
Pred2 = predict(MSH.nlme113A77, data.frame(T = T))
#olivet
col1 = "red1"
col2 = "royalblue"
col3 = "green3"
col4 = "black"
col5 = "yellow"
col6 = "purple"
cores = c(col1,col2,col3)

#_____ Vetor de Médias _____#
#
# Médias dos volumes para os valores de T
# T = 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84, 90
# Criando os vetores de Médias
fatorT = as.factor(T)
nivel = as.numeric( levels(fatorT) )
med.vol = med.pred1 = vector( length = length(nivel) )
cont = 1
for(i in nivel){
  med.vol[cont] = mean( Vol[ T==i ] )
  med.pred1[cont] = mean( Pred1[ T==i ] )
  cont = cont + 1
}
med.pred2 = vector(length = length(nivel) )
cont = 1
for(i in nivel){
  med.vol[cont] = mean( Vol[ T==i ] )
  med.pred2[cont] = mean( Pred2[ T==i ] )
  cont = cont + 1
}
medias = as.data.frame( cbind(med.vol, med.pred1, med.pred2,nivel) )
colnames(medias) = c("Med.Vol","Med.Pred1","Med.Pred2","Tempo")
attach(medias)

#_____ Médias dos volumes _____#

xyplot(Med.Vol + Med.Pred1 + Med.Pred2 ~ Tempo , data=medias, type="a",
  xlab="Tempo", ylab="Médias dos volumes",
  strip = strip.custom(strip.names=F, strip.levels=F),
  col=c(col1, col2, col3),
  key =
  list(text = list("", font = 2),
  lines = list(col=cores[1]), text = list("Volume"),
  lines = list(col=cores[2]), text = list("MSH"),
  lines = list(col=cores[3]), text = list("MNLM"),
  border=T),
  panel=function(...) {
    panel.abline(h=(0:3)*0.05,v=(0:4)*20, col="gray92")
    panel.xyplot(...)
  }
)

#_____ Colocando as Legendas _____#

legenda = c("Volume","MSH","MNLM")
cores

```

```

xyplot(Med.Vol + Med.Pred1 + Med.Pred2 ~ Tempo , data=medias, type="a",
  xlab="Tempo (meses)", ylab="Médias dos volumes (metros cúbicos)",
  strip = strip.custom(strip.names=F, strip.levels=F),
  col=c(col1, col2, col3),
  key = list( space="top", columns=3,
  lines = list( col=cores ),
  text = list( legenda ), border=T ),
  panel=function(...) {
    panel.abline(h=(0:3)*0.05,v=(0:4)*20, col="gray92")
    panel.xyplot(...)
  }
)

```

Comparação do crescimento entre volume ao longo do tempo para as árvores:

