Abordagem bayesiana dos modelos de
regressão hipsométricos não lineares
utilizados em biometria florestal

Monica Fabiana Bento Moreira Thiersch

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura:_____

Abordagem bayesiana dos modelos de regressão hipsométricos não lineares utilizados em biometria florestal

Monica Fabiana Bento Moreira Thiersch

Orientador: *Prof. Dr. Marinho Gomes de Andrade*

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências - Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA.*

**USP – São Carlos
Abril de 2011**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

F689a Fabiana Bento Moreira Thiersch, Monica
Abordagem bayesiana dos modelos de regressão
hipsométricos não lineares utilizados em biometria
florestal / Monica Fabiana Bento Moreira Thiersch;
orientador Marinho Gomes de Andrade -- São Carlos,
2010.
162 p.

Tese (Doutorado - Programa de Pós-Graduação em
Ciências de Computação e Matemática Computacional) --
Instituto de Ciências Matemáticas e de Computação,
Universidade de São Paulo, 2010.

1. Modelos não lineares. 2. Inferência Bayesiana.
3. Eucalyptus sp. 4. densidade a priori. I. Gomes
de Andrade, Marinho, orient. II. Título.

DEDICO

Aos meus pais, em especial à minha mãe Neide, pela
lição de vida, dedicação e amor a seus filhos.

Ao Cláudio, meu marido, a Fabiana minha linda filha
e aos meus irmãos Sidney e Rodrigo (*in memoriam*),
pelo que representam para mim.

OFEREÇO

*Há duas formas para viver sua vida:
Uma é acreditar que não existe milagre.
A outra é acreditar que todas as coisas
são um milagre.*
Albert Einstein

AGRADECIMENTOS

À DEUS por tudo;

A Capes e ao CNPQ, pelo apoio financeiro;

À Universidade de São Paulo e ao Departamento de Matemática e Estatística pela oportunidade de realização de minha pós-graduação;

Ao Professor Marinho Gomes de Andrade pela orientação, apoio, ensinamentos e pelos esforços despendidos durante a realização deste trabalho, que certamente marcarão minha formação como pesquisadora e como ser humano;

Aos Professores Antonio Tavares Costa e Iraziet Charret pela imensa amizade, apoio e ensinamentos durante toda minha graduação e pós-graduação;

Aos professores Sérgio, Maria do Carmo, Telma e Roberto Braga por todo o tempo concedido em ensinar e ajudar durante minha graduação (UFLA);

A todos os professores do Departamento de Matemática Aplicada e Estatística (ICMC);

A todos os colegas de pós-graduação, pela oportunidade de convivência e amizade;

Aos amigos Sidney, Rosa Thiersch, Márcia Thiersch, Adenisio, Maria de Fátima, Agui-naldo, Marília, Carlinhos, Jaqueline, Nilson, Meiriany, Guilherme Bento, Fábio Lúcio, Renata Athayde, Juliano César, Giovanna Tangerino, Debora, Nathalia, Michele Horta, pela amizade e apoio;

Aos meu sobrinhos Laura, Mariana, Marcelo, Carlos Henrique, Duda e Carolina que são uma preciosidade na minha vida.

A toda minha família, em especial a mãe e a minha vó Ormezinda;

A todos que, de alguma maneira, contribuíram para a realização deste trabalho.

**E a vocês Cláudio Thiersch, meu marido e companheiro e
Fabiana que é a luz da minha vida,
MUITO OBRIGADA POR ESTAREM SEMPRE AO MEU LADO!!!!**

RESUMO

THIERSCH, Monica Fabiana Bento Moreira. **Abordagem bayesiana dos modelos de regressão não lineares utilizados em biometria florestal**. 2010. 162p. Monografia (Doutorado) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2010.

Neste trabalho está sendo proposto uma abordagem bayesiana para resolver o problema de inferência com restrição nos parâmetros para os modelos de Petterson, Prodan, Stofel e Curtis, utilizados para representar a relação hipsométrica em clones de *Eucalyptus sp.* Consideramos quatro diferentes densidades de probabilidade *a priori*, entre as quais, a densidade *a priori* não informativa de Jeffreys, a densidade *a priori* vaga normal *flat*, uma densidade *a priori* construída empiricamente e a densidade *a priori* potência. As estimativas bayesianas foram calculadas com a técnica de simulação de Monte Carlo em Cadeia de Markov (MCMC). Os métodos propostos foram aplicados em vários conjuntos de dados reais e os resultados foram comparados aos obtidos com os estimadores de máxima verossimilhança. Os resultados obtidos com as densidades *a priori* não informativa e vaga foram semelhantes aos resultados encontrados com os estimadores de máxima verossimilhança, porém, para vários conjuntos de dados, as estimativas não apresentaram coerência biológica. Por sua vez, as densidades *a priori* informativas empírica e *a priori* potência sempre produziram resultados coerentes biologicamente, independentemente do comportamento dos dados na parcela, destacando a superioridade desta abordagem.

ABSTRACT

THIERSCH, Monica Fabiana Bento Moreira. **Bayesian approach for nonlinear regression models used in forest biometrics.** 2010. 162p. Monograph (Doutorado) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2010.

In this work we propose a Bayesian approach to solve the inference problem with restriction on parameters for the models of Petterson, Prodan, Stofel and Curtis used to represent the hypsometric relationship in clones of *Eucalyptus sp.* We consider four different prior probability densities, the non informative Jeffreys' prior, a vague prior with flat normal probability density, a prior constructed empirically and a power prior density. The Bayesian estimates were calculated using the Monte Carlo Markov Chain (MCMC) simulation technique. The proposed methods were applied to several real data sets and the results were compared to those obtained with the maximum likelihood estimators. The results obtained with a non informative prior and prior vague showed similar results to those found with the maximum likelihood estimators, but, for various data sets, the estimates did not show biological coherence. In turn, the methods a prior empirical informative and a prior power, always produce biologically consistent results, regardless of the behavior of the data in the plot, highlighting the superiority of this approach.

Sumário

Lista de Tabelas

Lista de Figuras

1	Introdução	20
2	Abordagem clássica	24
2.1	Inferência sobre os parâmetros	27
2.2	Análise dos resíduos	28
2.2.1	Tipos de resíduos	28
2.2.2	Análise de diagnóstico	29
2.3	Aplicação do método clássico	30
3	Abordagem bayesiana	38
3.1	Densidade de probabilidade <i>a priori</i>	39
3.1.1	Densidade de probabilidade <i>a priori</i> não informativa de Jeffreys . .	39
3.1.1.1	Densidade de probabilidade <i>a posteriori</i> obtida através da <i>priori</i> não informativa de Jeffreys	40
3.1.1.2	Aplicação	42
3.1.2	Densidade de probabilidade <i>a priori</i> não informativa Normal <i>flat</i> .	51
3.1.2.1	Densidade de probabilidade <i>a posteriori</i> obtida através da <i>priori</i> não informativa Normal <i>flat</i>	51
3.1.2.2	Densidade de probabilidade <i>a posteriori</i> marginais	52
3.1.2.3	Aplicação	53

3.1.3	Densidade de probabilidade <i>a priori</i> empírica	62
3.1.3.1	A densidade de probabilidade <i>a posteriori</i> obtida através da <i>priori</i> empírica	63
3.1.3.2	Modelos de Stofel e Curtis	64
3.1.3.3	Modelos de Prodan e Petterson	64
3.1.3.4	Aplicação	65
3.1.4	Densidade de probabilidade <i>a priori</i> potência	75
3.1.4.1	A densidade de probabilidade <i>a posteriori</i> obtida através da <i>priori</i> potência	76
3.1.4.2	Modelos de Stofel e Curtis	76
3.1.4.3	Modelos de Petterson e Prodan	78
3.1.4.4	Aplicação	80
4	Seleção de modelos	89
4.1	Critério de informação de Akaike	89
4.2	Critério de informação Bayesiano	90
4.3	Critério desvio-informação	91
4.4	Critério da densidade preditiva ordenada	91
4.5	Seleção da densidade de probabilidade <i>a priori</i> para modelos hipsométricos	92
5	Conclusões e Propostas Futuras	94
	Referências	96
	Apêndice A – Resultados do método clássico	100
	Apêndice B – Priori empírica para os modelos hipsométricos	113
B.1	Modelos de Stofel e Curtis	113
B.2	Modelo de Petterson	114
B.3	Modelo de Prodan	116

Apêndice C – Resultados do método bayesiano	118
C.1 Resultados do método bayesiano utilizando densidade de probabilidade a <i>priori</i> empírica	118
C.2 Resultados do método bayesiano utilizando densidade de probabilidade a <i>priori</i> potência	141

Lista de Tabelas

1.1	Relação dos modelos de hipsometria.	22
2.1	Descrição das parcelas utilizadas na aplicação.	31
2.2	Estimativas dos parâmetros pelo método de máxima verossimilhança para os modelos hipsométricos dados na Tabela (1.1)	32
3.1	Derivada dos modelos de hipsometria.	40
3.2	Densidades <i>a priori</i> não informativa de Jeffreys assumindo independência.	40
3.3	Densidades <i>a posteriori</i> condicionais obtidas utilizando <i>a priori</i> não informativa de Jeffreys.	41
3.4	Estimativas bayesianas utilizando densidade <i>a priori</i> de Jeffreys para os modelos hipsométricos dados na Tabela (1.1).	44
3.5	Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade <i>a priori</i> de Jeffreys para os modelos hipsométricos dados na Tabela (1.1).	45
3.6	Estimativas bayesianas utilizando densidade <i>a priori</i> Normal <i>flat</i> para os modelos hipsométricos dados na Tabela (1.1).	55
3.7	Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade <i>a priori</i> normal flat para os modelos hipsométricos dados na Tabela (1.1).	56
3.8	Parâmetros da densidade <i>a priori</i> empírica	63
3.9	Estimativas bayesianas utilizando densidade <i>a priori</i> empírica para os modelos hipsométricos dados na Tabela (1.1).	67
3.10	Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade <i>a priori</i> empírica para os modelos hipsométricos dados na Tabela (1.1).	68

3.11	Estimativas bayesianas utilizando densidade <i>a priori</i> potência para os modelos hipsométricos dados na Tabela (1.1).	81
3.12	Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade <i>a priori</i> potência para os modelos hipsométricos dados na Tabela (1.1).	82
4.1	Critério da densidade preditiva ordenada para os modelos hipsométricos selecionados para cada densidade <i>a priori</i> e parcela utilizada.	93

Lista de Figuras

1.1	Comportamento dos modelos adotados quando $\beta_i > 0$, $i = 0, 1, 2$	22
2.1	Ajuste do modelo de Petterson para a parcela A (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	34
2.2	Ajuste do modelo de Petterson para a parcela B (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	35
2.3	Ajuste do modelo de Petterson para a parcela C (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	36
2.4	Ajuste do modelo de Petterson para a parcela D (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	37
3.1	Ajuste bayesiano utilizando densidade <i>a priori</i> de Jeffreys para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.	46
3.2	Ajuste bayesiano utilizando densidade <i>a priori</i> de Jeffreys para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4.	47
3.3	Ajuste bayesiano utilizando densidade <i>a priori</i> de Jeffreys para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4.	48
3.4	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> de Jeffreys.	49

3.5	Trajétória (A1) e histograma (A2) das amostras geradas para σ^{-2} com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela A, com distribuição <i>a priori</i> de Jeffreys.	50
3.6	Ajuste bayesiano utilizando <i>a priori</i> Normal <i>flat</i> para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.	57
3.7	Ajuste bayesiano utilizando <i>a priori</i> Normal <i>flat</i> para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4.	58
3.8	Ajuste bayesiano utilizando <i>a priori</i> Normal <i>flat</i> para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4.	59
3.9	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> Normal <i>flat</i>	60
3.10	Trajétória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela A, com distribuição <i>a priori</i> Normal <i>flat</i>	61
3.11	Ajuste bayesiano utilizando <i>a priori</i> empírica para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.	70
3.12	Ajuste bayesiano utilizando <i>a priori</i> empírica para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4.	71
3.13	Ajuste bayesiano utilizando <i>a priori</i> empírica para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4.	72
3.14	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> empírica.	73
3.15	Trajétória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela A, com distribuição <i>a priori</i> empírica.	74
3.16	Ajuste bayesiano utilizando <i>a priori</i> potência para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.	84

3.17	Ajuste bayesiano utilizando <i>a priori</i> potência para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4. . . .	85
3.18	Ajuste bayesiano utilizando <i>a priori</i> potência para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4. . . .	86
3.19	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (C1 e C2), parcela A, com distribuição <i>a priori</i> potência.	87
3.20	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela A, com distribuição <i>a priori</i> potência.	88
A.1	Ajuste do modelo de Curtis para a parcela A (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	101
A.2	Ajuste do modelo de Stofel para a parcela A (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	102
A.3	Ajuste do modelo de Prodan para a parcela A (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	103
A.4	Ajuste do modelo de Curtis para a parcela B (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	104
A.5	Ajuste do modelo de Stofel para a parcela B (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	105
A.6	Ajuste do modelo de Prodan para a parcela B (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	106
A.7	Ajuste do modelo de Curtis para a parcela C (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	107

A.8	Ajuste do modelo de Stofel para a parcela C (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	108
A.9	Ajuste do modelo de Prodan para a parcela C (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	109
A.10	Ajuste do modelo de Curtis para a parcela D (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	110
A.11	Ajuste do modelo de Stofel para a parcela D (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	111
A.12	Ajuste do modelo de Prodan para a parcela D (A), resíduo ordinário <i>versus</i> Dap (B), resíduo estudentizado <i>versus</i> Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).	112
C.1	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> empírica.	119
C.2	Trajетória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela A, com distribuição <i>a priori</i> empírica.	120
C.3	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> empírica.	121
C.4	Trajетória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela A, com distribuição <i>a priori</i> empírica.	122
C.5	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela A, com distribuição <i>a priori</i> empírica.	123

C.6	Trajatória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela A, com distribuição <i>a priori</i> empírica.	124
C.7	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição <i>a priori</i> empírica.	125
C.8	Trajatória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela B, com distribuição <i>a priori</i> empírica.	126
C.9	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição <i>a priori</i> empírica.	127
C.10	Trajatória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela B, com distribuição <i>a priori</i> empírica.	128
C.11	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição <i>a priori</i> empírica.	129
C.12	Trajatória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela B, com distribuição <i>a priori</i> empírica.	130
C.13	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela B, com distribuição <i>a priori</i> empírica.	131
C.14	Trajatória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela B, com distribuição <i>a priori</i> empírica.	132

C.15	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição <i>a priori</i> empírica.	133
C.16	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela C, com distribuição <i>a priori</i> empírica.	134
C.17	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição <i>a priori</i> empírica.	135
C.18	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela C, com distribuição <i>a priori</i> empírica.	136
C.19	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição <i>a priori</i> empírica.	137
C.20	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela C, com distribuição <i>a priori</i> empírica.	138
C.21	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela C, com distribuição <i>a priori</i> empírica.	139
C.22	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela C, com distribuição <i>a priori</i> empírica.	140
C.23	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> potência.	141

C.24 Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela A, com distribuição <i>a priori</i> potência.	142
C.25 Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> potência.	143
C.26 Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela A, com distribuição <i>a priori</i> potência.	144
C.27 Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição <i>a priori</i> potência.	145
C.28 Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela A, com distribuição <i>a priori</i> potência.	146
C.29 Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição <i>a priori</i> potência.	147
C.30 Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela B, com distribuição <i>a priori</i> potência.	148
C.31 Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição <i>a priori</i> potência.	149
C.32 Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela B, com distribuição <i>a priori</i> potência.	150

C.33	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição <i>a priori</i> potência.	151
C.34	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela B, com distribuição <i>a priori</i> potência.	152
C.35	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela B, com distribuição <i>a priori</i> potência.	153
C.36	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela B, com distribuição <i>a priori</i> potência.	154
C.37	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição <i>a priori</i> potência.	155
C.38	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela C, com distribuição <i>a priori</i> potência.	156
C.39	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição <i>a priori</i> potência.	157
C.40	Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela C, com distribuição <i>a priori</i> potência.	158
C.41	Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição <i>a priori</i> potência.	159

C.42	Trajeto�ria (A1) e histograma (A2) das amostras geradas com o algoritmo de simula��o MCMC ap�s ser verificada a converg�ncia para o modelo de Petterson, parcela C, com distribui��o <i>a priori</i> pot�ncia.	160
C.43	Histograma e trajet�ria das amostras geradas com o algoritmo de simula��o MCMC ap�s ser verificada a converg�ncia para o modelo de Prodan, para os par�metros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela C, com distribui��o <i>a priori</i> pot�ncia.	161
C.44	Trajeto�ria (A1) e histograma (A2) das amostras geradas com o algoritmo de simula��o MCMC ap�s ser verificada a converg�ncia para o modelo de Prodan, parcela C, com distribui��o <i>a priori</i> pot�ncia.	162

1 *Introdução*

A predição e projeção da produção florestal desde os primórdios das atividades florestais planejadas sempre foram desafios aos biometristas florestais, os quais continuam desenvolvendo métodos que visam aumentar o grau de exatidão das estimativas. O estudo da eficiência das estimativas e determinações obtidas para árvores e povoamentos tem se tornado cada vez mais imperioso. Esses estudos têm sido fortalecidos principalmente pelo fato da atividade florestal ter se tornado um negócio cada vez mais rentável e competitivo economicamente.

A quantificação da altura das árvores é fundamental principalmente para o cálculo do volume de madeira dos povoamentos e para se fazer classificações destes quanto às suas produtividades. A altura é uma medida indireta e de difícil mensuração, a qual é obtida geralmente através de hipsômetros que sob condições inadequadas de visibilidade da copa das árvores, topografia, e sub-bosque podem acarretar procedimentos demorados e de alto custo, onerando seu uso e apresentando erros graves de medição, proporcionando considerável aumento do viés no inventário (SCOLFORO, 1997).

Para contornar estes problemas, tem-se como prática medir na parcela a altura de parte das árvores e o diâmetro de todas e, através dos pares altura-diâmetro mensurados, estabelecer uma relação matemática que possibilite estimar as alturas das demais árvores contidas nas parcelas. Essas relações matemáticas são conhecidas como equações hipsométricas, que buscam estimar a altura das árvores através da relação *DAP* (Diâmetro do tronco da árvore a 1,3m do solo) e *H* (altura total da árvore). Essa técnica de mensuração na qual mede-se o diâmetro de todas as árvores da parcela e a altura total de parte delas, já descrita por Ker e Smith (1957), resulta em uma sensível redução nos custos do inventário florestal e em uma operacionalização mais eficaz (RIBEIRO et al., 2010).

Para que as estimativas destas equações hipsométricas sejam precisas, devem ser observados os diversos fatores que influenciam essa relação, tais como: idade, região de plantio, sistema de amostragem, densidade de plantio, tamanho da copa, posição sociológica, material genético e práticas silviculturais. Neste sentido, espera-se que os dados

que serão utilizados para obtenção das estimativas das alturas das árvores que não foram mensuradas captem os diversos fatores que afetam essa relação. Assim, caso o tamanho da amostra dos pares diâmetro-altura medidos seja insuficiente, ou seja, caso não represente toda a amplitude de variação dos dados da região de estudo, poderão acarretar estimativas distorcidas e/ou pouco realistas.

O número de árvores para a determinação de uma relação hipsométrica é influenciada diretamente pela heterogeneidade que o plantio apresenta, e que por mais homogêneo que seja, uma quantidade mínima de pares de altura-diâmetro devem ser amostradas. De acordo com Miguel (2009) este valor deve estar acima de 30 pares de altura-diâmetro por hectare. Silva et al. (2007) com estudos de simulação afirma que este número deve ser superior a 27 pares de altura-diâmetro. Estas recomendações nem sempre são seguidas pelas empresas florestais, que no intuito de reduzirem os custos com o inventário florestal, trabalham com amostras que giram em torno de 20 a 30 pares de altura-diâmetro mensurados por hectare. Isso pode acarretar em alguns casos, resultados super ou subestimados e até estimativas inconsistentes biologicamente o que inviabilizaria seu uso. Para tentar resolver esse problema, uma alternativa seria aumentar a amostragem. Como esse procedimento é economicamente inviável, outra alternativa seria utilizar informações *a priori* sobre a relação hipsométrica na região em estudo (MOREIRA et al., 2007).

No mesmo sentido, buscando melhorar as estimativas das alturas, os métodos e modelos adotados tem tornado-se cada vez mais sofisticados (SOARES et al., 2004). A grande dificuldade da escolha do melhor modelo para representar estas relações deve-se a não linearidade da relação entre as variáveis envolvidas e a grande variedade de modelos que se pode utilizar (MACHADO; BASSO; BEVILACQUA, 1993). Estes dois fatores aliados as restrições impostas aos parâmetros dos modelos, por razões biológicas, torna o ajuste de modelos matemáticos para representar as relações hipsométricas um problema de regressão não linear com restrições nos parâmetros. A abordagem desse problema com os métodos convencionais de inferência estatística, tais como método de máxima verossimilhança e método de mínimos quadrados torna-se ineficiente para vários modelos propostos. Uma abordagem bayesiana mostra-se como uma abordagem alternativa promissora por permitir incorporar, *a priori*, restrições biológicas aos parâmetros dos modelos considerados (MOREIRA et al., 2007).

De forma geral é possível representar a relação entre o diâmetro (Dap) do tronco da árvore e a altura total da árvore (H) usando a relação

$$y_i = \boldsymbol{\eta}(x_i; \boldsymbol{\beta}) + \epsilon_i \quad (1.1)$$

onde $i = 1, \dots, n$, y_i e x_i são funções monotônicas da altura H_i e do diâmetro Dap_i , respectivamente, ϵ_i é um erro aleatório.

Existe uma profusão de modelos que descrevem essa relação (COUTO; BASTOS, 1986; CURTIS, 1967; SOARES et al., 2004; MIGUEL, 2009; RIBEIRO et al., 2010; SCOLFORO, 1997; RUFINO et al., 2010). Não foi foco deste trabalho questionar os modelos hipsométricos mais utilizados no setor florestal, mas tentar estimar de forma mais precisa e eficiente seus parâmetros com amostras de tamanho reduzido. Dessa forma, serão considerados quatro desses modelos que podem ser vistos na Tabela (1.1) e na Figura (1.1).

Tabela 1.1: Relação dos modelos de hipsometria.

Autor	y	x	$\eta(x, \beta)$
Stofel	$\ln(H)$	$\ln(Dap)$	$\beta_0 + \beta_1 x$
Curtis	$\ln(H)$	$1/Dap$	$\beta_0 - \beta_1 x$
Petterson	$\ln(H)$	$1/Dap$	$-3 \ln(\beta_0 + \beta_1 x)$
Prodan	$\ln(H)$	$1/Dap$	$-\ln(\beta_0 x^2 + \beta_1 x + \beta_2)$

Para todos os modelos apresentados na Tabela (1.1) devemos considerar a restrição $\beta > 0$ para os parâmetros, necessárias para que os modelos descrevam curvas biologicamente coerentes, como pode ser visto na Figura (1.1).

Figura 1.1: Comportamento dos modelos adotados quando $\beta_i > 0$, $i = 0, 1, 2$.

Neste trabalho, em lugar das observações pareadas $\{(H_1, Dap_1), \dots, (H_n, Dap_n)\}$ usaremos a relação apropriada para cada modelo dado na Tabela (1.1), e vamos considerar

as observações transformadas $\{(y_1, x_1), \dots, (y_n, x_n)\}$, onde $y_i = g_1(H_i)$ e $x_i = g_2(Dap_i)$, com $g_1(\cdot)$ e $g_2(\cdot)$ dados na Tabela (1.1).

Vamos denotar por $\boldsymbol{\beta} = (\beta_0, \dots, \beta_s)'$ o vetor de parâmetros de cada modelo e escrever de forma geral o modelo para as n observações transformadas, na forma vetorial, como se segue:

$$\mathbf{Y} = \boldsymbol{\eta}(\mathbf{x}; \boldsymbol{\beta}) + \boldsymbol{\epsilon} \quad (1.2)$$

em que, $\mathbf{Y} = (y_1, \dots, y_n)'$, $\mathbf{x} = (x_1, \dots, x_n)'$, $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)'$ e $\boldsymbol{\eta}(\mathbf{x}; \boldsymbol{\beta}) = (\eta(x_1; \boldsymbol{\beta}), \dots, \eta(x_n; \boldsymbol{\beta}))'$.

Essa forma geral de representação dos modelos apresentados na Tabela (1.1) será utilizada neste trabalho nas diferentes abordagens propostas para obter a estimação e inferência dos parâmetros dos modelos.

Esta tese está estruturada da seguinte forma: No capítulo 2 é apresentada a abordagem clássica que utiliza o método de máxima verossimilhança para fazer a estimação e inferência sobre os parâmetros dos modelos. No capítulo 3 apresentamos uma abordagem bayesiana com densidade *a priori* não informativa de Jeffreys, uma densidade *a priori* vaga normal *flat*, e densidades *a priori* informativas empírica e potência, destacando a superioridade dos métodos bayesianos com densidades *a priori* informativa sobre os demais, no capítulo 4 apresentamos métodos de seleção de modelos com diferentes densidades de probabilidade *a priori* para os modelos hipsométricos aplicados em clones de *Eucalyptus sp*, e finalizando, no capítulo 5 as conclusões e proposta para trabalhos futuros.

A grande contribuição deste trabalho foi a adoção de uma metodologia bayesiana com densidades *a priori* informativas para estimar os parâmetros dos modelos hipsométricos, agregando mais informação ao estudo presente. A proposta de uma densidade *a priori* potência foi feita por Ibrahim e Chen (2000). Neste trabalho e nos demais que sucederam a essa idéia foram utilizados apenas modelos lineares. Nesta tese, a inovação foi trabalhar com modelos não lineares obtendo as estimativas dos parâmetros através de um algoritmo MCMC.

2 Abordagem clássica

Neste capítulo vamos considerar o método da máxima verossimilhança para estimar os parâmetros dos modelos hipsométricos descritos na Tabela (1.1).

Para escrever a função de verossimilhança dos modelos hipsométricos dados na Tabela (1.1), será assumido que o erro ϵ_i é um erro aleatório independente e identicamente distribuído, com $\epsilon_i \sim N(0, \sigma^2)$ e $i = 1, \dots, n$. Desta forma a função de verossimilhança para os dados $D = \{(H_i, Dap_i), i = 1, 2, \dots, n\}$ pode ser escrita como:

$$L(\beta, \tau | \mathbf{D}) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \sum_{i=1}^n (y_i - \eta(x_i; \beta))^2 \right\} \quad (2.1)$$

em que $\tau = 1/\sigma^2 > 0$.

Para simplificar os cálculos é comum estimar os parâmetros a partir da função log-verossimilhança, uma vez que maximizar $L(\beta, \tau | \mathbf{D})$ equivale maximizar $\ln(L(\beta, \tau | \mathbf{D})) = l(\beta, \tau | \mathbf{D})$. Para os modelos da Tabela (1.1) a função log-verossimilhança é dada por:

$$l(\beta, \tau | \mathbf{D}) = \frac{n}{2} \ln \left(\frac{\tau}{2\pi} \right) - \frac{\tau}{2} \sum_{i=1}^n (y_i - \eta(x_i; \beta))^2. \quad (2.2)$$

Devido a relação não linear envolvendo o vetor de parâmetros $\beta = (\beta_0, \dots, \beta_p)$, as estimativas destes parâmetros só podem ser calculadas numericamente. Para isso será considerada a função score U_r , $r = 1, \dots, p$, dada por:

$$U_r = \frac{\partial l(\beta, \tau; \mathbf{D})}{\partial \beta_r} = \sum_{i=1}^n \frac{\partial l(\beta, \tau; \mathbf{D})}{\partial \eta} \frac{\partial \eta(x_i; \beta)}{\partial \beta_r}$$

logo,

$$U_r = \tau \sum_{i=1}^n (y_i - \eta(x_i; \beta)) \frac{\partial \eta(x_i; \beta)}{\partial \beta_r}. \quad (2.3)$$

A matriz de informação de Fisher é dada por $K = \{K_{r,s}; r, s = 1, \dots, p\}$ onde

$$K_{r,s} = E \left\{ \frac{\partial l(\boldsymbol{\beta}, \tau; \mathbf{D})}{\partial \beta_r} \frac{\partial l(\boldsymbol{\beta}, \tau; \mathbf{D})}{\partial \beta_s} \right\}$$

$$K_{r,s} = \tau^2 \sum_{i=1}^n E(y_i - \eta(x_i; \boldsymbol{\beta}))^2 \frac{\partial \eta(x_i; \boldsymbol{\beta})}{\partial \beta_r} \frac{\partial \eta(x_i; \boldsymbol{\beta})}{\partial \beta_s}$$

como $E(y_i - \eta(x_i; \boldsymbol{\beta}))^2 = 1/\tau$, $K_{r,s}$ será dado por

$$K_{r,s} = \tau \sum_{i=1}^n \frac{\partial \eta(x_i; \boldsymbol{\beta})}{\partial \beta_r} \frac{\partial \eta(x_i; \boldsymbol{\beta})}{\partial \beta_s} \quad (2.4)$$

Na forma matricial, a matriz K será dada por

$$K = \tau X' X \quad (2.5)$$

onde a matriz X é definida como

$$X = \begin{pmatrix} \frac{\partial \eta(x_1, \boldsymbol{\beta})}{\partial \beta_1} & \frac{\partial \eta(x_1, \boldsymbol{\beta})}{\partial \beta_2} & \cdots & \frac{\partial \eta(x_1, \boldsymbol{\beta})}{\partial \beta_p} \\ \frac{\partial \eta(x_2, \boldsymbol{\beta})}{\partial \beta_1} & \frac{\partial \eta(x_2, \boldsymbol{\beta})}{\partial \beta_2} & \cdots & \frac{\partial \eta(x_2, \boldsymbol{\beta})}{\partial \beta_p} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \eta(x_n, \boldsymbol{\beta})}{\partial \beta_1} & \frac{\partial \eta(x_n, \boldsymbol{\beta})}{\partial \beta_2} & \cdots & \frac{\partial \eta(x_n, \boldsymbol{\beta})}{\partial \beta_p} \end{pmatrix}_{n \times p}. \quad (2.6)$$

O método escore de Fisher é usado para calcular as estimativas do vetor de parâmetros $\boldsymbol{\beta}$ e é dado por:

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + (K^{-1})^{(m)} U^{(m)} \quad (2.7)$$

em que $U^{(m)}$ e $(K^{-1})^{(m)}$ são respectivamente a matriz K e o vetor escore U avaliados substituindo-se $\boldsymbol{\beta}$ por suas estimativas $\boldsymbol{\beta}^{(m)}$ calculadas na m -ésima iteração.

Multiplicando-se ambos os membros de da equação (2.7) por $K^{(m)}$, tem-se

$$K^{(m)} \boldsymbol{\beta}^{(m+1)} = K^{(m)} \boldsymbol{\beta}^{(m)} + U^{(m)} \quad (2.8)$$

considerando as equações (2.3) e (2.6) podemos escrever

$$U^{(m)} = \tau (X^{(m)})' \boldsymbol{\epsilon}^{(m)} \quad (2.9)$$

em que

$$\boldsymbol{\epsilon}^{(m)} = \begin{pmatrix} y_1 - \eta(x_1, \boldsymbol{\beta}^{(m)}) \\ y_2 - \eta(x_2, \boldsymbol{\beta}^{(m)}) \\ \vdots \\ y_n - \eta(x_n, \boldsymbol{\beta}^{(m)}) \end{pmatrix}_{n \times 1}. \quad (2.10)$$

Substituindo (2.5) e (2.9) na equação (2.8), tem-se que:

$$(X^{(m)})' X^{(m)} \boldsymbol{\beta}^{(m+1)} = (X^{(m)})' X^{(m)} \boldsymbol{\beta}^{(m)} + (X^{(m)})' \boldsymbol{\epsilon}^{(m)}.$$

Definindo o vetor $\mathbf{z}^{(m)}$ por:

$$\mathbf{z}^{(m)} = X^{(m)} \boldsymbol{\beta}^{(m)} + \boldsymbol{\epsilon}^{(m)} \quad (2.11)$$

tem se:

$$(X^{(m)})' X^{(m)} \boldsymbol{\beta}^{(m+1)} = (X^{(m)})' \mathbf{z}^{(m)}$$

$$\boldsymbol{\beta}^{(m+1)} = \left[(X^{(m)})' X^{(m)} \right]^{-1} (X^{(m)})' \mathbf{z}^{(m)}. \quad (2.12)$$

Portanto as equações (2.11) e (2.12) podem ser usadas para calcular iterativamente as estimativas do vetor de parâmetros $\boldsymbol{\beta}$.

Para a estimativa do parâmetro τ vamos considerar a função escore:

$$U_\tau = \frac{\partial l(\boldsymbol{\beta}, \tau; \mathbf{D})}{\partial \tau}$$

logo,

$$U_\tau = -\frac{n}{2} \tau^{-1} - \frac{1}{2} \sum_{i=1}^n (y_i - \eta(x_i; \boldsymbol{\beta}))^2. \quad (2.13)$$

Fazendo $U_\tau = 0$ tem-se que:

$$\tau^{-1} = \frac{1}{n} \sum_{i=1}^n (y_i - \eta(x_i; \boldsymbol{\beta}))^2.$$

Denotando-se por $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \dots, \hat{\beta}_p)'$ o vetor de estimativas de máxima verossimilhança do vetor de parâmetros $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ e denotando por $\phi = 1/\tau$, tem-se o estimador de máxima verossimilhança $\hat{\phi}$ para ϕ , dado por:

$$\hat{\phi} = \frac{1}{n} \sum_{i=1}^n (y_i - \eta(x_i; \hat{\boldsymbol{\beta}}))^2. \quad (2.14)$$

Utilizando a notação vetorial dada em (2.10) e denotando-se por $\hat{\epsilon}$ o vetor ϵ avaliado com

as estimativas $\hat{\beta}$, tem-se que:

$$\hat{\phi} = \frac{1}{n} \hat{\epsilon}' \hat{\epsilon} \quad (2.15)$$

Os parâmetros iniciais utilizados para encontrar as estimativas dos parâmetros de cada modelo dado na Tabela 1.1 foram obtidos de maneira empírica de forma muito semelhante a realizada para encontrar os parâmetros da densidade *a priori* empírica na abordagem bayesiana com densidade *a priori* empírica (Capítulo 3). Para maiores detalhes veja apêndice B.

2.1 Inferência sobre os parâmetros

Inferências sobre os parâmetros são feitas usualmente através da normalidade assintótica dos estimadores de máxima verossimilhança (BICKEL; DOKSUM, 2006). Dessa forma, para amostras suficientemente grandes (De acordo com Silva et al. (2007), para dados hipsométricos, a amostra deveria ter um mínimo de 27 pares de altura-diâmetro) e sob certas condições de regularidade (SEBER; WILD, 1989), a inferência sobre os parâmetros $\beta = (\beta_1, \dots, \beta_p)$ pode ser feita usando a propriedade assintótica onde a matriz de covariância de β é aproximada por $\phi(X'X)^{-1}$. Assim, tem-se que $Var(\hat{\beta}) = diag\{\phi(X'X)^{-1}\}$. Para fazer inferência sobre $\hat{\phi}$ usa-se a aproximação assintótica $Var(\hat{\phi}) = 2\phi^2/n$. As variâncias para $\hat{\beta}$ e $\hat{\phi}$ podem ser estimadas de forma consistente substituindo-se β e ϕ por suas estimativas de máxima verossimilhança $\hat{\beta}$ e $\hat{\phi}$ para calcular (sob a teoria de grandes amostras) intervalos de confiança para os parâmetros.

Outra maneira de calcular intervalos de confiança para os parâmetros é utilizar *Bootstrap* (EFRON; TIBSHIRANI, 1993). Esse método possui vantagens por não utilizar resultados assintóticos para a distribuição dos estimadores dos parâmetros β 's. Esse tipo de intervalo é apropriado quando se tem uma amostra “pequena”, ou quando for crítica a suposição de normalidade. O método *Bootstrap* consiste em, a partir das estimativas de máxima verossimilhança $\hat{\beta}$, gerar novas variáveis respostas y_i , como se fosse realizado um novo experimento, usando o fato que $y_i \sim N(\eta(x_i, \hat{\beta}), \hat{\sigma}^2)$. Para cada “novo” experimento gerado $B_j = (y_{j1}, y_{j2}, \dots, y_{jd})$, com j variando até M grande, por exemplo, $M = 10000$ recalcula-se as estimativas $\hat{\beta}^*$, que são denominadas de estimativas *Bootstrap*. Dessa forma o intervalo de confiança percentílico *Bootstrap* será dado por

$$ICBoot = (\hat{\beta} - D_2; \hat{\beta} + D_1), \quad (2.16)$$

onde $D_1 = \hat{\beta} - \hat{\beta}^*_{\alpha/2}$ e $D_2 = \hat{\beta}^*_{1-\alpha/2} - \hat{\beta}$, sendo $\hat{\beta}^*_{\alpha/2}$ e $\hat{\beta}^*_{1-\alpha/2}$ os quantis amostrais de

ordem $\alpha/2$ e $1 - \alpha/2$ das estimativas *Bootstrap* $\hat{\beta}^*$.

Crterios de seleo de modelos tais como AIC e BIC podem ser usados para selecionar o melhor modelo para descrever a relao hipsomtrica. Neste captulo ser adotado o critrio BIC, proposto por Schwarz (1978), dado por

$$BIC = -2l(\hat{\beta}, \hat{\phi}; \mathbf{D}) + (p + 1) \quad (2.17)$$

onde p a dimenso do vetor de parmetros β .

2.2 Anlise dos resduos

Como j foi dito, o ajuste de modelos hipsomtricos so muito teis para estimar a altura das rvores que no foram medidas em uma determinada parcela. Entretanto, de suma importncia a verificao do ajuste desses modelos para detectar possveis transgresses as suposies feitas e conseqentemente o uso de modelos que acarretariam pouca utilidade e baixa confiabilidade nos seus resultados.

Uma das maneiras de investigar o problema estudar o comportamento do modelo no conjunto de dados observados, principalmente as discrepncias entre os valores observados e os valores ajustados, ou seja, pela anlise de resduos.

Os resduos dos modelos dados na Tabela (1.1) so estimados como a equao (2.10) substituindo β pela respectiva estimativa de mxima verossimilhana $\hat{\beta}$. A anlise desses resduos mostra a qualidade do ajuste e pode ser utilizada para detectar a presena de pontos atpicos que podem influenciar o ajuste do modelo.

2.2.1 Tipos de resduos

Na literatura so encontrados vrios tipos de resduos (veja por exemplo Cook e Weisberg (1982), Cook e Tsai (1985)), os quais tem fundamental importncia na verificao do ajuste do modelo.

a) Resduos ordinrios

Os resduos advindos do processo de ajuste so dados por:

$$\hat{\epsilon}_i = y_i - \eta(x_i; \hat{\beta}). \quad (2.18)$$

b) Resduos projetados

Definindo a matriz de projeção ortogonal no subespaço tangente a $\hat{\boldsymbol{\eta}}$, expressa por

$$H = \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}',$$

em que $\tilde{X} = \{\partial\eta_i(\boldsymbol{\beta})/\partial\beta_r\}$ é a matriz Jacobiana da transformação de $\boldsymbol{\eta}(\boldsymbol{\beta})$ em $\boldsymbol{\beta}$ a qual deve ser avaliada em $\hat{\boldsymbol{\beta}}$. Essa matriz é uma matriz de projeção local pois é estimada em $\hat{\boldsymbol{\beta}}$. O critério $\hat{h}_{ii} \geq 2p/n$ continua sendo adotado como guia para detectar pontos suspeitos de serem influentes (CORDEIRO; PRUDENTE; DEMÉTRIO, 2009).

Segundo (CORDEIRO; PRUDENTE; DEMÉTRIO, 2009) o resíduo projetado é dado por:

$$rp = (I - H_2)\epsilon \quad (2.19)$$

onde $H_2 = V(V'V)^{-1}V'$, sendo $V = (\tilde{X}, S)$, $S = (I - H)T$ e $T = (\tilde{x}_{11}, \dots, \tilde{x}_{pp})$.

c) Resíduos projetados estudentizados

Ainda segundo (CORDEIRO; PRUDENTE; DEMÉTRIO, 2009) os resíduos projetados estudentizados são definidos como,

$$s_i = \frac{\{(I - \hat{H}_2)\epsilon\}_i}{\tilde{\sigma}(I - \hat{H}_2)_{ii}^{1/2}} \quad (2.20)$$

onde $\tilde{\sigma} = (\epsilon'(I - \hat{H}_2)\epsilon / \text{tr}(I - \hat{H}_2))^{1/2}$.

2.2.2 Análise de diagnóstico

Em alguns casos, um pequeno número de observações de um determinado conjunto de dados exerce uma influência desproporcional nos coeficientes do modelo e em suas propriedades. Em casos extremos, os parâmetros estimados dependem mais deste pequeno conjunto de dados atípico do que da maioria dos dados. Essa situação é nitidamente indesejável; aspira-se que o modelo de regressão seja igualmente representativo para todas as observações, não apenas para poucas. Consequentemente, torna-se necessário encontrar esses pontos influentes e verificar seu impacto no modelo. Se esses pontos influentes forem realmente “ruins”, então eles deverão ser retirados da amostra. Caso contrário, se não houver nada de errado com esses pontos, então eles devem permanecer no conjunto de dados, mas seu efeito deve ser conhecido antes da utilização do modelo de regressão.

As estatísticas mais utilizadas para a verificação de pontos influentes são:

- Medida de inconsistência ($re_{(i)}$):

Um ponto é considerado inconsistente se $s_{(i)}$ é grande, isto é, tal que $|s_{(i)}| \geq$

$t_{\gamma/(2n);n-p-1}$, com nível de significância igual a $100\gamma\%$;

- Medida de alavanca (h_{ii}):

Um ponto é considerado de alavanca se h_{ii} é grande, ou seja, tal que $h_{ii} \geq 2p/n$. Pode ser classificado como bom, quando consistente, ou ruim, quando inconsistente;

- Medidas de influência geral:

A distância de Cook mede a alteração provocada no valor do ajuste pela retirada da observação i . É dada por

$$D_{(i)} = \frac{(\hat{\beta} - \hat{\beta}_{(i)})'((X^{(nt)})'X^{(nt)})(\hat{\beta} - \hat{\beta}_{(i)})}{ps^2} \quad (2.21)$$

onde,

$$\beta_{(i)} = \beta - \frac{(\tilde{X}'\tilde{X})^{-1}\tilde{x}_i\epsilon_i}{1 - \hat{h}_{ii}},$$

e $\tilde{x}_i$ é a i -ésima linha de $\tilde{X}$.

Um ponto é considerado influente se $D_{(i)}$ for grande, isto é, se $D_{(i)} \geq 2\sqrt{p/n}$.

2.3 Aplicação do método clássico

Os dados utilizados para testar a metodologia proposta foram coletados em florestas plantadas de clones de *Eucalyptus sp.* da Votorantim Celulose e Papel (VCP), na região de Itapetininga, sul do estado de São Paulo, com coordenadas geográficas $23^{\circ}56'S$ de latitude e $42^{\circ}21'W$ de longitude, a uma altitude média de $655m$.

Nesta região a precipitação média anual é de $1600mm$ e temperatura média anual de $20^{\circ}C$. Os solos são predominantemente Latossolo Vermelho distrófico (GOLFARI; CASER; MOURA, 1978).

A amostragem na região de estudo é sistemática com intensidade amostral de uma parcela a cada $10ha$. As parcelas são mensuradas anualmente a partir do segundo ano dos povoamentos. Estas unidades amostrais são circulares de $500m^2$, nas quais, são medidos os diâmetros a $1,3m$ do solo de todas as árvores, a altura total de 10 árvores centrais e de 5 dominantes (árvores com maiores diâmetros e conseqüentemente com maiores alturas).

Foram utilizadas várias parcelas locadas em povoamentos com idades variando de dois a oito anos, das quais foram selecionadas quatro (vide Tabela (2.1)) para exemplificar os resultados obtidos.

Tabela 2.1: Descrição das parcelas utilizadas na aplicação.

Parcela	Idade (anos)	Espaçamento	Número de fustes*
<i>A</i>	4,34	3,0x2,0m	14
<i>B</i>	4,08	3,0x3,0m	12
<i>C</i>	6,12	3,0x3,3m	15
<i>D</i>	5,10	3,0x2,0m	14

* Tronco de árvore.

As estimativas dos parâmetros obtidas através do método escore de Fisher, os intervalos de confiança assintótico e Bootstrap, e o critério de seleção de modelos BIC para os modelos hipsométricos dados na Tabela (1.1) foram calculados através do software R (R Development Core Team, 2010) e são apresentados na Tabela (2.2).

Em geral, o método escore de Fisher produziu bons resultados para os modelos hipsométricos descritos na Tabela (1.1), sempre que os dados utilizados possuíam coerência biológica. Quando isso não ocorre, as estimativas do(s) parâmetro(s) dos modelos apresentava(m) sinal negativo(s), inviabilizando seu uso para estimar a altura das demais árvores da parcela em estudo. Isso pode ser visto nas parcelas B e C para todos os modelos e no modelo de Prodan para a parcela A. Nesses casos os parâmetros negativos resultam em uma relação biologicamente inconsistente para representar a relação diâmetro-altura. Dessa forma, esses ajustes não podem ser utilizados, pois não geraram estimativas coerentes para a altura das árvores.

No caso da parcela B (Tabela (2.2)), o intervalo de confiança assintótico para o parâmetro β_0 no modelo de Prodan engloba o valor zero. Isso mostra que esse parâmetro não possui significância estatística e/ou que a teoria assintótica não pode ser aplicada para fazer inferência sobre estes parâmetros, pois gera intervalos de confiança insatisfatórios. Isso ocorre porque, em geral, nesse problema a suposição que é feita sobre possuir uma amostra grande é violada. Como o tamanho da amostra é pequeno, construir os intervalos de confiança utilizando a técnica Bootstrap mostra-se mais adequado, pois esses intervalos não utilizam a teoria de grandes amostras e não necessitam de nenhuma suposição sobre distribuições assintóticas para sua construção. Isso fica evidenciado no modelo de Prodan para a parcela D (Tabela (2.2)) em que o intervalo de confiança Bootstrap é de menor amplitude do que o intervalo de confiança assintótico e não contém o valor zero.

Para as quatro parcelas utilizadas neste texto, o melhor modelo pelo critério BIC foi o modelo de Petterson, como pode ser visto na Tabela (2.2). Os gráficos desse modelo para as parcelas A, B, C e D são apresentados nas Figuras (2.1), (2.2), (2.3) e (2.4), respectivamente. Os gráficos para os demais modelos são mostrados no apêndice (A).

Tabela 2.2: Estimativas dos parâmetros pelo método de máxima verossimilhança para os modelos hipsométricos dados na Tabela (1.1)

Par.	Modelo	Parâmetros	IC(95%) Bootstrap	IC(95%) Assintótico	BIC
A	1	$\beta_0 = 3,5222$	(3, 5213; 3, 5233)	(3, 5193; 3, 5252)	87, 2366
		$\beta_1 = 6,1521$	(6, 1399; 6, 1643)	(5, 7074; 6, 5966)	
	2	$\beta_0 = 1,6694$	(1, 6647; 1, 6737)	(1, 6316; 1, 7073)	80, 4487
		$\beta_1 = 0,5329$	(0, 5311; 0, 5347)	(0, 5271; 0, 5386)	
	3	$\beta_0 = 0,3071$	(0, 3043; 0, 3099)	(0, 3070; 0, 3072)	2, 7373
		$\beta_1 = 0,7143$	(0, 6790; 0, 7506)	(0, 7048; 0, 7238)	
	4	$\beta_0 = 5,3032$	(4, 9980; 5, 5802)	(3, 6897; 6, 9167)	20, 7905
		$\beta_1 = -0,6514$	(-0, 6983; -0, 5981)	(-0, 7016; -0, 6012)	
		$\beta_2 = 0,0647$	(0, 0625; 0, 0667)	(0, 0646; 0, 0648)	
B	1	$\beta_0 = 2,8308$	(2, 8195; 2, 8439)	(2, 6850; 2, 9766)	63, 7717
		$\beta_1 = -3,9396$	(-4, 1203; -3, 7263)	(-40, 9055; 33, 0263)	
	2	$\beta_0 = 3,7576$	(3, 7244; 3, 7889)	(2, 6414; 4, 8736)	63, 7613
		$\beta_1 = -0,2453$	(-0, 2567; -0, 2333)	(-0, 3905; -0, 1000)	
	3	$\beta_0 = 0,3904$	(0, 3533; 0, 4293)	(0, 3880; 0, 3927)	-11,3978
		$\beta_1 = -0,5119$	(-1, 1240; 0, 0827)	(-1, 0962; 0, 0723)	
	4	$\beta_0 = -9,6642$	(-83, 574; 64, 359)	(-9086, 161; 9066, 833)	-8, 8912
		$\beta_1 = 1,0189$	(-8, 2519; 10, 2698)	(-142, 6267; 144, 6646)	
		$\beta_2 = 0,02023$	(-0, 2687; 0, 3093)	(-0, 1209; 0, 1614)	
C	1	$\beta_0 = 3,3764$	(3, 3752; 3, 3775)	(3, 3654; 3, 3873)	107, 8634
		$\beta_1 = -1,2161$	(-1, 2426; -1, 1886)	(-7, 1431; 4, 7108)	
	2	$\beta_0 = 3,6065$	(3, 6028; 3, 6102)	(3, 4944; 3, 7185)	107, 9262
		$\beta_1 = -0,0564$	(-0, 0576; -0, 0552)	(-0, 0677; -0, 0451)	
	3	$\beta_0 = 0,3242$	(0, 3202; 0, 3282)	(0, 3240; 0, 3243)	4, 9464
		$\beta_1 = -0,1233$	(-0, 2164; -0, 0323)	(-0, 1901; -0, 0565)	
	4	$\beta_0 = 17,0885$	(9, 9086; 24, 9146)	(-492, 6162; 526, 7932)	8, 7625
		$\beta_1 = -1,5391$	(-2,2287; -0, 8941)	(-5, 4844; 2, 4062)	
		$\beta_2 = 0,0668$	(0, 0527; 0, 0818)	(0, 0649; 0, 0687)	
D	1	$\beta_0 = 3,3930$	(3, 3928; 3, 3932)	(3, 3925; 3, 3935)	106, 3011
		$\beta_1 = 5,6844$	(5, 6810; 5, 6876)	(5, 5601; 5, 8088)	
	2	$\beta_0 = 1,8969$	(1, 8957; 1, 8982)	(1, 8874; 1, 9066)	97, 7404
		$\beta_1 = 0,4057$	(0, 4053; 0, 4061)	(0, 4045; 0, 4069)	
	3	$\beta_0 = 0,3198$	(0, 3192; 0, 3205)	(0, 3197; 0, 3199)	20, 0477
		$\beta_1 = 0,6986$	(0, 6868; 0, 7097)	(0, 6956; 0, 7016)	
	4	$\beta_0 = 1,0719$	(0, 8479; 1, 289)	(-0, 1889; 2, 3327)	23, 0988
		$\beta_1 = 0,1319$	(0, 1005; 0, 1633)	(0, 1054; 0, 1586)	
		$\beta_2 = 0,0354$	(0, 0343; 0, 0364)	(0, 0353; 0, 0355)	

1 = Modelo de Curtis, 2 = Modelo de Stofel, 3 = Modelo de Petterson, 4 = Modelo de Prodan, Par. é a parcela utilizada, IC o intervalo de confiança e BIC é o critério de seleção de modelos de Schwarz.

Neste momento vamos nos ater somente aos casos em que é possível a utilização do método clássico de máxima verossimilhança, ou seja, as parcelas A e D do modelo selecionado de Petterson os quais possuem consistência biológica. Nas Figuras (2.1) e (2.4), parcelas A e D respectivamente, podemos visualizar o ajuste do modelo, os resíduos ordinário e estudentizado, o gráfico q-q plot, assim como os gráficos que representam os pontos de alavanca e a medida da distância de Cook. No caso da parcela A, Figura (2.1), percebemos no gráfico F a presença de pontos influentes (2, 5, 10 e 12). Esses pontos são também considerados consistentes pois seus resíduos são pequenos e sempre inferiores a aos limites $(-t_{0,05/(2n);n-p-1}; t_{0,05/(2n);n-p-1})$, como pode ser visto no gráfico C. Pelo gráfico D, podemos notar que a suposição de normalidade para a variável resposta está satisfeita.

No caso da parcela D (Figura (2.4)) os resultados são semelhantes aos obtidos na parcela A. A suposição de normalidade foi satisfeita e não existem pontos aberrantes que possam estar mascarando os resultados obtidos com o ajuste. Notem no gráfico (2.4) C que embora exista um ponto fora dos limites, o ponto 13, o mesmo possui resíduo ordinário menor que um metro, valor esse aceito no âmbito florestal como erro admissível nas medições. Dessa forma não existe a necessidade de retirada deste ponto para a realização do ajuste.

Em todas as parcelas utilizadas cujos dados não apresentavam coerência biológica, nem um dos modelos foram bem ajustados pelo método clássico e conseqüentemente não podem ser usados de forma adequada para o planejamento da exploração dos recursos florestais. Para que o uso de algum modelo seja possível, restrições sobre os parâmetros devem ser consideradas. Dessa forma, a abordagem bayesiana apresenta-se como uma abordagem conveniente para tratar o problema de ajuste de relações hipsométricas com restrições impostas aos parâmetros.

Figura 2.1: Ajuste do modelo de Petterson para a parcela A (A), resíduo ordinário *versus* Dap (B), resíduo estudentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura 2.2: Ajuste do modelo de Peterson para a parcela B (A), resíduo ordinário *versus* Dap (B), resíduo estudentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).

Figura 2.3: Ajuste do modelo de Petterson para a parcela C (A), resíduo ordinário *versus* Dap (B), resíduo estudntizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura 2.4: Ajuste do modelo de Peterson para a parcela D (A), resíduo ordinário *versus* Dap (B), resíduo studentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

3 *Abordagem bayesiana*

Como discutido no capítulo anterior, as estimativas dos parâmetros de forma clássica são obtidas por método iterativo e a inferência sobre os parâmetros são obtidas de maneira assintótica, o que para amostras pequenas (como é o caso), cálculos de intervalos de confiança, testes de hipóteses, previsões, etc. podem não ser confiáveis.

Dessa forma, uma abordagem que incorpore algum conhecimento *a priori* sobre os parâmetros é sempre preferível. Nesse sentido, propomos uma abordagem bayesiana para resolver o problema de ajuste e inferência nos parâmetros sujeitos a certas restrições. Neste contexto assume-se que estes parâmetros são variáveis aleatórias e qualquer informação inicial sobre eles pode ser modelada por uma função densidade de probabilidade *a priori*. Combinando-se estas densidades *a priori* com a função de verossimilhança, através do teorema de Bayes, chega-se a função densidade de probabilidade *a posteriori*. Denotando-se o vetor de parâmetros de um modelo por $\boldsymbol{\theta}$, a densidade *a priori* por $\pi_0(\boldsymbol{\theta})$ e a função de verossimilhança associada a um conjunto de observações $\mathbf{D}$ por $L(\boldsymbol{\theta}|\mathbf{D})$, então a função densidade de propabilidade *a posteriori* é dada por:

$$\pi(\boldsymbol{\theta}|\mathbf{D}) = \frac{L(\boldsymbol{\theta}|\mathbf{D})\pi_0(\boldsymbol{\theta})}{\int_{\Theta} L(\boldsymbol{\theta}|\mathbf{D})\pi_0(\boldsymbol{\theta})d\boldsymbol{\theta}} \quad (3.1)$$

A integral em (3.1) é uma integral múltipla sobre o domínio de definição dos parâmetros $\boldsymbol{\theta} \in \Theta$ que representa a constante normalizadora da densidade *a posteriori*, portanto é uma função somente dos dados $\mathbf{D}$ observados. É comum se adotar a notação de proporcionalidade para representar a densidade *a posteriori*, como faremos neste texto, dada por:

$$\pi(\boldsymbol{\theta}|\mathbf{D}) \propto L(\boldsymbol{\theta}|\mathbf{D})\pi_0(\boldsymbol{\theta}). \quad (3.2)$$

A inferência bayesiana tem sido amplamente utilizada em diversas áreas de conhecimento, como por exemplo na modelagem de crescimento (veja Filho et al. (2008), Andrade et al. (2010), amaral (2008)), na estimativa da altura de Eucalypto (MOREIRA et al., 2007; BARROS, 2010) dentre outros. Para mais detalhes sobre inferência bayesiana ver por

exemplo (BOX; TIAO, 1973).

3.1 Densidade de probabilidade *a priori*

Um dos grandes proveitos do uso de métodos bayesianos, em relação a métodos clássicos, é que pode-se incorporar na análise, informação *a priori* sobre os parâmetros. Através dela, a opinião *a priori* de um especialista pode ser incorporada na forma de uma distribuição de probabilidade. Quando não existem informações *a priori* palpáveis ou de qualquer natureza substancial, as distribuições *a priori* não informativas são as mais indicadas.

Neste trabalho será considerada *a priori* não informativa de Jeffreys, uma distribuição *a priori* vaga normal *flat* assim como outras duas formas de distribuição *a priori* informativas que serão detalhadas a seguir.

A implementação dos algoritmos MCMC para a obtenção das estimativas dos parâmetros bayesianos, assim como os critérios de seleção utilizados foram realizados no software R (R Development Core Team, 2010).

3.1.1 Densidade de probabilidade *a priori* não informativa de Jeffreys

Nesta seção, será considerada para a análise bayesiana proposta, *a priori* de Jeffreys, que se baseia no uso da medida de informação de Fisher sobre $\boldsymbol{\theta} \in \mathfrak{R}$, de elementos

$$I_{sr}(\boldsymbol{\beta}) = E \left[\left(\frac{\partial \ln f(X|\boldsymbol{\beta})}{\partial \beta_s} \frac{\partial \ln f(X|\boldsymbol{\beta})}{\partial \beta_r} \right) \middle| \boldsymbol{\beta} \right],$$

e é definida como uma densidade proporcional à raiz quadrada do determinante de $I(\boldsymbol{\beta})$,

$$\pi(\boldsymbol{\beta}) \propto |I(\boldsymbol{\beta})|^{\frac{1}{2}}. \quad (3.3)$$

Considerando os erros independentes e identicamente distribuídos, $e_i \sim N(0, 1/\tau)$, a forma geral da matriz de informação de Fisher $I(\boldsymbol{\theta})$, para os modelos da Tabela (1.1), pode ser calculada como

$$I_{sr} = -E \left\{ \frac{\partial^2 l(\mathbf{D}|\boldsymbol{\beta})}{\partial \beta_s \partial \beta_r} \right\} = -\frac{1}{\sigma^2} \sum_{i=1}^n \frac{\partial \eta(x_i; \boldsymbol{\beta})}{\partial \beta_s} \frac{\partial \eta(x_i; \boldsymbol{\beta})}{\partial \beta_r} \quad (3.4)$$

A Tabela (3.1) a seguir, mostra as derivadas para construção das densidades *a priori* de Jeffreys para cada um dos modelos.

Tabela 3.1: Derivada dos modelos de hipsometria.

Autor	$\frac{\partial \eta(x; \beta)}{\partial \beta_0}$	$\frac{\partial \eta(x; \beta)}{\partial \beta_1}$	$\frac{\partial \eta(x; \beta)}{\partial \beta_2}$
Stofel	1	x	—
Curtis	1	$-x$	—
Petterson	B	Bx	—
Prodan	zx^2	zx	z
$B = -3(\beta_0 + \beta_1 x)^{-4}; z = -(\beta_0 x^2 + \beta_1 x + \beta^2)^{-2}.$			

Adotando uma *priori* não informativa para τ dada por $\pi_0(\tau) \propto 1/\tau$, a densidade *a priori* de Jeffreys para os parâmetros β e τ , assumindo independência entre eles são mostradas na Tabela (3.2).

Tabela 3.2: Densidades *a priori* não informativa de Jeffreys assumindo independência.

Autor	densidades $\pi_0(\beta, \tau) \propto$
Stofel	$1/\tau$
Curtis	$1/\tau$
Petterson	$\tau^{-1} \sqrt{\sum_{i=1}^n B_i^2(\beta) \sum_{i=1}^n x_i^2 B_i^2(\beta)}$
Prodan	$\tau^{-1} \sqrt{\sum_{i=1}^n z_i^2(\beta) x_i^4 \sum_{i=1}^n z_i^2(\beta) x_i^2 \sum_{i=1}^n z_i^2(\beta)}$

3.1.1.1 Densidade de probabilidade *a posteriori* obtida através da *priori* não informativa de Jeffreys

A função densidade de probabilidade *a posteriori* para β e τ é encontrada através da equação (3.2).

Utilizando as densidades *a priori* da Tabela (3.2), a densidade de probabilidade *a posteriori* conjunta para β e τ é dada por:

$$\pi(\beta, \tau | \mathbf{D}) \propto \tau^{\frac{n}{2}} \exp \left\{ -\frac{\tau}{2} (\mathbf{Y}'\mathbf{Y} + \mathbf{G}(\beta)) \right\} \pi_0(\beta) \pi_0(\tau), \quad (3.5)$$

em que $\mathbf{G}(\beta) = -2\boldsymbol{\mu}(\mathbf{x}; \beta)'[\mathbf{Y} + (1/2)\boldsymbol{\mu}(\mathbf{x}; \beta)]$. A densidade *a posteriori* dos parâmetros β e τ não tem um núcleo que possa ser identificado como uma das distribuições tradicionalmente conhecidas e os seus momentos só podem ser obtidos numericamente. Para calcular os estimadores bayesianos dos parâmetros dos modelos analisados, será utilizado o algoritmo de simulação de Monte Carlo em Cadeia de Markov (MCMC). A construção desse algoritmo é baseada nas densidades de probabilidade condicionais *a posteriori*, dadas na Tabela (3.3).

O algoritmo de Metropolis-Hastings utilizado para gerar amostras da densidade *a posteriori* conjunta é dado a seguir:

Tabela 3.3: Densidades *a posteriori* condicionais obtidas utilizando *a priori* não informativa de Jeffreys.

Autor	$\pi(\boldsymbol{\beta} \tau, \mathbf{D}) \propto$	$\pi(\tau \boldsymbol{\beta}, \mathbf{D}) \propto$
Stofel	$\exp\{c\}$	$Gama(a, b(\boldsymbol{\beta}))$
Curtis	$\exp\{c\}$	$Gama(a, b(\boldsymbol{\beta}))$
Petterson	$\exp\{c\} \sqrt{\sum_{i=1}^n B_i^2(\boldsymbol{\beta}) \sum_{i=1}^n x_i^2 B_i^2(\boldsymbol{\beta})}$	$Gama(a, b(\boldsymbol{\beta}))$
Prodan	$\exp\{c\} \sqrt{\sum_{i=1}^n z_i^2(\boldsymbol{\beta}) x_i^4 \sum_{i=1}^n z_i^2(\boldsymbol{\beta}) x_i^2 \sum_{i=1}^n z_i^2(\boldsymbol{\beta})}$	$Gama(a, b(\boldsymbol{\beta}))$

$c = -\frac{\tau}{2} \mathbf{G}(\boldsymbol{\beta}), a = \frac{n}{2}, b(\boldsymbol{\beta}) = \frac{1}{2}(\mathbf{Y}'\mathbf{Y} + \mathbf{G}(\boldsymbol{\beta}))$.

• Algoritmo *Metrópolis-Hasting*:

1. Faça $j = 0$ e considere a condição inicial para $\tau^{(j)}$ e $\boldsymbol{\beta}^{(j)}$.
2. Gere $\boldsymbol{\beta}^{(j+1)}$ de um núcleo normal $N(\mu_0, \Sigma_0)$ com $\mu_0 \in \mathbb{R}^p$ e $\Sigma_0 = \text{diag}(\sigma_1^2, \dots, \sigma_p^2)$, em que μ_0 e Σ_0 são hiperparâmetros conhecidos.
3. Calcula-se a probabilidade de aceitação de $\boldsymbol{\beta}^{(j+1)}$, por $\alpha = \min\{1, p\}$, em que

Autor	p
Stofel / Curtis	p_{sc}
Petterson	$p_{sc} \frac{\sqrt{\sum_{i=1}^n B_i^2(\boldsymbol{\beta}^{(j+1)}) \sum_{i=1}^n x_i^2 B_i^2(\boldsymbol{\beta}^{(j+1)})}}{\sqrt{\sum_{i=1}^n B_i^2(\boldsymbol{\beta}^{(j)}) \sum_{i=1}^n x_i^2 B_i^2(\boldsymbol{\beta}^{(j)})}}$
Prodan	$p_{sc} \frac{\sqrt{\sum_{i=1}^n z_i^2(\boldsymbol{\beta}^{(j+1)}) x_i^4 \sum_{i=1}^n z_i^2(\boldsymbol{\beta}^{(j+1)}) x_i^2 \sum_{i=1}^n z_i^2(\boldsymbol{\beta}^{(j+1)})}}{\sqrt{\sum_{i=1}^n z_i^2(\boldsymbol{\beta}^{(j)}) x_i^4 \sum_{i=1}^n z_i^2(\boldsymbol{\beta}^{(j)}) x_i^2 \sum_{i=1}^n z_i^2(\boldsymbol{\beta}^{(j)})}}$

$p_{sc} = \exp\left\{-\frac{\tau}{2}[\mathbf{G}(\boldsymbol{\beta}^{(j+1)}) - \mathbf{G}(\boldsymbol{\beta}^{(j)})]\right\}$.

4. Gere u de uma distribuição uniforme $U(0, 1)$ e use a seguinte regra de aceitação:

$$\boldsymbol{\beta}^{(j+1)} = \begin{cases} \boldsymbol{\beta}^{(j+1)} & \text{se } u \leq \alpha \\ \boldsymbol{\beta}^{(j)} & \text{se } u > \alpha \end{cases}$$

5. Gere $\tau^{(j)}$ da posteriori condicional gama, $\tau^{(j)} \sim Gama(a, b(\boldsymbol{\beta}^{(j)}))$
6. Faça $j \leftarrow j + 1$ e volte para o passo (2).

Para encontrar as estimativas bayesianas obtidas via simulação de Monte Carlo será adotado o procedimento sugerido em Gilks (GILKS; RICHARDSON; SPIEGELHALTER, 1998). Considera-se um período de aquecimento (burn-in) de 50% das iterações iniciais. Os passos 2 e 3 do algoritmo serão repetidos até que haja evidências de convergência, monitorada graficamente e com o critério de Geweke (GEWEKE, 1992). Então, uma amostra será selecionada a cada s iterações, a qual será obtida com base no cálculo da amostra essencial em função da autocorrelação entre os valores gerados. O tamanho da amostra essencial é

dado por:

$$AE_k = \frac{M}{a_k} \quad (3.6)$$

em que M é o tamanho da amostra gerada pelo MCMC após a retirada do *burn in* e

$$a_k = \left\{ 1 + 2 \sum_{s=1}^M \left(1 - \frac{s}{M} \right) \rho_{ks} \right\} \quad (3.7)$$

sendo ρ_{ks} a autocorrelação da amostra gerada pelo MCMC de defasagem (lag) s .

Definindo $\theta^{(j)} = \{\beta_0^{(j)}, \beta_1^{(j)}, \tau^{(j)}\}$, $j = 1, \dots, M$ para os modelos de Curtis, Stofel e Petterson e $\theta^{(j)} = \{\beta_0^{(j)}, \beta_1^{(j)}, \beta_2^{(j)}, \tau^{(j)}\}$, $j = 1, \dots, M$ para o modelo de Prodan, a amostra resultante do algoritmo, os estimadores bayesianos obtidos via simulação de Monte Carlo são dados por

$$\widehat{g}(\theta_s^{MC}) = \frac{1}{M} \sum_{j=1}^M g(\theta_s^{(j)}). \quad (3.8)$$

E os intervalos de credibilidade $[\widehat{\theta}_s^{MC}(\alpha), \widehat{\theta}_s^{MC}(1 - \alpha)]$ serão estimados com base nos percentis da amostra, dados por:

$$P[\widehat{\theta}_s^{MC}(\alpha) \leq \theta_s^{MC} \leq \widehat{\theta}_s^{MC}(1 - \alpha)] = 1 - \alpha \quad (3.9)$$

3.1.1.2 Aplicação

Os dados utilizados para testar a abordagem bayesiana com distribuição *a priori* de Jeffreys foram coletados em florestas plantadas de clones de *Eucalyptus sp.* descritos na seção 2.3. Dentre as várias parcelas utilizadas, foram selecionadas três, as parcelas A, B, C (Tabela (2.1)) para exemplificar os resultados obtidos. Não será apresentado os resultados para a parcela D, como foi feito no capítulo anterior, porque os resultados dessa parcela com a metodologia bayesiana são muito similares aos obtidos com a parcela A.

Os resultados das estimativas bayesianas são apresentadas na Tabela (3.4) assim como o critério de convergência de Geweke (1992) utilizado para verificar a convergência da cadeia gerada pelo algoritmo de Metrópolis-Hastings, lembrando que como critério de Geweke a convergência ocorre quando $|Z| < Zc$ para $Zc = 1,96$. Podemos verificar nessa tabela que nas parcelas A, para o modelo de Prodan, B e C a estimativa do parâmetro $\beta_i, i = 0, 1, 2$ foi negativa. Isso provoca resultados biologicamente inconsistentes, inviabilizando seu uso. Esse resultado pode ser verificado nas Figuras (3.1), (3.2) e (3.3) onde o parâmetro estimado negativo resulta em um ajuste decrescente.

Podemos verificar na Tabela (3.4) para a parcela A, modelo de Curtis, que mesmo

havendo coerência biológica nos dados da parcela, o intervalo de credibilidade para o parâmetro β_1 engloba o valor zero. Dessa forma seu possível uso é descartado, principalmente se compararmos esse resultado com o obtido com o método clássico de máxima verossimilhança (Tabela (2.2)), no qual o intervalo de confiança é de menor amplitude e não engloba o valor zero.

Para a seleção de modelos, foram utilizados o critério de informação de Akaike (AIC), o critério de informação bayesiano (BIC), o critério desvio-informação (DIC) e o critério da densidade preditiva ordenada (maiores detalhes sobre seleção de modelo no capítulo 4). Os resultados são apresentados na Tabela (3.5). Para as três parcelas utilizadas o melhor modelo selecionado por todos os critérios foi o modelo de Petterson.

Como exemplo, apresentamos o histograma e trajetória das amostras geradas com o algoritmo de Metrópolis-Hastings após verificada a convergência, para os parâmetros β e τ nas Figuras (3.4) e (3.5), respectivamente. Essa convergência para o parâmetro β pode ser observada nos gráficos *A1* e *B1* nos quais se verifica a tendência estacionária das séries simuladas dos parâmetros em torno das estimativas. Os histogramas *A2* e *B2* nos mostra que as densidades *a posteriori* continuam normalmente distribuídas entre os valores estimados, tratando-se de uma distribuição unimodal. Na Figura (3.5) *A1* temos a trajetória das amostras geradas para o parâmetro τ pelo algoritmo MCMC o qual nos indica que a convergência da série foi alcançada. O histograma *A2* nos mostra que a distribuição *a posteriori* para τ é assimétrica positiva e unimodal. O resultado para as demais parcelas e modelos apresentam resultados semelhantes.

As estimativas bayesianas utilizando densidade *a priori* de Jeffreys não produziram resultados satisfatórios quando comparadas as estimativas de máxima verossimilhança, pois embora as estimativas pontuais tenham quase sempre coincidido, os intervalos de credibilidade, em geral, foram de maior amplitude do que os intervalos de confiança clássicos encontrados (veja por exemplo o modelo de Curtis para a parcela A, Tabela (2.2) e Tabela (3.4)). Neste contexto, nosso próximo passo será utilizar uma densidade de probabilidade *a priori* normal *flat* para tentar melhorar os resultados obtidos até o presente momento.

Tabela 3.4: Estimativas bayesianas utilizando densidade *a priori* de Jeffreys para os modelos hipsométricos dados na Tabela (1.1).

Par.	Modelo	Est. Bayesiana	D.P.	ICred.(95%)	CCG(Z)
A	Curtis	$\beta_0 = 3,5165$	0,8081	(1,9714; 5,0426)	0,0360
		$\beta_1 = 6,1082$	9,9328	(-13,2789; 25,1234)	0,0394
		$S_2 = 0,6752$	1,6297	(0,0344; 4,7969)	0,0441
	Stofel	$\beta_0 = 1,6643$	0,6963	(0,3079; 3,0273)	0,0002
		$\beta_1 = 0,5351$	0,2703	(0,0056; 1,0604)	0,0029
		$S_2 = 2,9077$	1,2305	(0,3332; 11,3178)	0,0264
	Petterson	$\beta_0 = 0,3071$	0,0032	(0,3005; 0,3135)	0,0286
		$\beta_1 = 0,7146$	0,0429	(0,6304; 0,8019)	0,0331
		$S_2 = 0,0109$	0,0863	(0,0179; 0,0894)	0,0385
	Prodan	$\beta_0 = 5,3066$	0,3512	(4,6078; 5,9879)	0,0149
		$\beta_1 = -0,6519$	0,0621	(-0,7724; -0,5283)	0,0198
		$\beta_2 = 0,0647$	0,0026	(0,0594; 0,0698)	0,0224
		$S_2 = 0,1296$	0,0615	(0,0588; 0,2801)	0,0315
B	Curtis	$\beta_0 = 2,8489$	5,0198	(-6,8845; 12,7296)	0,0067
		$\beta_1 = -3,7515$	79,9554	(-158,6787; 153,7024)	0,0068
		$S_2 = 0,3354$	0,0582	(0,0163; 2,3527)	0,0387
	Stofel	$\beta_0 = 3,7451$	3,0022	(-2,2091; 9,6366)	0,0211
		$\beta_1 = -0,2407$	1,0829	(-2,3723; 1,9200)	0,02184
		$S_2 = 3,0333$	1,3965	(0,3790; 11,0577)	0,0367
	Petterson	$\beta_0 = 0,3904$	0,0041	(0,3823; 0,3985)	0,0438
		$\beta_1 = -0,5119$	0,0659	(-0,6422; -0,3846)	0,0449
		$S_2 = 0,0195$	0,0096	(0,0084; 0,0439)	0,0362
	Prodan	$\beta_0 = -9,3389$	2,7272	(-14,6201; -3,9682)	0,0201
		$\beta_1 = 0,9791$	0,3424	(0,2998; 1,6467)	0,0204
		$\beta_2 = 0,0214$	0,0108	(0,0004; 0,0429)	0,0207
		$S_2 = 0,0219$	0,0120	(0,0088; 0,0535)	0,0216
C	Curtis	$\beta_0 = 3,3658$	3,5851	(-3,8210; 105365)	0,0035
		$\beta_1 = -1,4851$	83,3377	(-168,8528; 162,1927)	0,0023
		$S_2 = 0,3792$	0,0474	(0,0125; 2,4798)	0,0065
	Stofel	$\beta_0 = 3,6327$	2,4969	(-1,3539; 8,5506)	0,01377
		$\beta_1 = -0,0648$	0,7921	(-1,6261; 1,1561)	0,0130
		$S_2 = 4,5585$	1,1729	(0,3288; 25,2759)	0,0265
	Petterson	$\beta_0 = 0,3241$	0,0024	(0,3195; 0,3288)	0,0194
		$\beta_1 = -0,1224$	0,0556	(-0,2302; -0,0142)	0,0176
		$S_2 = 0,0453$	0,0196	(0,0214; 0,0959)	0,0397
	Prodan	$\beta_0 = 17,5167$	1,3326	(14,9308; 20,1027)	0,0304
		$\beta_1 = -1,5769$	0,1171	(-1,8025; -1,3489)	0,0301
		$\beta_2 = 0,0676$	0,0026	(0,0627; 0,0726)	0,0293
		$S_2 = 0,0493$	0,0208	(0,0233; 0,1025)	0,0041

Par. é a parcela utilizada, Est. bayesiana é a estimativa bayesiana obtida via MCMC, D.P. é o desvio padrão, ICred. o intervalo de credibilidade e CCG(Z) é o critério de convergência de Geweke (1992).

Tabela 3.5: Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade *a priori* de Jeffreys para os modelos hipsométricos dados na Tabela (1.1).

Parcela	Modelos	AIC	BIC	Log(CPO)	DIC
A	Curtis	-8,29	1,55	92,83	5,14
	Stofel	8,79	18,63	82,40	8,73
	Petterson	-45,70	-35,87	108,39	-50,55
	Prodan	-28,04	-14,93	100,57	-35,01
B	Curtis	18,10	27,01	80,64	61,58
	Stofel	9,40	18,31	71,84	8,13
	Petterson	-45,43	-36,52	98,31	-50,41
	Prodan	-42,41	-30,53	97,83	-49,25
C	Curtis	16,27	26,52	80,84	71,37
	Stofel	11,07	21,32	68,07	13,55
	Petterson	-48,09	-37,85	95,63	-53,17
	Prodan	-45,14	-31,48	95,16	-52,18

AIC é o critério de informação de Akaike (1973), BIC o critério de informação bayesiano (SCHWARZ, 1978), DIC é o critério desvio-informação (SPIEGELHALTER et al., 2002) e Log(CPO) é o logaritmo do critério da densidade preditiva ordenada (GELFAND; DEY, 1994).

(A1)

(A2)

(A3)

(A4)

Figura 3.1: Ajuste bayesiano utilizando densidade *a priori* de Jeffreys para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.

(B1)

(B2)

(B3)

(B4)

Figura 3.2: Ajuste bayesiano utilizando densidade *a priori* de Jeffreys para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4.

(C1)

(C2)

(C3)

(C4)

Figura 3.3: Ajuste bayesiano utilizando densidade *a priori* de Jeffreys para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4.

Figura 3.4: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição *a priori* de Jeffreys.

Figura 3.5: Trajetória (A1) e histograma (A2) das amostras geradas para σ^{-2} com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela A, com distribuição *a priori* de Jeffreys.

3.1.2 Densidade de probabilidade *a priori* não informativa Normal *flat*

Nesta seção será considerada uma densidade de probabilidade Normal *flat* para o vetor de parâmetro β , dada por $N(\hat{\beta}, (\alpha\Sigma)^{-1})$. Assim temos:

$$\pi_0(\beta) \propto \exp \left\{ -\frac{1}{2}(\beta - \hat{\beta})' \alpha \Sigma (\beta - \hat{\beta}) \right\} \quad (3.10)$$

em que $\hat{\beta} = (X'X)^{-1}X'Y$ obtido através da aproximação de Laplace (MAZUCHELI, 1995), Σ é a inversa da matrix Hessiana (BAILEY, 2001; BHATTACHARJEE; DUNSMORE, 1991; DANIELS; POURAHMADI, 2002) e α é um fator que controla a variância da densidade *a priori*, podendo torná-la mais ou menos flat, conforme as informações disponíveis. Para τ vamos considerar a densidade *a priori*

$$\pi_0(\tau) \propto \frac{1}{\tau}, \quad (3.11)$$

A priori conjunta considerando independência entre os parâmetros β e τ é dada por

$$\pi_0(\beta, \tau) = \pi_0(\beta|\tau)\pi_0(\tau) \quad (3.12)$$

3.1.2.1 Densidade de probabilidade *a posteriori* obtida através da *priori* não informativa Normal *flat*

A função densidade de probabilidade *a posteriori* de β e τ é encontrada através da equação (3.2).

Assumindo a priori não informativa dada pela equação (3.12), a qual é obtida substituindo as equações (3.10) e (3.11), a função densidade de probabilidade *a posteriori* para os modelos de Curtis e Stofel será:

$$\pi(\beta, \tau | \mathbf{D}) = \tau^{\frac{n}{2}-1} \exp \left\{ -\frac{\tau}{2} \left[(\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) + S^2(\hat{\beta}) \right] \right\} \cdot \exp \left\{ -\frac{1}{2}(\beta - \hat{\beta})' \alpha \Sigma (\beta - \hat{\beta}) \right\}$$

simplificando, temos:

$$\pi(\beta, \tau | \mathbf{D}) = \tau^{\frac{n}{2}-1} \exp \left\{ -\frac{1}{2} \left[(\beta - \hat{\beta})' (\tau X' X + \alpha \Sigma) (\beta - \hat{\beta}) \right] - \frac{\tau}{2} S^2(\hat{\beta}) \right\} \quad (3.13)$$

Para os modelos de Petterson e Prodan a função densidade de probabilidade *a posteriori* é obtida substituindo as equações (2.1) e (3.12) (utilizando (3.10) e (3.11)) na

equação (3.2), logo,

$$\pi(\boldsymbol{\beta}, \tau | \mathbf{D}) \propto \left(\frac{\tau}{2\pi}\right)^{n/2-1} \exp\left\{-\frac{\tau}{2}[\mathbf{Y}'\mathbf{Y} - \mathbf{G}(\boldsymbol{\beta})]\right\} \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \alpha \Sigma (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right\} \quad (3.14)$$

onde $\mathbf{G}(\boldsymbol{\beta}) = 2\eta(\mathbf{x}; \boldsymbol{\beta})'[\mathbf{Y} - \frac{1}{2}\eta(\mathbf{x}; \boldsymbol{\beta})]$.

3.1.2.2 Densidade de probabilidade *a posteriori* marginais

As densidades *a posteriori* marginais para os parâmetros dos modelos de Curtis e Stofel são encontradas considerando-se um algoritmo de simulação de Monte Carlo em Cadeia de Markov (MCMC) que utiliza as densidades condicionais *a posteriori* dadas por:

$$\boldsymbol{\beta} | \tau, \mathbf{D} \propto N(\hat{\boldsymbol{\beta}}, (\tau X'X + \alpha \Sigma)^{-1}) \quad (3.15)$$

e

$$\tau | \boldsymbol{\beta}, \mathbf{D} \propto \text{Gama}\left(\frac{n}{2}, \frac{1}{2} \left((\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' X'X (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + S^2(\hat{\boldsymbol{\beta}})\right)\right) \quad (3.16)$$

Agora é possível elaborar o algoritmo de Gibbs Sampling para encontrar os estimadores Bayesianos de $\boldsymbol{\beta}$ e τ utilizando as densidades *a posteriori* condicionais (3.15) e (3.16), dado por:

- Algoritmo Gibbs Sampling:

1. Faça $j = 0$ e considere a condição inicial, $\tau^{(0)}$.
2. Gerar $\boldsymbol{\beta}^{(j+1)}$ de $\boldsymbol{\beta}^{(j+1)} | \tau^{(j)}, \mathbf{D} \propto N(\hat{\boldsymbol{\beta}}, (\tau X'X + \alpha \Sigma)^{-1})$
3. Gerar $\tau^{(j+1)}$ de $\tau^{(j+1)} | \boldsymbol{\beta}^{(j)}, \mathbf{D} \propto \text{Gama}\left(\frac{n}{2}, \frac{1}{2} \left((\boldsymbol{\beta}^{(j)} - \hat{\boldsymbol{\beta}})' X'X (\boldsymbol{\beta}^{(j)} - \hat{\boldsymbol{\beta}}) + S^2(\hat{\boldsymbol{\beta}})\right)\right)$
4. Faça $j \leftarrow j + 1$ e repita os passos (2) e (3) até atingir a convergência.

Para os modelos de Petterson e Prodan as densidades *a posteriori* marginais são encontradas considerando-se as densidades *a posteriori* condicionais dadas por:

$$\pi(\boldsymbol{\beta} | \tau, \mathbf{D}) \propto \exp\left\{-\frac{\tau}{2}\mathbf{G}(\boldsymbol{\beta}) - \frac{1}{2}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \alpha \Sigma (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right\} \quad (3.17)$$

e

$$\pi(\tau | \boldsymbol{\beta}, \mathbf{D}) \propto \text{Gama}\left(\frac{n}{2}, \left(-\frac{1}{2}(\mathbf{Y}'\mathbf{Y} - \mathbf{G}(\boldsymbol{\beta}))\right)\right) \quad (3.18)$$

Para esses modelos a densidade *a posteriori* do parâmetro $\boldsymbol{\beta}$ não tem um núcleo que possa ser identificado como uma das distribuições tradicionalmente conhecidas e os seus momentos só podem ser obtidos numericamente. Para calcular os estimadores bayesianos

dos parâmetros desses modelos, será utilizado o algoritmo de simulação de Metropolis-Hastings, dado por:

- Algoritmo Metrópolis-Hasting:

1. Faça $j = 0$ e considere a condição inicial para $\tau^{(j)}$ e $\beta^{(j)}$.
2. Gere $\beta^{(j+1)}$ de um núcleo normal $N(\mu, \alpha\Sigma)$ com $\mu \in \mathbb{R}^p$, em que μ e Σ são hiperparâmetros conhecidos.

3. Calcula-se a probabilidade de aceitação de $\beta^{(j+1)}$, por $\alpha = \min\{1, p\}$, em que

$$p = \exp \left\{ -\frac{\tau^{(j)}}{2} \left[G(\beta^{(j+1)}) - G(\beta^{(j)}) \right] \right\} \quad (3.19)$$

4. Gere u de uma distribuição uniforme $U(0, 1)$ e use a seguinte regra de aceitação:

$$\beta^{(j+1)} = \begin{cases} \beta^{(j+1)} & \text{se } u \leq \alpha \\ \beta^{(j)} & \text{se } u > \alpha \end{cases}$$

5. Gere $\tau^{(j+1)}$ da posteriori condicional gama, $\tau^{(j+1)} \sim G\left(\frac{n}{2}, \left(-\frac{1}{2}(\mathbf{Y}'\mathbf{Y} - G(\beta^{(j+1)}))\right)\right)$
6. Faça $j \leftarrow j + 1$ e volte para o passo (2).

Para encontrar as estimativas bayesianas obtidas via simulação de Monte Carlo será adotado o mesmo procedimento adotado na seção 3.1.1.1. Considera-se um período de aquecimento (burn-in) de 50% das iterações iniciais. Todos os passos do algoritmo serão repetidos até que haja evidências de convergência, monitorada graficamente e com o critério de Geweke (GEWEKE, 1992).

Os estimadores bayesianos obtidos via simulação de Monte Carlo são calculados pela equação (3.8) e os intervalos de credibilidade por (3.9).

3.1.2.3 Aplicação

Os dados utilizados para testar a abordagem bayesiana com distribuição *a priori* normal *flat* foram coletadas em florestas plantadas de clones de *Eucalyptus sp.* descritos na seção 2.3. Dentre as várias parcelas utilizadas, foram selecionadas três, as parcelas A, B, C (Tabela (2.1)) para exemplificar os resultados obtidos.

Os resultados das estimativas bayesianas são apresentadas na Tabela (3.6) assim como o critério de convergência de Geweke (1992) utilizado para verificar a convergência da cadeia gerada pelo algoritmo de Metrópolis-Hastings, lembrando que como critério de

Geweke a convergência ocorre quando $|Z| < Z_c$ para $Z_c = 1,96$. Podemos verificar nessa tabela que os resultados obtidos utilizando a densidade *a priori* normal *flat* são bastante semelhantes aos encontrados utilizando a densidade de probabilidade *a priori* de Jeffreys. Assim como ocorre na seção 3.1.1.2, quando não havia coerência biológica dos dados da parcela utilizada, o parâmetro $\beta_i, i = 0, 1, 2$ foi negativo, originando resultados inconsistentes que inviabilizam seu uso. Esse resultado também pode ser visto nas Figuras (3.6), (3.7) e (3.8) onde o parâmetro estimado negativo resulta em um ajuste decrescente.

Assim como ocorreu com as estimativas para a parcela A no modelo de Curtis utilizando a densidade de probabilidade *a priori* de Jeffreys, os resultados obtidos utilizando a densidade *a priori* normal *flat* (Tabela (3.6)) produziu intervalo de credibilidade para o parâmetro β_1 que engloba o valor zero. Esse resultado também descartou seu possível uso, lembrando mais uma vez que no método clássico de máxima verossimilhança o intervalo de confiança foi de menor amplitude e não conteve o valor zero (Tabela (2.2)).

Para a seleção de modelos, foram utilizados o critério de informação de Akaike (AIC), o critério de informação bayesiano (BIC), o critério desvio-informação (DIC) e o critério da densidade preditiva ordenada (maiores detalhes sobre seleção de modelo no capítulo 4). Os resultados são apresentados na Tabela (3.7). Para as três parcelas utilizadas, o melhor modelo selecionado por todos os critérios foi o modelo de Prodan. No entanto, em nenhuma das três parcelas seu uso é viável pois pelo menos um dos parâmetros ($\beta_i, i = 0, 1, 2$) foi negativo, o que provocaria, caso fossem utilizados, resultados inconsistentes biologicamente. Neste contexto, o melhor modelo selecionado por todos os critérios de seleção utilizados foi o modelo de Stofel (segundo melhor modelo). Vale lembrar que seu uso só é possível no caso da parcela A, a qual possui coerência biológica, pois os parâmetros do modelo obedecem a restrição nos parâmetros $\beta_i > 0; i = 0, 1$ imposta biologicamente.

Como exemplo, apresentamos o histograma e trajetória das amostras geradas com o algoritmo de Metrópolis-Hastings após verificada a convergência, para os parâmetros β e τ nas Figuras (3.9) e (3.10), respectivamente. Essa convergência para o parâmetro β pode ser observada nos gráficos A1 e B1 nos quais se verifica a tendência estacionária das séries simuladas dos parâmetros em torno das estimativas. Os histogramas A2 e B2 nos mostra que as densidades *a posteriori* continuam normalmente distribuídas entre os valores estimados, tratando-se de uma distribuição unimodal. Na Figura (3.10) A1 temos a trajetória das amostras geradas para o parâmetro τ pelo algoritmo MCMC o qual nos indica que a convergência da série foi alcançada. O histograma A2 nos mostra que a distribuição *a posteriori* para τ é assimétrica positiva e unimodal. O resultado para as demais parcelas e modelos apresentam resultados semelhantes.

Tabela 3.6: Estimativas bayesianas utilizando densidade *a priori* Normal *flat* para os modelos hipsométricos dados na Tabela (1.1).

Par.	Modelo	Est. Bayesiana	D.P.	ICred.(95%)	CCG(Z)
A	Curtis	$\beta_0 = 3,5273$	1,1502	(1,0543; 6,0717)	0,0386
		$\beta_1 = 6,2813$	13,7493	(-23,1967; 37,7363)	0,0461
		$S_2 = 2,5161$	0,0362	(0,0121; 15,8198)	0,0142
	Stofel	$\beta_0 = 1,6049$	4,4024	(-7,9695; 11,3951)	0,0093
		$\beta_1 = 0,5565$	1,7086	(-3,2314; 4,2695)	0,0058
		$S_2 = 2,0336$	0,0260	(0,0091; 11,8323)	0,0031
	Petterson	$\beta_0 = 0,3065$	0,0103	(0,2861; 0,3262)	0,0553
		$\beta_1 = 0,7223$	0,1353	(0,4660; 0,9961)	0,0516
		$S_2 = 2,7 * 10^{-5}$	$1,0 * 10^{-4}$	$(1,2 * 10^{-5}; 5,9 * 10^{-5})$	0,0194
	Prodan	$\beta_0 = 4,9241$	6,1353	(-6,0832; 16,5078)	0,0046
		$\beta_1 = -0,5955$	1,0866	(-2,6316; 1,3635)	0,0012
		$\beta_2 = 0,0626$	0,0462	(-0,0213; 0,1491)	0,0016
		$S_2 = 0,030$	0,0276	(0,0045; 0,1041)	0,0092
B	Curtis	$\beta_0 = 2,8401$	0,4989	(1,7689; 3,9865)	0,0151
		$\beta_1 = -3,9396$	0,5876	(-5,1243; -2,7981)	0,0268
		$S_2 = 2,8929$	0,0329	(0,0133; 13,2813)	0,0024
	Stofel	$\beta_0 = 3,7501$	2,7704	(-1,6449; 9,2622)	0,0047
		$\beta_1 = -0,2429$	1,0112	(-2,2818; 1,7422)	0,0031
		$S_2 = 2,1638$	0,0494	(0,0199; 10,3059)	0,0154
	Petterson	$\beta_0 = 0,3904$	0,0015	(0,3875; 0,3932)	0,0373
		$\beta_1 = -0,5119$	0,0232	(-0,5577; -0,4657)	0,0381
		$S_2 = 3,5 * 10^{-5}$	$1,75 * 10^{-5}$	$(1,5 * 10^{-5}; 7,7 * 10^{-5})$	0,0043
	Prodan	$\beta_0 = -7,2753$	5,5851	(-17,3299; 3,0019)	0,0057
		$\beta_1 = 0,7214$	0,7032	(-0,5704; 1,9843)	0,0055
		$\beta_2 = 0,0294$	0,0222	(-0,0103; 0,0703)	0,0053
		$S_2 = 0,0196$	0,0098	(0,0086; 0,0447)	0,0558
C	Curtis	$\beta_0 = 3,3745$	0,7532	(1,6963; 5,0438)	0,0066
		$\beta_1 = -1,1366$	5,3217	(-11,4673; 9,2251)	0,0329
		$S_2 = 6,6948$	0,0098	(0,0039; 40,1661)	0,0476
	Stofel	$\beta_0 = 3,5716$	3,1602	(-2,9316; 10,0419)	0,0211
		$\beta_1 = -0,0469$	1,0155	(-2,1732; 2,0306)	0,0263
		$S_2 = 3,8352$	0,0161	(0,0073; 28,4456)	0,0084
	Petterson	$\beta_0 = 0,3241$	0,0053	(0,3134; 0,3345)	0,0213
		$\beta_1 = -0,1208$	0,1232	(-0,3648; 0,1217)	0,0218
		$S_2 = 1,1 * 10^{-5}$	$4,6 * 10^{-6}$	$(5 * 10^{-6}; 2,2 * 10^{-5})$	0,0119
	Prodan	$\beta_0 = 20,0009$	4,7810	(11,0942; 28,7263)	0,0172
		$\beta_1 = -1,7951$	0,4207	(-2,5604; -1,0159)	0,0159
		$\beta_2 = 0,0724$	0,0093	(0,0552; 0,0893)	0,0148
		$S_2 = 0,0484$	0,0207	(0,0232; 0,1021)	0,0564

Par. é a parcela utilizada, Est. bayesiana é a estimativa bayesiana obtida via MCMC, D.P. é o desvio padrão, ICred. o intervalo de credibilidade e CCG(Z) é o critério de convergência de Geweke (1992).

Tabela 3.7: Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade *a priori* normal flat para os modelos hipsométricos dados na Tabela (1.1).

Parcela	Modelos	AIC	BIC	Log(CPO)	DIC
A	Curtis	678,51	688,35	80,02	1388,75
	Stofel	130,39	140,23	85,82	284,18
	Petterson	754,33	764,16	63,52	795,24
	Prodan	-30,61	-17,5	101,91	-37,33
B	Curtis	322,94	331,85	70,82	665,94
	Stofel	193,73	202,64	71,08	406,39
	Petterson	1214,5	1223,41	-8,74	1209,95
	Prodan	-43,38	-31,5	98,28	-50,36
C	Curtis	2942,38	2952,63	81,76	5931,13
	Stofel	1043,3	1053,555	83,71	2126,54
	Petterson	623,78	634,03	79,08	663,51
	Prodan	-45,17	-31,51	95,15	-52,15

AIC é o critério de informação de Akaike (1973), BIC o critério de informação bayesiano (SCHWARZ, 1978), DIC é o critério desvio-informação (SPIEGELHALTER et al., 2002) e Log(CPO) é o logaritmo do critério da densidade preditiva ordenada (GELFAND; DEY, 1994).

Um dos principais objetivos da proposta de uma abordagem bayesiana em modelos hipsométricos é resolver o problema de estimação e inferência com restrições nos parâmetros. Até o presente momento, nenhuma das abordagens bayesianas, seja com a densidade *a priori* de Jeffreys ou com densidade *a priori* normal *flat*, foram capazes de solucionar. Seus resultados em muitos casos (vide por exemplo as parcelas B e C nas Tabelas (3.4) e (3.6)) proporcionam resultados biologicamente inconsistentes. Neste sentido, nosso próximo passo será utilizar uma densidade *a priori* informativa, buscando incorporar as informações *a priori* disponíveis.

A primeira densidade de probabilidade *a priori* informativa proposta será uma densidade *a priori* construída empiricamente, a qual considera as restrições nos parâmetros impostas biologicamente. Sua descrição e resultados são expostos na próxima seção.

(A1)

(A2)

(A3)

(A4)

Figura 3.6: Ajuste bayesiano utilizando *a priori* Normal *flat* para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.

(B1)

(B2)

(B3)

(B4)

Figura 3.7: Ajuste bayesiano utilizando *a priori* Normal *flat* para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4.

(C1)

(C2)

(C3)

(C4)

Figura 3.8: Ajuste bayesiano utilizando *a priori* Normal *flat* para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4.

Figura 3.9: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição *a priori* Normal *flat*.

Figura 3.10: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela A, com distribuição *a priori* Normal *flat*.

3.1.3 Densidade de probabilidade *a priori* empírica

Até o presente momento, foi utilizada apenas densidades de probabilidade *a priori* não informativas ou vagas, desconsiderando o conhecimento prévio dos parâmetros dos modelos (Tabela (1.1)). Nessa seção, será considerada uma densidade de probabilidade *a priori* empírica a qual procura modelar as restrições impostas aos parâmetros dos modelos, por razões biológicas, e incluir o conhecimento anterior que se tem dos parâmetros envolvidos.

Para considerar as restrições do tipo $\beta \in R$, onde $R = \{a_i < \beta_i < b_i, i = 0, \dots, p\}$, será considerado densidades *a priori* conjugadas Normal-Gama Truncada para os parâmetros, considera-se $\beta|\tau \propto \text{NT}(\mu, (\tau P)^{-1})$ e $\tau \propto \text{G}(\nu_0, \lambda_0)$. Assim temos:

$$\pi_1(\beta|\tau) \propto \frac{\tau^{1/2}}{T(\mu, P)} \exp \left\{ -\frac{\tau}{2} [(\beta - \mu)' P (\beta - \mu)] \right\} I_R(\beta) \quad (3.20)$$

Em ((3.20)) $\mu = (\mu_0, \mu_1)'$ (para os modelos de Curtis, Stofel e Petterson) e $\mu = (\mu_0, \mu_1, \mu_2)'$ (para o modelo de Prodan) é o vetor de hiperparâmetros de locação e P é a matriz de precisão, ambos conhecidos e $T(\mu, P)$ é uma constante normalizadora que não precisa ser calculada quando se usa o algoritmo *Gibbs sampling*. $I_R(\beta)$ é uma função indicadora tal que $I_R(\beta) = 1$ se $\beta \in R$ e $I_R(\beta) = 0$ caso contrário. A densidade *a priori* para τ é dada por:

$$\pi_2(\tau) \propto \tau^{\nu_0-1} \exp\{-\lambda_0 \tau\}. \quad (3.21)$$

Em ((3.21)) ν_0 e λ_0 são hiperparâmetros também conhecidos. A *a priori* conjunta é dada por:

$$\pi_0(\beta, \tau) = \pi_1(\beta|\tau)\pi_2(\tau) \quad (3.22)$$

Neste trabalho vamos adotar uma abordagem empírica para a determinação dos hiperparâmetros μ e P . Para isso consideramos os percentis (H_p, Dap_p) e (H_q, Dap_q) , com $p < q$, encontrados em dados históricos coletados anteriormente às observações da amostra D , os quais não ocorreram obrigatoriamente de forma pareada (neste trabalho consideramos nas aplicações $p = 25\%$ e $q = 75\%$, mas outros percentis podem ser considerados sem perda de generalidade).

Assumindo que os modelos da Tabela (1.1) são válidos para esses percentis, podemos escrever os parâmetros da densidade *a priori* empírica como descrito na Tabela (3.8). Para maiores detalhes sobre os parâmetros *a priori* veja o apêndice B.

A matriz de precisão P também pode ser encontrada de forma empírica considerando

Tabela 3.8: Parâmetros da densidade *a priori* empírica

Modelo	Parâmetros da <i>priori</i>				
	μ_0^*	μ_1^*	μ_2^*	y_s	x_s
Stofel	$y_p + \left(\frac{y_q - y_p}{x_q - x_p}\right) x_p$	$-\left(\frac{y_q - y_p}{x_q - x_p}\right)$	—	$\ln(H_s)$	$\ln(Dap_s)$
Curtis	$y_p + \left(\frac{y_q - y_p}{x_q - x_p}\right) x_p$	$-\left(\frac{y_q - y_p}{x_q - x_p}\right)$	—	$\ln(H_s)$	$1/Dap_s$
Petterson	$\frac{\alpha_0 \alpha_1^2}{2}$	$\alpha_0 \alpha_1$	—	$\ln\left(\frac{1}{H_s^{1/3}}\right)$	$1/Dap_s$
Prodan	$\frac{\alpha_0 \alpha_1^2}{2}$	$\alpha_0 \alpha_1$	α_0	H_s	$1/Dap_s$

$$s = p, q; \alpha_0 = \exp\left\{y_p + \left(\frac{y_q - y_p}{x_q - x_p}\right) x_p\right\}; \alpha_1 = -\left(\frac{y_q - y_p}{x_q - x_p}\right).$$

um coeficiente de variação constante $CV = \gamma_j/\mu_j^*$, $j = 0, 1, 2$ que expresse a confiança nos hiperparâmetros $\boldsymbol{\mu}$. Desta forma temos:

$$P_{j,j} = \frac{1}{\gamma_j^2} = \frac{1}{(CV\mu_j^*)^2}, j = 0, 1, 2 \quad (3.23)$$

Vamos assumir também que os coeficientes β_i , $i = 0, 1, 2$ são independentes *a priori*, de forma que $P_{i,j} = 0$ para $i \neq j$, ($i, j = 0, 1, 2$).

Nesta modelagem um passo importante é estabelecer a relação entre a variável aleatória normal truncada e uma versão desta variável com distribuição normal. Esta relação é dada em (ALBERT; CHIB, 1996), considerando uma variável aleatória uniforme $u_i \sim U(0, 1)$, então podemos escrever que:

$$\beta_i = \mu_i + \varpi_i \Phi^{-1} \left[\Phi \left(\frac{a_i - \mu_i}{\varpi_i} \right) + u_i \left(\Phi \left(\frac{b_i - \mu_i}{\varpi_i} \right) - \Phi \left(\frac{a_i - \mu_i}{\varpi_i} \right) \right) \right] \quad (3.24)$$

para $i = 0, \dots, p$

Estamos assumindo independência *a priori* entre os parâmetros μ_i , portanto $\varpi_i = \sqrt{\sigma_{ii}}$ é a raiz quadrada do i -ésimo elemento da diagonal de $(\tau \mathbf{P})^{-1}$.

As equações (3.24) indicam como podemos gerar valores dos parâmetros restritos da distribuição truncada a partir da geração de variáveis aleatórias uniformes $u_i \propto U(0, 1)$.

3.1.3.1 A densidade de probabilidade *a posteriori* obtida através da *priori* empírica

A função densidade de probabilidade *a posteriori* de $\boldsymbol{\beta}$ e τ é dada por

$$\pi(\boldsymbol{\beta}, \tau | \mathbf{D}) \propto L(\boldsymbol{\beta}, \tau | \mathbf{D}) \pi_0(\boldsymbol{\beta}, \tau) \quad (3.25)$$

3.1.3.2 Modelos de Stofel e Curtis

Para os modelos de Stofel e Curtis (MOREIRA et al., 2007) (Tabela (1.1)), a função densidade de probabilidade *a posteriori* de β e τ é obtida substituindo-se (3.20) e (3.21) em (3.22), e (2.1) e (3.22) em (3.25). Assim,

$$\pi(\beta, \tau | \mathbf{D}) \propto \tau^{n/2 + \nu_0} \exp \left\{ -\frac{\tau}{2} \left[(\beta - \tilde{\beta})' \mathbf{V} (\beta - \tilde{\beta}) + \mathbf{W} + 2\lambda_0 \right] \right\} \mathbf{I}_R(\beta), \quad (3.26)$$

onde $\mathbf{W} = \mathbf{Y}'\mathbf{Y} + \tilde{\beta}'\mathbf{V}\tilde{\beta}$, $\mathbf{V} = (\mathbf{X}'\mathbf{X} + \mathbf{P})$ e $\tilde{\beta}$ é o estimador bayesiano irrestrito, dado por

$$\tilde{\beta} = (\mathbf{X}'\mathbf{X} + \mathbf{P})^{-1}(\mathbf{X}'\mathbf{X}\hat{\beta} + \mathbf{P}\mu) \quad (3.27)$$

sendo que $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ é o estimador de mínimos quadrados.

O próximo passo é elaborar um algoritmo amostrador de Gibbs para encontrar o estimador bayesiano de β e τ . Será adotada a seguinte notação:

$$\beta | \tau, \mathbf{D} \propto \text{NT} \left(\tilde{\beta}, (\tau \mathbf{V})^{-1} \right) \quad (3.28)$$

$$\tau | \beta, \mathbf{D} \propto \text{G}(n/2 + \nu_0 + 1, \Lambda(\beta) + 2\lambda_0) \quad (3.29)$$

onde $\Lambda(\beta) = (\beta - \tilde{\beta})' \mathbf{V} (\beta - \tilde{\beta}) + \mathbf{W}$.

O algoritmo amostrador de Gibbs utilizado para gerar amostras da densidade *a posteriori* conjunta $\pi(\beta, \tau | \mathbf{D})$ é dado a seguir:

- Algoritmo *Gibbs Sampling*:

1. Faça $j = 0$ e considere a condição inicial, $\tau^{(0)}$.
2. Gere $\beta^{(j+1)}$ de $\beta^{(j+1)} | \tau^{(j)}, \mathbf{D} \propto \text{NT} \left(\tilde{\beta}, (\tau^{(j)} \mathbf{V})^{-1} \right)$.
3. Gerar $\tau^{(j+1)}$ de $\tau^{(j+1)} | \beta^{(j)}, \mathbf{D} \propto \text{G}(n/2 + \nu_0 + 1, \Lambda(\beta^{(j)}) + 2\lambda_0)$.
4. Faça $j \leftarrow j + 1$ e repita os passos (2) e (3) até atingir a convergência.

3.1.3.3 Modelos de Prodan e Petterson

Para os modelos de Prodan e Petterson (Tabela (1.1)), a função densidade *a posteriori* é obtida substituindo na equação (3.22) as equações (3.20) e (3.21), e (3.22) e (2.1) na equação (3.25).

Em seguida será elaborado um algoritmo de simulação de Monte Carlo em Cadeia de Markov (MCMC) para encontrar o estimador Bayesiano de β e τ . As funções densidade condicionais usadas no algoritmo de MCMC são dadas por:

$$\beta|\tau, \mathbf{D} \propto L(\beta, \tau|\mathbf{D})\text{NT}(\beta; \beta_0, (\tau\mathbf{P})^{-1}) \quad (3.30)$$

$$\tau|\beta, \mathbf{D} \propto \Psi(\beta, \tau)\text{G}(\tau; \nu_0, \lambda_0) \quad (3.31)$$

onde $\Psi(\beta, \tau) = L(\beta, \tau|\mathbf{D})\text{NT}(\beta; \beta_0, (\tau\mathbf{P})^{-1})$. O algoritmo MCMC utilizado para gerar amostras da densidade *a posteriori* conjunta $\pi(\beta, \tau|\mathbf{D})$ é dado por:

- Algoritmo Metrópolis-Hasting:

1. Faça $j = 0$ e considere a condição inicial, $\tau^{(0)}$.
2. Gere $\beta^{(j+1)}$ de $\beta^{(j+1)}|\tau^{(j)} \propto \text{NT}(\beta; \beta_0, (\tau^{(j)}\mathbf{P})^{-1})$.
3. Aceite $\beta^{(j+1)}$ com probabilidade

$$\alpha(\beta^{(j)}, \beta^{(j+1)}) = \min\{1, L(\beta^{(j+1)}, \tau^{(j)}|\mathbf{D})/L(\beta^{(j)}, \tau^{(j)}|\mathbf{D})\}.$$

4. Gere $\tau^{(j+1)}$ de $\tau^{(j+1)} \propto \text{G}(\tau; \nu_0, \lambda_0)$.
5. Aceite $\tau^{(j+1)}$ com probabilidade

$$\alpha(\tau^{(j)}, \tau^{(j+1)}) = \min\{1, \Psi(\beta^{(j+1)}, \tau^{(j+1)}|\mathbf{D})/\Psi(\beta^{(j+1)}, \tau^{(j)}|\mathbf{D})\}.$$

6. Faça $j \leftarrow j + 1$ e repita os passos (2)-(6) até a convergência ser atingida.

Para encontrar as estimativas bayesianas obtidas via simulação de Monte Carlo será adotado o mesmo procedimento adotado na seção 3.1.1.1. Considera-se um período de aquecimento (burn-in) de 50% das iterações iniciais. Todos os passos do algoritmo serão repetidos até que haja evidências de convergência, monitorada graficamente e com o critério de Geweke (GEWEKE, 1992).

Os estimadores bayesianos obtidos via simulação de Monte Carlo são calculados pela equação (3.8) e os intervalos de credibilidade por (3.9).

3.1.3.4 Aplicação

Os dados utilizados para testar a abordagem bayesiana com distribuição *a priori* empírica foram coletadas em florestas plantadas de clones de *Eucalyptus sp.* descritos na

seção 2.3. Dentre as várias parcelas utilizadas, foram selecionadas três, as parcelas A, B e C (Tabela (2.1)) para exemplificar os resultados obtidos.

Os resultados das estimativas bayesianas são apresentadas na Tabela (3.9), assim como o critério de convergência de Geweke (1992) utilizado para verificar a convergência da cadeia gerada pelo algoritmo MCMC, lembrando que como critério de Geweke a convergência ocorre quando $|Z| < Z_c$ para $Z_c = 1,96$. Podemos verificar nessa tabela que os resultados obtidos utilizando a densidade de probabilidade *a priori* empírica são sempre coerentes biologicamente. Todos os resultados respeitam as restrições impostas biologicamente, $\beta_i > 0, i = 0, 1, 2$, ao contrário do que ocorre com as densidades *a priori* de Jeffreys e normal *flat* que em muitos casos produz resultados inconsistentes. Esse resultado também pode ser visto nas Figuras (3.11), (3.12) e (3.13) onde os parâmetros estimados positivos resultam em ajustes crescentes.

Para a seleção de modelos, foram utilizados o critério de informação de Akaike (AIC), o critério de informação bayesiano (BIC), o critério desvio-informação (DIC) e o critério da densidade preditiva ordenada (maiores detalhes sobre seleção de modelo no capítulo 4). Os resultados são apresentados na Tabela (3.10). Para a parcela A, o melhor modelo selecionado por todos os critérios de seleção utilizados foi o modelo de Curtis e para a parcela B o modelo de Stofel. No caso da parcela C, os critérios DIC e CPO selecionaram o modelo de Petterson enquanto que os critérios AIC e BIC ficaram com o modelo de Prodan. Embora tenha havido essa pequena divergência de escolha do melhor modelo pelos critérios adotados, a mesma não pode ser considerada significativa quando olhamos para os valores obtidos por cada critério (Tabela (3.10)). O AIC para o modelo de Petterson foi $-49,53$ e para o modelo de Prodan $-49,90$. O DIC para o modelo de Petterson foi $-55,83$ e para o modelo de Prodan $-54,76$. Para os outros dois critérios, BIC e Log(CPO), a situação foi a mesma, seus valores são próximos o que nos permite inferir que esses dois modelos são igualmente bons. A fim de optar por apenas um modelo, ficaremos com o modelo de Petterson pois foi o selecionado pelo critério da densidade preditiva ordenada além de ser mais parcimonioso que o modelo de Prodan por possuir um parâmetro a menos.

Como exemplo, apresentamos o histograma e trajetória das amostras geradas com o algoritmo MCMC após verificada a convergência no modelo de Curtis, para os parâmetros β e τ nas Figuras (3.14) e (3.15), respectivamente. O resultado para as demais parcelas e modelos apresentam resultados semelhantes e são apresentados no apêndice C.1.

Como exemplo, apresentamos o histograma e trajetória das amostras geradas com o algoritmo de Metrópolis-Hastings após verificada a convergência, para os parâmetros β e

Tabela 3.9: Estimativas bayesianas utilizando densidade *a priori* empírica para os modelos hipsométricos dados na Tabela (1.1).

Par.	Modelo	Est. Bayesiana	D.P.	ICred.(95%)	CCG(Z)
A	Curtis	$\beta_0 = 3,5222$	0,0096	(3,5034; 3,5402)	0,0405
		$\beta_1 = 6,1502$	0,1007	(5,9537; 6,3528)	0,0181
		$S_2 = 0,00034$	$8,9 * 10^{-6}$	(0,00032; 0,00036)	0,0074
	Stofel	$\beta_0 = 1,6690$	0,0237	(1,6226; 1,7145)	0,0599
		$\beta_1 = 0,5330$	0,0093	(0,5144; 0,5506)	0,0756
		$S_2 = 5,6 * 10^{-4}$	$2,2 * 10^{-5}$	$(5,2 * 10^{-4}; 6,1 * 10^{-4})$	0,0449
	Petterson	$\beta_0 = 0,3066$	0,0021	(0,3023; 0,3108)	0,0101
		$\beta_1 = 0,7211$	0,0279	(0,6668; 0,7772)	0,0096
		$S_2 = 0,3748$	0,1349	(0,2019; 0,7014)	0,0214
	Prodan	$\beta_0 = 0,4718$	0,0377	(0,3979; 0,5443)	0,0187
		$\beta_1 = 0,1367$	0,0075	(0,1222; 0,1513)	0,0165
		$\beta_2 = 0,0340$	0,0006	(0,0329; 0,0352)	0,0099
		$S_2 = 0,5849$	0,2028	(0,3157; 1,1062)	0,0233
B	Curtis	$\beta_0 = 3,5826$	0,0178	(3,5478; 3,6171)	0,1119
		$\beta_1 = 8,0469$	0,2474	(7,5624; 8,5129)	0,0339
		$S_2 = 9,3 * 10^{-4}$	$4,03 * 10^{-5}$	$(8,5 * 10^{-4}; 1,0 * 10^{-3})$	0,0574
	Stofel	$\beta_0 = 1,6592$	0,0341	(1,5885; 1,7267)	0,0194
		$\beta_1 = 0,5112$	0,0121	(0,4878; 0,5347)	0,0189
		$S_2 = 9,5 * 10^{-4}$	$5,4 * 10^{-5}$	$(8,6 * 10^{-4}; 1,07 * 10^{-3})$	0,0484
	Petterson	$\beta_0 = 0,3505$	0,0033	(0,3439; 0,3569)	0,0473
		$\beta_1 = 0,1224$	0,0526	(0,0239; 0,2251)	0,0395
		$S_2 = 1,1814$	0,3898	(0,6299; 2,1674)	0,0365
	Prodan	$\beta_0 = 1,7265$	0,1682	(1,3974; 2,0585)	0,0543
		$\beta_1 = 0,2106$	0,0162	(0,1781; 0,2422)	0,0076
		$\beta_2 = 0,0262$	0,0011	(0,0240; 0,0284)	0,0399
		$S_2 = 1,8599$	0,6322	(1,0117; 3,4646)	0,0224
C	Curtis	$\beta_0 = 3,8089$	0,0052	(3,7985; 3,8187)	0,0755
		$\beta_1 = 8,8607$	0,1006	(8,6689; 9,0637)	0,0195
		$S_2 = 1,27 * 10^{-4}$	$2,07 * 10^{-6}$	$(1,24 * 10^{-4}; 1,32 * 10^{-4})$	0,0466
	Stofel	$\beta_0 = 2,2368$	0,0119	(2,2136; 2,2598)	0,0343
		$\beta_1 = 0,3781$	0,0039	(0,3705; 0,3857)	0,0232
		$S_2 = 1,28 * 10^{-4}$	$2,09 * 10^{-6}$	$(1,24 * 10^{-4}; 1,32 * 10^{-4})$	0,0106
	Petterson	$\beta_0 = 0,2923$	0,0011	(0,2901; 0,2945)	0,0040
		$\beta_1 = 0,6186$	0,0268	(0,5668; 0,6711)	0,0075
		$S_2 = 0,6549$	0,2096	(0,3609; 1,1634)	0,0007
	Prodan	$\beta_0 = 1,5971$	0,1601	(1,2735; 1,9061)	0,0210
		$\beta_1 = 0,1727$	0,0141	(0,1459; 0,2002)	0,0326
		$\beta_2 = 0,0221$	0,0006	(0,0209; 0,0233)	0,0432
		$S_2 = 1,1261$	0,3644	(0,6236; 2,0411)	0,0585

Par. é a parcela utilizada, Est. bayesiana é a estimativa bayesiana obtida via MCMC, D.P. é o desvio padrão, ICred. o intervalo de credibilidade e CCG(Z) é o critério de convergência de Geweke (1992).

Tabela 3.10: Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade *a priori* empírica para os modelos hipsométricos dados na Tabela (1.1).

Parcela	Modelos	AIC	BIC	Log(CPO)	DIC
A	Curtis	-54,73	-44,89	121,72	-54,55
	Stofel	-27,56	-17,72	121,45	-6,21
	Petterson	-14,94	-5,1	92,83	-20,26
	Prodan	-6,49	6,62	89,58	-13,81
B	Curtis	47,81	58,06	90,05	47,29
	Stofel	72,19	82,43	120,93	99,56
	Petterson	-15,57	-5,32	94,19	-20,86
	Prodan	2,33	15,99	90,06	-5,01
C	Curtis	17,43	27,68	85,89	73,42
	Stofel	11,40	21,64	72,77	13,71
	Petterson	-49,53	-35,86	102,57	-55,83
	Prodan	-49,90	-39,65	101,56	-54,76

AIC é o critério de informação de Akaike (1973), BIC o critério de informação bayesiano (SCHWARZ, 1978), DIC é o critério desvio-informação (SPIEGELHALTER et al., 2002) e Log(CPO) é o logaritmo do critério da densidade preditiva ordenada (GELFAND; DEY, 1994).

τ nas Figuras (3.4) e (3.5), respectivamente. Essa convergência pode ser observada nos gráficos *A1* e *B1* da Figura (3.4) e *A1* da Figura (3.5) nos quais se verifica a tendência estacionária das séries simuladas dos parâmetros em torno das estimativas. Os histogramas *A2* e *B2* da Figura (3.4) e *A2* da Figura (3.5) nos mostra que as densidades *a posteriori* são normalmente distribuídas entre os valores estimados, tratando-se de uma distribuição unimodal. O resultado para as demais parcelas e modelos apresentam resultados semelhantes.

Comparando as estimativas bayesianas com densidade *a priori* empírica (EBPE) com as estimativas bayesianas com *priori* de Jeffreys (EBPJ), normal *flat* (EBPN) e as estimativas de máxima verossimilhança (EMV), podemos verificar que EBPE resultam sempre em valores positivos (Tabela (3.9)), ao contrário do que ocorre nos outros três casos (Tabelas (2.2), (3.4) e (3.6), respectivamente EMV, EBPJ e EBPN). Os intervalos de credibilidade dados na Tabela (3.9) mostram a precisão dessas estimativas. Além disso, o resultado das EBPE coincide com os resultados das EMV quando estas são factíveis. Isso pode ser verificado na Figura (3.11) para os modelos de Curtis, Stofel e Petterson, respectivamente gráficos (A1), (A2) e (A3).

Por tudo isso, as estimativas bayesianas com densidade *a priori* empírica mostraram-se superiores as demais técnicas bayesianas e clássica utilizadas até aqui. No entanto, a quantidade de informação *a priori* disponível sobre relação hipsométrica em clones de

Eucalyptus sp. é muitas vezes superior a quantidade de informação *a priori* utilizada na construção da *priori* empírica. Desse modo, com o intuito de melhorar ainda mais as estimativas bayesianas encontradas, propomos uma densidade *a priori* potência buscando agregar mais informação histórica nas análises *a posteriori* e com isso melhorar as estimativas encontradas.

(A1)

(A2)

(A3)

(A4)

Figura 3.11: Ajuste bayesiano utilizando *a priori* empírica para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.

(B1)

(B2)

(B3)

(B4)

Figura 3.12: Ajuste bayesiano utilizando *a priori* empírica para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4.

(C1)

(C2)

(C3)

(C4)

Figura 3.13: Ajuste bayesiano utilizando *a priori* empírica para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4.

Figura 3.14: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição *a priori* empírica.

Figura 3.15: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela A, com distribuição *a priori* empírica.

3.1.4 Densidade de probabilidade *a priori* potência

A abordagem utilizando a densidade de probabilidade *a priori* potência provem da classe de *prioris* informativas na inferência bayesiana. A idéia básica é usar um parâmetro de potência a ($0 \leq a \leq 1$) para controlar a influência dos dados históricos no estudo presente. A idéia inicial do uso deste parâmetro potência pode ser encontrado em Diaconis e Ylvisaker (1979) e Morris (1983) os quais consideraram esse parâmetro apenas como uma constante. Ibrahim e Chen desenvolveram essa idéia em estudos teóricos em Ibrahim e Chen (2000), Chen, Ibrahim e Shao (2000), Ibrahim, Chen e Sinha (2003) e Chen, Ibrahim e Shao (2003), utilizando o parâmetro potência como uma variável aleatória, assim como será considerado nesse trabalho.

Seja os dados para o estudo corrente denotados por $\mathbf{D} = (n, \mathbf{y}, \mathbf{X})$, onde n é o tamanho da amostra, $\mathbf{y}$ é um vetor ($n \times 1$) da variável resposta e $\mathbf{X}$ é a matrix ($n \times p$) de covariáveis explanatórias. Além disso, será denotado por $L(\boldsymbol{\theta}; \mathbf{D})$ a função de verossimilhança dos dados correntes, onde $\boldsymbol{\theta}$ é o vetor de parâmetros do modelo. Agora, supõe que tem-se dados históricos similares aos que estão sendo utilizados no presente estudo, estes serão denotados por $\mathbf{D}_0 = (n_0, \mathbf{y}_0, X_0)$. Será simbolizado como $\pi(\theta)$ a distribuição *a priori* para θ antes dos dados históricos serem observados. Então a densidade de probabilidade *a priori* potência de (θ, a) para os dados correntes será (IBRAHIM; CHEN, 2000):

$$\pi(\boldsymbol{\theta}, a | \mathbf{D}_0) \propto [L(\boldsymbol{\theta} | \mathbf{D}_0)]^a \pi_0(\boldsymbol{\theta}) \pi(a). \quad (3.32)$$

O parâmetro a mede o poder da informação histórica disponível no estudo corrente. O caso em que $a = 0$ significa que nenhuma porção dos dados históricos está sendo utilizada no presente estudo; enquanto $a = 1$ dá igual peso a $L(\boldsymbol{\theta} | \mathbf{D}_0)$ e a função de verossimilhança do estudo corrente $L(\boldsymbol{\theta} | \mathbf{D})$, resultando na incorporação de todo o conhecimento dos dados históricos no estudo presente. Então a equação (3.32) pode ser vista como uma generalização da usual atualização bayesiana de $\pi(\boldsymbol{\theta})$ (para maiores detalhes ver Ibrahim e Chen (2000)). O parâmetro potência a pode ser interpretado como parâmetro de precisão. (DUAN; YE; SMITH, 2006) Por exemplo, considere o caso de uma amostra normal com variância conhecida. Suponha que $\mathbf{D}_0$ consiste em n_0 observações, $X_1, \dots, X_{n_0}$, de uma população com distribuição normal com parâmetro $\boldsymbol{\theta}$ e variância σ^2 conhecida. Se a distribuição *a priori* é assumida uniforme (*priori* não informativa), $\pi(\boldsymbol{\theta} | \mathbf{D}_0, a) \propto L(\boldsymbol{\theta} | \mathbf{D}_0)^a \pi(\boldsymbol{\theta})$ implica numa distribuição *a priori* de $\boldsymbol{\theta}$ para os dados correntes $\mathbf{D}$, $\boldsymbol{\theta} | \mathbf{D}_0, a \sim N(\bar{x}_0, \sigma^2 / a n_0)$, onde $\bar{x}_0$ é a média amostral dos dados históricos. Portanto, a pode ser visto como parte do parâmetro de precisão, pois quanto menor a

maior será a variância da *priori* potência, enquanto a grande implica em menor variância da *priori* potência.

Uma escolha natural para $\pi(a)$ é uma distribuição de probabilidade beta. No entanto, outras distribuições poderiam ser usadas como a distribuição gama truncada ou uma normal truncada. De acordo com Ibrahim e Chen (2000), essas três distribuições tem propriedades teóricas e computacionais similares. Na prática dão resultados semelhantes desde que os hiperparâmetros sejam apropriadamente escolhidos. Dessa forma, neste trabalho a escolha será utilizar uma distribuição beta para $\pi(a)$, $a \sim \text{Beta}(\gamma_0, \lambda_0)$, onde (γ_0, λ_0) são hiperparâmetros conhecidos.

A priori potência definida em (3.32) pode ser facilmente generalizada para vários conjuntos de dados históricos. Se existem L_0 conjuntos de dados históricos, será definido como $D_{0s} = (n_{0s}, \mathbf{y}_{0s}, X_{0s})$ o s -ésimo conjunto de dados históricos, sendo $s = 1, \dots, L_0$ e $D_0 = (D_{01}, \dots, D_{0L_0})$. Neste caso, é desejável definir o peso para o parâmetro a_{0s} para cada conjunto de dados históricos, e utilizar densidades de probabilidade *a priori* i.i.d.. Sendo $a = (a_1, \dots, a_{L_0})$, a distribuição *a priori* generalizada será

$$\pi(\boldsymbol{\theta}, a | \mathbf{D}_0) \propto \left(\prod_{s=1}^{L_0} L(\boldsymbol{\theta} | \mathbf{D}_{0s})^{a_s} \pi(a_s) \right) \pi_0(\boldsymbol{\theta}). \quad (3.33)$$

3.1.4.1 A densidade de probabilidade *a posteriori* obtida através da *priori* potência

A função densidade de probabilidade *a posteriori* de $\boldsymbol{\beta}$, τ e a é dada por

$$\pi(\boldsymbol{\beta}, \tau, a | \mathbf{D}, \mathbf{D}_0) \propto L(\boldsymbol{\beta}, \tau | \mathbf{D}) \pi(\boldsymbol{\beta}, \tau, a | \mathbf{D}_0), \quad (3.34)$$

onde $L(\boldsymbol{\beta}, \tau | \mathbf{D})$ é função de verossimilhança associada ao conjunto de observações $\mathbf{D}$ e $\pi(\boldsymbol{\beta}, \tau, a | \mathbf{D}_0)$ dado pela equação (3.32).

3.1.4.2 Modelos de Stofel e Curtis

Para encontrar a função densidade de probabilidade *a posteriori* de $\boldsymbol{\beta}$, τ e a para os modelos de Stofel e Curtis será substituída a equação (3.32) na equação (3.34). Assim,

$$\pi(\boldsymbol{\beta}, \tau, a | \mathbf{D}, \mathbf{D}_0) \propto L(\boldsymbol{\beta}, \tau | \mathbf{D}) L(\boldsymbol{\beta}, \tau | \mathbf{D}_0)^a \pi(\boldsymbol{\beta}, \tau) \pi(a | \gamma_0, \lambda_0) \quad (3.35)$$

A densidade *a priori* inicial de $\boldsymbol{\beta}$ e τ será considerada constante, ou seja, $\pi(\boldsymbol{\beta}, \tau) \propto 1/\tau$, para a uma densidade de probabilidade beta, $\pi(a) \propto a^{\gamma_0-1} (1-a)^{\lambda_0-1}$ como mencionado

anteriormente, $L(\boldsymbol{\beta}, \tau | \mathbf{D}) \propto \tau^{\frac{n}{2}} \exp \left\{ -\frac{\tau}{2} \left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' X' X (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + S^2 \right] \right\}$ sendo $S^2 = (\mathbf{Y} - X\hat{\boldsymbol{\beta}})'(\mathbf{Y} - X\hat{\boldsymbol{\beta}})$ e $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\mathbf{Y}$, $L(\boldsymbol{\beta}, \tau | \mathbf{D}_0) \propto \tau^{\frac{n_0}{2}} \exp \left\{ -\frac{\tau}{2} \left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_0)' X'_0 X_0 (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_0) + S_0^2 \right] \right\}$ sendo $S_0^2 = (\mathbf{Y}_0 - X_0\hat{\boldsymbol{\beta}}_0)'(\mathbf{Y}_0 - X_0\hat{\boldsymbol{\beta}}_0)$ e $\hat{\boldsymbol{\beta}}_0 = (X'_0 X_0)^{-1}X'_0 \mathbf{Y}_0$, logo

$$\begin{aligned}
\pi(\boldsymbol{\beta}, \tau, a | \mathbf{D}, \mathbf{D}_0) &\propto \tau^{\frac{n}{2}} \exp \left\{ -\frac{\tau}{2} \left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' X' X (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + S^2 \right] \right\} \\
&\times \left\{ \tau^{\frac{n_0}{2}} \exp \left\{ -\frac{\tau}{2} \left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_0)' X'_0 X_0 (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_0) + S_0^2 \right] \right\} \right\}^a \\
&\times \frac{1}{\tau} a^{\gamma_0-1} (1-a)^{\lambda_0-1} \\
&\propto \tau^{\frac{n+an_0}{2}-1} \exp \left\{ -\frac{\tau}{2} \left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' X' X (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + S^2 \right. \right. \\
&\quad \left. \left. + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_0)' a X'_0 X_0 (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_0) + a S_0^2 \right] \right\} a^{\gamma_0-1} (1-a)^{\lambda_0-1} \\
&\propto \tau^{\frac{n+an_0}{2}-1} \exp \left\{ -\frac{\tau}{2} \left(\boldsymbol{\beta}' X' X \boldsymbol{\beta} - \boldsymbol{\beta}' X' X \hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}' X' X \boldsymbol{\beta} + \hat{\boldsymbol{\beta}}' X' X \hat{\boldsymbol{\beta}} \right. \right. \\
&\quad \left. \left. + \boldsymbol{\beta}' a X'_0 X_0 \boldsymbol{\beta} - \boldsymbol{\beta}' a X'_0 X_0 \hat{\boldsymbol{\beta}}_0 - \hat{\boldsymbol{\beta}}_0' a X'_0 X_0 \boldsymbol{\beta} + \hat{\boldsymbol{\beta}}_0' a X'_0 X_0 \hat{\boldsymbol{\beta}}_0 + S^2 \right. \right. \\
&\quad \left. \left. + a S_0^2 \right) \right\} a^{\gamma_0-1} (1-a)^{\lambda_0-1} \\
&\propto \tau^{\frac{n+an_0}{2}-1} \exp \left\{ \boldsymbol{\beta}' (X' X + a X'_0 X_0) \boldsymbol{\beta} - \boldsymbol{\beta}' (X' X \hat{\boldsymbol{\beta}} + a X'_0 X_0 \hat{\boldsymbol{\beta}}_0) \right. \\
&\quad \left. - (\hat{\boldsymbol{\beta}}' X' X + \hat{\boldsymbol{\beta}}_0' a X'_0 X_0) \boldsymbol{\beta} + \hat{\boldsymbol{\beta}}' X' X \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}_0' a X'_0 X_0 \hat{\boldsymbol{\beta}}_0 + S^2 \right. \\
&\quad \left. + a S_0^2 \right\} a^{\gamma_0-1} (1-a)^{\lambda_0-1}
\end{aligned}$$

Finalmente,

$$\pi(\boldsymbol{\beta}, \tau, a | \mathbf{D}, \mathbf{D}_0) \propto \tau^{\frac{n+an_0}{2}-1} \exp \left\{ -\frac{\tau}{2} \left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_1(a))' \Sigma(a) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_1(a)) + S_1^2(a) \right] \right\} a^{\gamma_0-1} (1-a)^{\lambda_0-1}, \quad (3.36)$$

onde $\Sigma(a) = X'X + aX'_0X_0$, $\hat{\boldsymbol{\beta}}_1(a) = \Sigma(a)^{-1}(X'X\hat{\boldsymbol{\beta}} + aX'_0X_0\hat{\boldsymbol{\beta}}_0)$ e $S_1^2(a) = \hat{\boldsymbol{\beta}}'X'X\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}_0'aX'_0X_0\hat{\boldsymbol{\beta}}_0 + S^2 + aS_0^2 - \hat{\boldsymbol{\beta}}_1'(a)\Sigma(a)\hat{\boldsymbol{\beta}}_1(a)$.

Tendo como base a distribuição *a posteriori* (3.36), obtemos as distribuições condicionais necessárias na elaboração do algoritmo de Monte Carlo em Cadeia de Markov para encontrar os estimadores bayesianos de $\boldsymbol{\beta}$, τ e a ,

$$\boldsymbol{\beta} | \tau, a, \mathbf{D}, \mathbf{D}_0 \sim N(\hat{\boldsymbol{\beta}}_1(a), (\tau \Sigma(a))^{-1}), \quad (3.37)$$

$$a | \tau, \boldsymbol{\beta}, \mathbf{D}, \mathbf{D}_0 \propto \tau^{\frac{n_1}{2}} \exp \left\{ -\frac{\tau}{2} \psi(\boldsymbol{\beta}, a) \right\} a^{\gamma_0-1} (1-a)^{\lambda_0-1}, \quad (3.38)$$

$$\tau | \boldsymbol{\beta}, a, \mathbf{D}, \mathbf{D}_0 \sim G\left(\frac{n_1}{2}, \psi(\boldsymbol{\beta}, a)\right), \quad (3.39)$$

onde, $n_1 = n + an_0$ e $\psi(\boldsymbol{\beta}, a) = (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_1(a))' \Sigma(a) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_1(a)) + S_1^2(a)$.

O algoritmo utilizado para calcular os estimadores bayesianos dos parâmetros dos modelos de Stofel e Curtis é dado por:

- Algoritmo Metrópolis-Hasting:

1. Faça $j = 0$ e considere a condição inicial $\tau^{(0)}$, $\beta^{(0)}$ e $a^{(0)}$.
2. Gere $a^{(j+1)}$ de $Beta(\gamma_0, \lambda_0)$, onde (γ_0, λ_0) são hiperparâmetros conhecidos.
3. Calcula-se a probabilidade de aceitação de $a^{(j+1)}$, por $\alpha = \min\{1, p\}$, em que

$$p = (\tau^{(j)})^{\frac{n_0(a^{(j+1)} - a^{(j)})}{2}} \exp \left\{ -\frac{\tau^{(j)}}{2} \left[\psi(\beta^{(j)}, a^{(j+1)}) - \psi(\beta^{(j)}, a^{(j)}) \right] \right\}$$

4. Gere u de uma distribuição uniforme $U(0, 1)$ e use a seguinte regra de aceitação:

$$a^{(j+1)} = \begin{cases} a^{(j+1)} & \text{se } u \leq \alpha \\ a^{(j)} & \text{se } u > \alpha \end{cases} \quad (3.40)$$

5. Gere $\beta^{(j+1)}$ de $\beta^{(j+1)} | \tau^{(j)}, a^{(j+1)}, \mathbf{D}, \mathbf{D}_0 \sim N(\hat{\beta}_1(a^{(j+1)}), (\tau^{(j)} \Sigma(a^{(j+1)}))^{-1})$,
6. Gere $\tau^{(j+1)}$ de $\tau^{(j+1)} | \beta^{(j+1)}, a^{(j+1)}, \mathbf{D}, \mathbf{D}_0 \sim G\left(\frac{n_1}{2}, \psi(\beta^{(j+1)}, a^{(j+1)})\right)$
7. Faça $j \leftarrow j + 1$ e repita os passos (2)-(6) até a convergência ser atingida.

3.1.4.3 Modelos de Petterson e Prodan

Assim como nos modelos de Curtis e Stofel, para os modelos de Petterson e Prodan, a função densidade de probabilidade *a posteriori* de β , τ e a será dada por:

$$\pi(\beta, \tau, a | \mathbf{D}, \mathbf{D}_0) \propto L(\beta, \tau | \mathbf{D}) L(\beta, \tau | \mathbf{D}_0)^a \pi(\beta, \tau) \pi(a | \gamma_0, \lambda_0) \quad (3.41)$$

onde,

$$L(\beta, \tau | \mathbf{D}) \propto \tau^{\frac{n}{2}} \exp \left\{ -\frac{\tau}{2} (\mathbf{Y} - \eta(x; \beta))' (\mathbf{Y} - \eta(x; \beta)) \right\},$$

$$L(\beta, \tau | \mathbf{D}_0) \propto \tau^{\frac{n_0}{2}} \exp \left\{ -\frac{\tau}{2} (\mathbf{Y}_0 - \eta(x_0; \beta))' (\mathbf{Y}_0 - \eta(x_0; \beta)) \right\},$$

$$\pi(a) \propto a^{\gamma_0 - 1} (1 - a)^{\lambda_0 - 1}$$

e

$$\pi(\beta, \tau) \propto 1/\tau.$$

Então, por intermédio da equação (3.41) obtem-se as distribuições condicionais ne-

cessárias na elaboração do algoritmo MCMC, as quais são dadas por:

$$\pi(\boldsymbol{\beta}|\tau, a, \mathbf{D}, \mathbf{D}_0) \propto L(\boldsymbol{\beta}, \tau|\mathbf{D})L(\boldsymbol{\beta}, \tau|\mathbf{D}_0)^a \quad (3.42)$$

$$\pi(a|\tau, \boldsymbol{\beta}, \mathbf{D}, \mathbf{D}_0) \propto L(\boldsymbol{\beta}, \tau|\mathbf{D})L(\boldsymbol{\beta}, \tau|\mathbf{D}_0)^a \pi(a) \quad (3.43)$$

$$\tau|\boldsymbol{\beta}, a, \mathbf{D}, \mathbf{D}_0 \sim G\left(\frac{n_1}{2}, \vartheta(\boldsymbol{\beta}, a)\right) \quad (3.44)$$

onde, $\vartheta(\boldsymbol{\beta}, a) = \frac{1}{2} [(\mathbf{Y} - \eta(x; \boldsymbol{\beta}))'(\mathbf{Y} - \eta(x; \boldsymbol{\beta})) + a(\mathbf{Y}_0 - \eta(x_0; \boldsymbol{\beta}))'(\mathbf{Y}_0 - \eta(x_0; \boldsymbol{\beta}))]$.

O algoritmo utilizado para obtenção dos estimadores bayesianos é constituído dos seguintes passos,

- Algoritmo Metrópolis-Hasting:

1. Faça $j = 0$ e considere a condição inicial $\tau^{(0)}$, $\boldsymbol{\beta}^{(0)}$ e $a^{(0)}$.
2. Gere $a^{(j+1)}$ de $Beta(\gamma_0, \lambda_0)$, onde (γ_0, λ_0) são hiperparâmetros conhecidos.
3. Calcula-se a probabilidade de aceitação de $a^{(j+1)}$, por $\alpha = \min\{1, p\}$, em que

$$p = \frac{L(\boldsymbol{\beta}^{(j)}, \tau^{(j)}|\mathbf{D})L(\boldsymbol{\beta}^{(j)}, \tau^{(j)}|\mathbf{D}_0)^{a^{(j+1)}}}{L(\boldsymbol{\beta}^{(j)}, \tau^{(j)}|\mathbf{D})L(\boldsymbol{\beta}^{(j)}, \tau^{(j)}|\mathbf{D}_0)^{a^{(j)}}}$$

4. Gere u de uma distribuição uniforme $U(0, 1)$ e use a seguinte regra de aceitação:

$$a^{(j+1)} = \begin{cases} a^{(j+1)} & \text{se } u \leq \alpha \\ a^{(j)} & \text{se } u > \alpha \end{cases} \quad (3.45)$$

5. Gere $\boldsymbol{\beta}^{(j+1)}$ de $\boldsymbol{\beta}^{(j+1)}|\tau^{(j)}, a^{(j+1)}, \mathbf{D}, \mathbf{D}_0 \sim N(\hat{\boldsymbol{\beta}}, \tau^{(j)-1}I_2)$,
6. Calcula-se a probabilidade de aceitação de $\boldsymbol{\beta}^{(j+1)}$, por $\alpha = \min\{1, q\}$, em que

$$q = \frac{L(\boldsymbol{\beta}^{(j+1)}, \tau^{(j+1)}|\mathbf{D}_0)^{a^{(j+1)}}}{L(\boldsymbol{\beta}^{(j)}, \tau^{(j+1)}|\mathbf{D}_0)^{a^{(j+1)}}}$$

7. Gere u de uma distribuição uniforme $U(0, 1)$ e use a seguinte regra de aceitação:

$$\boldsymbol{\beta}^{(j+1)} = \begin{cases} \boldsymbol{\beta}^{(j+1)} & \text{se } u \leq \alpha \\ \boldsymbol{\beta}^{(j)} & \text{se } u > \alpha \end{cases} \quad (3.46)$$

8. Gere $\tau^{(j+1)}$ de $\tau^{(j+1)}|\boldsymbol{\beta}^{(j+1)}, a^{(j+1)}, \mathbf{D}, \mathbf{D}_0 \sim G\left(\frac{n_1}{2}, \vartheta(\boldsymbol{\beta}^{(j+1)}, a^{(j+1)})\right)$
9. Faça $j \leftarrow j + 1$ e repita os passos (2)-(6) até a convergência ser atingida.

Para encontrar as estimativas bayesianas obtidas via simulação de Monte Carlo será adotado o mesmo procedimento adotado na seção 3.1.1.1. Considera-se um período de aquecimento (burn-in) de 50% das iterações iniciais. Todos os passos do algoritmo serão repetidos até que haja evidências de convergência, monitorada graficamente e com o critério de Geweke (GEWEKE, 1992).

Os estimadores bayesianos obtidos via simulação de Monte Carlo são calculados pela equação (3.8) e os intervalos de credibilidade por (3.9).

3.1.4.4 Aplicação

Os dados utilizados para testar a abordagem bayesiana com distribuição *a priori* potência foram coletadas em florestas plantadas de clones de *Eucalyptus sp.* descritos na seção 2.3. Dentre as várias parcelas utilizadas, foram selecionadas três, as parcelas A, C, D (Tabela (2.1)) para exemplificar os resultados obtidos.

Os resultados das estimativas bayesianas com densidade *a priori* potência (EBPP) são apresentadas na Tabela (3.11), assim como o critério de convergência de Geweke (1992) utilizado para verificar a convergência da cadeia gerada pelo algoritmo MCMC, sendo que, como critério de Geweke a convergência ocorre quando $|Z| < Z_c$ para $Z_c = 1,96$. Podemos verificar nessa tabela que os resultados obtidos utilizando a densidade de probabilidade *a priori* potência são sempre coerentes biologicamente. Todos os resultados respeitam as restrições impostas biologicamente, $\beta_i > 0, i = 0, 1, 2$. Esse resultado também pode ser visto nas Figuras (3.16), (3.17) e (3.18) onde os parâmetros estimados positivos resultam em ajustes crescentes.

Para a seleção de modelos, foram utilizados o critério de informação de Akaike (AIC), o critério de informação bayesiano (BIC), o critério desvio-informação (DIC) e o critério da densidade preditiva ordenada (maiores detalhes sobre seleção de modelo no capítulo 4). Os resultados são apresentados na Tabela (3.12). Para a parcela A, o melhor modelo selecionado por todos os critérios de seleção utilizados foi o modelo de Prodan e para as parcelas B e C, o modelo de Petterson.

Como exemplo, apresentamos o histograma e trajetória das amostras geradas com o algoritmo de Metrópolis-Hastings após verificada a convergência no modelo de Prodan, para os parâmetros β e τ nas Figuras (3.19) e (3.20), respectivamente. Essa convergência para o parâmetro β pode ser observada nos gráficos A1, B1 e C1 nos quais se verifica a tendência estacionária das séries simuladas dos parâmetros em torno das estimativas. Os histogramas A2, B2 e C2 nos mostra que as densidades *a posteriori* são unimodais. Na

Tabela 3.11: Estimativas bayesianas utilizando densidade *a priori* potência para os modelos hipsométricos dados na Tabela (1.1).

Par.	Modelo	Est. Bayesiana	D.P.	ICred.(95%)	CCG(Z)
A	Curtis	$\beta_0 = 3,4895$	0,1445	(3,2066; 3,7774)	0,0027
		$\beta_1 = 6,0197$	1,7723	(2,5279; 9,5985)	0,0026
		$S_2 = 0,4273$	0,1947	(0,1422; 0,8935)	0,0053
	Stofel	$\beta_0 = 1,5361$	0,1547	(1,2289; 1,8459)	0,0371
		$\beta_1 = 0,5751$	0,0597	(0,4559; 0,6945)	0,0412
		$S_2 = 0,5300$	0,1998	(0,2703; 1,0037)	0,0429
	Petterson	$\beta_0 = 0,3037$	0,0077	(0,2876; 0,3199)	0,0153
		$\beta_1 = 0,7900$	0,1039	(0,5682; 1,0049)	0,0176
		$S_2 = 0,0027$	0,0011	(0,0013; 0,0054)	0,0159
	Prodan	$\beta_0 = 0,7031$	0,2025	(0,3445; 1,1854)	0,0139
		$\beta_1 = 0,2078$	0,0380	(0,1159; 0,2726)	0,0163
		$\beta_2 = 0,0282$	0,0016	(0,0255; 0,2726)	0,0146
		$S_2 = 0,0029$	0,0012	(0,0014; 0,0058)	0,0025
B	Curtis	$\beta_0 = 3,4624$	0,0593	(3,3475; 3,5742)	0,0286
		$\beta_1 = 5,9676$	0,7314	(4,5579; 7,3475)	0,0224
		$S_2 = 0,0368$	0,0562	(0,0121; 0,2426)	0,0271
	Stofel	$\beta_0 = 1,5658$	0,1897	(1,1981; 1,9542)	0,0006
		$\beta_1 = 0,5496$	0,0707	(0,4034; 0,6871)	0,0014
		$S_2 = 0,4947$	0,1858	(0,2482; 0,9588)	0,0418
	Petterson	$\beta_0 = 0,3076$	0,0062	(0,2942; 0,3215)	0,0017
		$\beta_1 = 0,7862$	0,0911	(0,5761; 0,9825)	0,0003
		$S_2 = 0,0023$	0,0009	(0,0011; 0,0045)	0,0073
	Prodan	$\beta_0 = 0,6953$	0,1396	(0,3929; 0,9472)	0,0449
		$\beta_1 = 0,2112$	0,0301	(0,1574; 0,2771)	0,0432
		$\beta_2 = 0,02943$	0,0014	(0,0263; 0,2771)	0,0423
		$S_2 = 0,0023$	0,0009	(0,0011; 0,0046)	0,0192
C	Curtis	$\beta_0 = 3,7666$	0,0709	(3,6326; 3,9032)	0,0132
		$\beta_1 = 8,1163$	1,1319	(5,9169; 10,2791)	0,0081
		$S_2 = 0,0478$	0,0766	(0,0160; 0,2531)	0,0299
	Stofel	$\beta_0 = 1,8393$	0,1868	(1,4687; 2,2207)	0,0398
		$\beta_1 = 0,5049$	0,0622	(0,3806; 0,6282)	0,0418
		$S_2 = 0,6285$	0,2376	(0,2993; 0,1817)	0,0276
	Petterson	$\beta_0 = 0,2812$	0,0027	(0,2749; 0,2872)	0,0064
		$\beta_1 = 0,8882$	0,0549	(0,7633; 1,0129)	0,0069
		$S_2 = 0,0020$	0,0008	(0,0010; 0,0039)	0,0039
	Prodan	$\beta_0 = 1,6775$	0,0998	(1,4937; 1,9028)	0,0144
		$\beta_1 = 0,1023$	0,0129	(0,0734; 0,1261)	0,0169
		$\beta_2 = 0,0251$	0,0004	(0,0244; 0,1261)	0,0198
		$S_2 = 0,0007$	0,0002	(0,0004; 0,0012)	0,0023

Par. é a parcela utilizada, Est. bayesiana é a estimativa bayesiana obtida via MCMC, D.P. é o desvio padrão, ICred. o intervalo de credibilidade e CCG(Z) é o critério de convergência de Geweke (1992).

Tabela 3.12: Critérios de seleção de modelos para o ajuste bayesiano utilizando densidade *a priori* potência para os modelos hipsométricos dados na Tabela (1.1).

Parcela	Modelos	AIC	BIC	Log(CPO)	DIC
A	Prodan	-71,34	-60,99	122,11	-78,27
	Petterson	-70,83	-58,23	122,01	-72,91
	Curtis	-19,20	-18,72	97,45	-17,74
	Stofel	-15,41	-14,63	90,44	-15,08
B	Petterson	-82,45	-72,20	131,75	-87,16
	Prodan	-76,74	-63,08	130,89	-83,88
	Curtis	-50,61	-40,37	116,08	-53,94
	Stofel	-13,83	-12,92	79,78	-13,74
C	Petterson	-42,48	-33,57	99,55	-47,08
	Curtis	-41,24	-32,33	96,59	-44,82
	Stofel	-7,78	3,88	94,67	-13,69
	Prodan	201,35	213,23	12,94	192,72

AIC é o critério de informação de Akaike (1973), BIC o critério de informação bayesiano (SCHWARZ, 1978), DIC é o critério desvio-informação (SPIEGELHALTER et al., 2002) e Log(CPO) é o logaritmo do critério da densidade preditiva ordenada (GELFAND; DEY, 1994).

Figura (3.20) A temos a trajetória das amostras geradas para o parâmetro τ pelo algoritmo MCMC o qual nos indica que a convergência da série foi alcançada. O histograma B nos mostra que a distribuição *a posteriori* para τ é assimétrica positiva e unimodal. O resultado para as demais parcelas e modelos apresentam resultados semelhantes e são apresentados no apêndice C.2.

Podemos perceber comparando as Tabela (2.2) (EMV) e (3.11) (EBPP) a forte influência dos dados históricos nos resultados bayesianos. No caso da parcela A, o ajuste da parcela pelo método de máxima verossimilhança proporciona resultados superestimados, enquanto o ajuste bayesiano utilizando a densidade *a priori* empírica pondera seu resultado através dos dados históricos minimizando esse problema (veja Figura (3.16)). Isso ocorre porque a parcela A é medida em uma unidade (fração) da fazenda muito pequena, não levando em consideração as áreas vizinhas. Dessa forma, essa pequena área (parcela A) mensurada provavelmente contém características que aumentam sua produtividade, como melhor solo, água disponível, etc., características essas não presente no restante da área (fazenda) que está sendo avaliada.

O número pequeno de indivíduos em cada parcela (amostra) acarreta, em alguns casos, que esses dados possuam uma tendência biologicamente inconsistente. É o que ocorre com as parcelas B e C, que possuem um comportamento decrescente (veja o ajuste EMV das parcelas nas Figuras (3.17) e (3.18)). Esse comportamento pode ser visto também na

Tabela (2.2) para a parcela A, modelo de Prodan, o qual mostra o resultado $\beta_1 < 0$. Uma vez que podemos utilizar dados históricos no ajuste bayesiano, esse "problema" é corrigido através do poder da informação histórica no estudo presente (veja o ajuste bayesiano nas Figuras (3.17) e (3.18)), mostrando a primazia do método bayesiano quando comparado ao método clássico de ajuste.

Uma questão que pode ser levantada neste momento é, qual deve ser a densidade de probabilidade *a priori* escolhida para obter-se as estimativas bayesianas dos parâmetros restritos em modelos hipsométricos? No próximo capítulo, tentaremos responder essa questão, utilizando o critério da densidade preditiva ordenada.

(A1)

(A2)

(A3)

(A4)

Figura 3.16: Ajuste bayesiano utilizando *a priori* potência para a parcela A dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras A1, A2, A3 e A4.

(B1)

(B2)

(B3)

(B4)

Figura 3.17: Ajuste bayesiano utilizando *a priori* potência para a parcela B dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras B1, B2, B3 e B4.

(C1)

(C2)

(C3)

(C4)

Figura 3.18: Ajuste bayesiano utilizando *a priori* potência para a parcela C dos modelos de Curtis, Stofel, Petterson e Prodan, respectivamente Figuras C1, C2, C3 e C4.

Figura 3.19: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (C1 e C2), parcela A, com distribuição *a priori* potência.

Figura 3.20: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela A, com distribuição *a priori* potência.

4 Seleção de modelos

Vários critérios de seleção de modelos para a escolha do modelo hipsométrico que melhor se ajusta a um conjunto de dados podem ser adotados, ver por exemplo Akaike (1974), Raftery (1996), Schwarz (1978), Chen et al. (2008), Chen, Ibrahim e Yiannoutsos (1999). Dentre eles, serão adotados neste trabalho os critérios a seguir:

4.1 Critério de informação de Akaike

Akaike (1973, 1974) mostrou como obter um rigoroso critério de seleção de modelos baseado na informação de Kullback-Leibler (KULLBACK; LEIBLER, 1951) e na teoria da verossimilhança (para maiores detalhes ver deLeeuw (1992)). Ao maximizar o valor do log-verossimilhança foi encontrado um viés da estimativa de $E_y E_x [\log(g(x|\hat{\theta}(y)))]$, o qual era aproximadamente igual a K , o número de parâmetros estimados pelo modelo. Esse resultado assintótico foi de fundamental importância para que Akaike encontrasse o critério de informação sem desvio assintótico, o qual foi dado por:

$$AIC = -2\ln(L(\hat{\theta}|\mathbf{D})) + 2p \quad (4.1)$$

No caso especial da estimação ser por mínimos quadrados com distribuição dos erros normal, o AIC pode ser expresso como

$$AIC = n\ln(L(\hat{\sigma}^2)) + 2p \quad (4.2)$$

onde

$$\hat{\sigma}^2 = \frac{\sum(\hat{\epsilon}_i)^2}{n},$$

e $(\hat{\epsilon}_i)^2$ é o resíduo do ajuste do modelo. Neste caso, p deve ser o número de parâmetros do modelo, incluindo o intercepto e σ^2 .

O critério de informação de Akaike de segunda ordem, muitas vezes chamado AICc (AIC corrigido), leva em conta o tamanho da amostra, aumentando a penalização relativa

a complexidade do modelo com pequeno conjunto de dados. Segundo Hurvich e Tsai (1989) o AICc é dado por,

$$AICc = -2\ln(L(\hat{\boldsymbol{\theta}}|\mathbf{D})) + 2p\frac{n}{n-p-1} \quad (4.3)$$

onde o termo “penalizante” é multiplicado pelo fator de correção $n/(n-p-1)$. Este pode também ser escrito como

$$AICc = -2\ln(L(\hat{\boldsymbol{\theta}}|\mathbf{D})) + 2p + \frac{2p(p+1)}{n-p-1}, \quad (4.4)$$

ou equivalentemente,

$$AICc = AIC + \frac{2p(p+1)}{n-p-1} \quad (4.5)$$

onde n é o tamanho da amostra. O uso do AICc é recomendado sempre que o tamanho da amostra não for grande comparada ao número de parâmetros estimados no modelo. Geralmente advoga-se pelo uso do AICc quando a razão n/p é pequena (< 40) (BURNHAM; ANDERSON, 2002).

No contexto bayesiano, o AIC e AICc podem ser utilizados fazendo uma modificação que considera a amostra *a posteriori* $\{\boldsymbol{\theta}^{(g)}, g = 1, \dots, G\}$ e considerando $l \equiv \ln L(\boldsymbol{\theta})$ como a função paramétrica de interesse, dessa forma temos

$$\hat{AIC} = -2\hat{l} + 2p \quad (4.6)$$

onde

$$\hat{l} \equiv E[\ln L(\boldsymbol{\theta})|\mathbf{y}] \approx \frac{1}{G} \sum_{g=1}^G \ln L(\boldsymbol{\theta}^{(g)}).$$

O melhor modelo será aquele que apresentar o menor valor de AIC.

4.2 Critério de informação Bayesiano

O BIC (*Bayesian Information Criterion*) é um critério de seleção de modelo proposto por Schwarz (1978) e modificado por Carlin e Louis (1996) para ser aplicado considerando a posteriori dos parâmetros do modelo ajustado. Este critério pondera entre o máximo da função log-verossimilhança e o número de parâmetros do modelo. O melhor modelo é aquele que apresentar o menor valor de BIC dado por

$$\widehat{BIC} = -2\hat{l} + 2p\ln(n) \quad (4.7)$$

onde $\hat{l}$, p e n foram citados anteriormente.

4.3 Critério desvio-informação

O DIC (*Deviance Information Criterion*) é uma generalização do BIC. Este critério é particularmente usual nos problemas bayesianos de seleção de modelos para os quais amostras da distribuição *a posteriori* dos parâmetros dos modelos foram obtidas por simulação de Monte Carlo em Cadeia de Markov (MCMC). Semelhante ao BIC este critério é uma aproximação assintótica para amostras grandes, isto é válido quando a distribuição *a posteriori* é aproximadamente uma distribuição normal multivariada. Define-se o desvio como:

$$D(\boldsymbol{\theta}) = -2\ln(L(\boldsymbol{\theta})) + C$$

em que $\boldsymbol{\theta}$ é o vetor de parâmetros desconhecidos do modelo; $L(\boldsymbol{\theta})$ a função de verossimilhança e C é uma constante que se cancela, e portanto, não precisa ser conhecida na comparação de dois modelos. O critério DIC introduzido por SPIEGELHALTER et al. (2002) é dado por:

$$DIC = D(\hat{\boldsymbol{\theta}}) + 2p_D \quad (4.8)$$

sendo $D(\hat{\boldsymbol{\theta}})$ o desvio avaliado na média *a posteriori* e $p_D = \bar{D} - D(\hat{\boldsymbol{\theta}})$, em que $\bar{D}$ é o desvio médio *a posteriori* que mede a qualidade do ajuste do modelo aos dados e é dado por

$$\bar{D} \equiv E[-2\ln L(\boldsymbol{\theta})] \approx -2 \frac{1}{G} \sum_{g=1}^G \ln L(\boldsymbol{\theta}^{(g)}).$$

Valores menores para DIC indicam melhores modelos.

4.4 Critério da densidade preditiva ordenada

Gelfand e Dey (1994) propuseram o critério da densidade preditiva ordenada (CPO), hoje bastante aceito entre os pesquisadores (ver por exemplo Chen et al. (2008)). Para aplicar este critério considera-se a densidade preditiva para uma observação $y_i = \ln(H_i)$ com o x_i , onde $x_i = \ln(Dap_i)$ para o modelo de Stofel e $x_i = 1/Dap_i$ para os demais modelos, dado as informações $\mathbf{D} = \{(y_1, x_1), \dots, (y_n, x_n)\}$. A densidade preditiva para y_i condicionado aos dados $D_{(-i)}$ e considerando-se o modelo m , é dada por:

$$c_i(m) = f(y_i|\mathbf{D}, m) = \int_0^\infty \int_{\Theta} f(y_i|\mathbf{D}, \boldsymbol{\theta}, m) \pi(\boldsymbol{\theta}|\mathbf{D}, m) d\boldsymbol{\theta} d\tau \quad (4.9)$$

em que $\pi(\boldsymbol{\theta}, \tau|\mathbf{D})$ é a densidade *a posteriori* para os parâmetros $\boldsymbol{\theta} \in \Theta$, dado os dados $\mathbf{D}$ e Θ que é o espaço paramétrico do modelo m (m é um índice que indica o modelo).

Usando as amostras geradas pelo algoritmo Metrópolis-Hastings, a equação (4.9) pode ser aproximada pelo estimador de Monte Carlo dado por:

$$\hat{f}(y_i|\mathbf{D}, m) = \frac{1}{G} \sum_{j=1}^G f(y_i|\mathbf{D}, \boldsymbol{\theta}^{(j)}, \tau^{(j)}, m) \quad (4.10)$$

onde $\boldsymbol{\theta}^{(j)}$ e $\tau^{(j)}$, com $j = 1, \dots, G$ são as amostras geradas no MCMC.

A estimativa $\hat{c}_i(m) = \hat{f}(y_i|\mathbf{D}, m)$ pode ser usada para a seleção do modelo. Isto pode ser visualizado considerando-se o gráfico de $\hat{c}_i(m)$ versus i ($i = 1, \dots, n$) para os diferentes modelos m . A forma de usar o CPO é escolher o modelo m tal que $C_m = \prod_{i=1}^n c_i(m)$ seja máximo.

Outra forma de usar o CPO é escolher o modelo m tal que $\ln(C_m) = \sum_{i=1}^n \ln c_i(m)$ seja máximo, onde,

$$\ln(c_i(m)) = \ln f(y_i|\mathbf{D}, m) = \ln \left\{ \int_0^\infty \int_{\Theta} f(y_i|\mathbf{D}, \boldsymbol{\theta}, m) \pi(\boldsymbol{\theta}|\mathbf{D}, m) d\boldsymbol{\theta} d\tau \right\} \quad (4.11)$$

a equação (4.11) pode ser aproximada pelo estimador de Monte Carlo dado por:

$$\ln \hat{f}(y_i|\mathbf{D}, m) = \ln \left\{ \frac{1}{G} \sum_{j=1}^G f(y_i|\mathbf{D}, \boldsymbol{\theta}^{(j)}, \tau^{(j)}, m) \right\}. \quad (4.12)$$

As equações (4.9) e (4.10) tem as mesmas propriedades das equações (4.11) e (4.12). O maior valor de c_i ou $\ln(c_i(m))$ (em média) indica o melhor modelo.

4.5 Seleção da densidade de probabilidade *a priori* para modelos hipsométricos

Uma questão que não ficou respondida no capítulo anterior é, qual é a melhor densidade de probabilidade *a priori* para os modelos hipsométricos descritos na Tabela (1.1)? Optamos por utilizar a densidade preditiva ordenada (seção 4.4) para responder essa questão. Através desse critério, pegamos os valores do logaritmo da CPO dos melhores modelos para cada situação analisada (densidade *a priori* $\times$ parcela) e colocamos os resultados na Tabela (4.1). Para as parcelas A e B a melhor densidade de probabilidade *a priori* foi a *a priori* potência seguida da *priori* empírica. No caso da parcela C, o resultado se inverteu, ficando em primeiro a densidade de probabilidade *a priori* empírica, seguida da *priori* potência.

Nos muitos casos analisados, o resultado foi sempre o mesmo, pelo critério da den-

Tabela 4.1: Critério da densidade preditiva ordenada para os modelos hipsométricos selecionados para cada densidade *a priori* e parcela utilizada.

Parcela	Densidade de probabilidade <i>a priori</i>	Log(CPO)
A	Potência ⁴	122,11
	Empírica ¹	121,72
	Jeffreys ³	108,39
	Normal <i>flat</i> ⁴	101,91
B	Potência ³	131,75
	Empírica ²	120,93
	Jeffreys ³	98,31
	Normal <i>flat</i> ⁴	98,28
C	Empírica ³	102,57
	Potência ³	99,55
	Jeffreys ³	95,63
	Normal <i>flat</i> ⁴	95,15

1 = Modelo de Curtis, 2 = Modelo de Stofel, 3 = Modelo de Petterson, 4 = Modelo de Prodan, Log(CPO) é o logaritmo do critério da densidade preditiva ordenada (GELFAND; DEY, 1994).

sidade preditiva ordenada, as densidades *a priori* informativas foram sempre superiores as não informativas e vagas, independentemente de haver ou não, coerência biológica nas amostras. Isso é explicado pelo número pequeno de indivíduos nas parcelas. Pela pouca informação disponível nos dados, o conhecimento *a priori* torna-se de grande importância nas análises *a posteriori*.

Desse modo, a incorporação de informações e conhecimentos *a priori* melhorou as estimativas bayesianas obtidas para todos os modelos hipsométricos (Tabela (1.1)) e parcelas analisadas .

5 *Conclusões e Propostas Futuras*

As estimativas clássicas em alguns casos, não apresentaram coerência biológica. Muito provavelmente por causa do número restrito de indivíduos na amostra. Nestes casos nenhum dos quatro modelos propostos tiveram resultados adequados, não respeitando as restrições nos parâmetros impostas biologicamente. Dessa forma, não podem ser utilizados de maneira adequada para o planejamento e exploração dos recursos florestais.

Nos casos em que as parcelas utilizadas possuíam coerência biológica, o estimador de máxima verossimilhança produziu resultados satisfatórios. No entanto, o uso de intervalos de confiança assintótico não foram adequados, provavelmente porque a suposição que é feita sobre possuir uma amostra grande não está sendo satisfeita. Em contrapartida, os intervalos de confiança bootstrap foram bastante satisfatórios. Em todas as parcelas analisadas, através do critério de seleção de Schwarz, o modelo de Petterson foi escolhido o mais adequado.

Na abordagem bayesiana utilizando densidades *a priori* não informativa ou vaga os resultados se assemelharam aos obtidos com a metodologia clássica, nas quais, em alguns casos não foram respeitadas as restrições nos parâmetros ($\beta > 0$). Nestes casos, as estimativas dos parâmetros obtidas não puderam ser utilizadas para estimar a altura das demais árvores da parcela que não foram mensuradas.

Na abordagem bayesiana com densidade *a priori* construída empiricamente não foram obtidos resultados incoerentes biologicamente em nenhuma parcela analisada. Além disso, seus resultados coincidiram com os encontrados com EMV quando os mesmos foram factíveis. Dessa forma, as estimativas bayesianas com densidade *a priori* empírica mostrou-se superior as EMV, EBPJ e EBPJN.

A última densidade de probabilidade *a priori* proposta foi a densidade *a priori* potência, a qual conseguiu agregar mais informação histórica nas análises quando comparada com a abordagem bayesiana empírica. Com esse enfoque minimizou-se problemas de superestimação que ocorreram com EMV em parcelas com coerência biológica, mas que

em média possuíam dados superiores aos encontrados no restante da área (fazenda) em estudo, como foi o caso da parcela A. Além disso, a restrição aos parâmetros do modelo foram atendidas assim como ocorreu com a densidade *a priori* empírica.

Através do critério da densidade preditiva ordenada, escolheu-se a melhor densidade de probabilidade *a priori*. Para as parcelas A e B a melhor escolha foi a densidade *a priori* potência e para a parcela C, a densidade *a priori* empírica. Isso mostra que no caso de relações hipsométricas, a incorporação de informações e conhecimentos *a priori* melhora as estimativas bayesianas.

É importante ressaltar que na maioria das parcelas analisadas (isso engloba as que não estão citadas neste texto), a densidade de probabilidade *a priori* potência foi na grande maioria das vezes selecionada em detrimento das demais.

Neste contexto, a abordagem bayesiana com densidade *a priori* potência mostra-se como uma ferramenta promissora no âmbito florestal uma vez que os dados disponíveis são quase sempre reduzidos. Como trabalhos futuros, a utilização de densidades *a priori* informativas para modelar o crescimento e produção florestal mostra-se bastante promissora, uma vez que os dados para esse fim são, em geral, reduzidos. Com as informações geradas por esses modelos é possível inferir sobre a melhor estratégia de plantio, a análise econômica de tratamentos silviculturais, o planejamento da época de corte, a alocação de parte das árvores ou povoamentos para diferentes produtos, dentre outros.

Referências

- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: PETROV, B. N.; CSAKI, F. (Ed.). Budapest: Akademiai Kiado, 1973. p. 267–81.
- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on automatic control**, v. 19, p. 716–23, 1974.
- ALBERT, J.; CHIB, S. Computation in bayesian econometrics: An introduction to markov chain monte carlo. **Advances in Econometrics Volume 11A: Computational Methods and Applications**, JAI Press, Greenwich, p. 3–24, 1996.
- AMARAL, M. T. reis. **Abordagem bayesiana para curva de crescimento com restrições nos parâmetros**. 109 p. Dissertação (Mestrado Biometria e Estatística Aplicada) — Universidade Federal Rural de Pernambuco, Recife-PE, 2008.
- ANDRADE, M. G. et al. Uma abordagem bayesiana para curvas de crescimento com diferentes densidades a priori. **Revista Brasileira de Biometria**, São Paulo, v. 28, n. 2, p. 161–181, 2010.
- BAILEY, M. Ideal point estimation with a small number of votes: A random-effects approach. **Political Analysis**, Oxford, v. 9, n. 3, p. 192–210, 2001.
- BARROS, K. N. N. de O. **Abordagem clássica e bayesiana em modelos simétricos transformados aplicados à estimativa de crescimento em altura de *Eucalyptus urophylla* no polo gesseiro do Araripe-PE**. 84 p. Dissertação (Mestrado Biometria e Estatística Aplicada) — Universidade Federal Rural de Pernambuco, Recife-PE, 2010.
- BHATTACHARJEE, S. K.; DUNSMORE, I. R. The influence of variables in a logistic model. **Biometrika**, Great Britain, v. 78, n. 4, p. 851–6, 1991.
- BICKEL, P.; DOKSUM, K. **Mathematical statistics: basic ideas and selected topics**. [S.l.]: Pearson Prentice Hall, 2006. (Mathematical Statistics).
- BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical Analysis**. New York: Addison-Wesley, 1973. 588 p.
- BURNHAM, K. P.; ANDERSON, D. R. **Model selection and multimodel inference: A practical information-theoretic approach**. New York: Springer Science, 2002. 497 p.
- CARLIN, B. P.; LOUIS, T. A. **Bayes and empirical bayes methods for data analysis**. London: Chapman & Hall, 1996. 399 p.
- CHEN, M.-H. et al. Bayesian variable selection and computation for generalized linear models with conjugate priors. **Bayesian Analysis**, v. 3, n. 3, p. 585–614, 2008.

- CHEN, M.-H.; IBRAHIM, J. G.; SHAO, Q.-M. Power prior distributions for generalized linear models. **Journal of statistical planning and inference**, v. 84, p. 121–137, 2000.
- CHEN, M.-H.; IBRAHIM, J. G.; SHAO, Q.-M. Prior elicitation for model selection and estimation in generalized linear mixed models. **Journal of statistical planning and inference**, v. 111, p. 57–76, 2003.
- CHEN, M.-H.; IBRAHIM, J. G.; YIANNOUTSOS, C. Prior elicitation, variable selection and bayesian computation for logistic regression models. **J. R. Statist. Soc. B**, v. 61, p. 223–242, 1999.
- COOK, R. D.; TSAI, C.-L. Residuals in nonlinear regression. **Biometrika**, v. 72, n. 1, p. 23–29, 1985.
- COOK, R. D.; WEISBERG, S. **Residuals and influence in regression**. London: Chapman & Hall, 1982. 230 p.
- CORDEIRO, G. M.; PRUDENTE, A. A.; DEMÉTRIO, C. G. B. Uma revisão dos modelos normais não-lineares. **Revista Brasileira de Biometria**, São Paulo, v. 27, n. 3, p. 360–393, 2009.
- COUTO, H. T. Z.; BASTOS, N. L. M. Modelos de equações de volume e relações hipsométricas para plantações de eucalyptus no estado de são paulo. **IPEF**, Piracicaba, SP, n. 37, p. 33–44, dez. 1986.
- CURTIS, R. O. Height diameter and height diameter age equations for second growth douglas-fir. **Forest Science**, v. 13, n. 4, p. 365–375, 1967.
- DANIELS, M. J.; POURAHMADI, M. Bayesian analysis of covariance matrices and dynamic models for longitudinal data. **Biometrika**, Great Britain, v. 89, n. 3, p. 553–566, 2002.
- DELEEUW, J. Introduction to akaike (1973) information theory and an extension of the maximum likelihood principle. In: _____. **Breakthroughs in Statistics**. London, 1992. p. 599–609.
- DIACONIS, P.; YLVISAKER, D. Conjugate priors for exponential families. **The Annals of Statistics**, v. 7, n. 2, p. 269–281, 1979.
- DUAN, Y.; YE, K.; SMITH, E. P. Evaluating water quality using power priors to incorporate historical information. **Environmetrics**, v. 17, p. 95–106, 2006.
- EFRON, B.; TIBSHIRANI, R. J. **An introduction to the Bootstrap**. New York: Chapman & Hall, 1993. 436 p.
- FILHO, S. M. et al. Abordagem bayesiana das curvas de crescimento de duas cultivares de feijoeiro. **Ciência Rural**, Santa Maria, RS, v. 38, n. 6, p. 1516–1521, 2008.
- GELFAND, A. E.; DEY, D. K. Bayesian model choice: Asymptotics and exact calculations. **Journal of the Royal Statistician Society B**, v. 56, n. 3, p. 501–514, 1994.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. Clarendon Press, Oxford, UK, 1992. In *Bayesian Statistics 4*.

GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. **Markov Chain Monte Carlo in Practice**. USA: Chapman & Hall, 1998. 486 p.

GOLFARI, L.; CASER, R. L.; MOURA, V. P. G. **Zoneamento ecológico esquemático para reflorestamento no Brasil: 2ª aproximação**. [S.l.]: Centro de pesquisas florestais da região do cerrado, 1978. 66 p.

HURVICH, C. M.; TSAI, C.-L. Regression an time series model selection in small samples. **Biometrika**, v. 76, n. 2, p. 297–307, 1989.

IBRAHIM, J. G.; CHEN, M.-H. Power prior distributions for regression models. **Statistical Science**, Institute of Mathematical Statistics, v. 15, n. 1, p. 46–70, 2000.

IBRAHIM, J. G.; CHEN, M.-H.; SINHA, D. On optimality properties of the power prior. **Journal of the american statistical association**, v. 98, p. 204–213, 2003.

KER, J.; SMITH, J. Sampling for height-diameter relationships. **Journal of Forestry**, Amsterdam, v. 55, n. 3, p. 205–207, 1957.

KULLBACK, S.; LEIBLER, R. A. On information and suficiency. **Annals of mathematical statistics**, v. 22, p. 79–86, 1951.

MACHADO, S. A.; BASSO, S. F.; BEVILACQUA, V. Teste de modelos matemáticos para o ajuste da relação hipsométrica em diferentes sítios e idades para plantações de *Pinus elliottii* no estado do paraná. In: CONGRESSO FLORESTAL PANAMERICANO 1, CONGRESSO FLORESTAL BRASILEIRO 7, 1993, Curitiba. **Anais...** São Paulo: Sociedade Brasileira de Silvicultura, 1993. v2, p. 553–556.

MAZUCHELI, J. **Análise bayesiana e discriminação de modelos não lineares**. 169 p. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) — Universidade de São Paulo, São Carlos, SP, 1995.

MIGUEL, E. P. **Avaliação biométrica e prognose da produção de Eucalyptus urophylla (S. T. Blake) na região norte do estado de Goiás**. 161 p. Dissertação (Mestrado em Engenharia Florestal) — Universidade Federal do Paraná, Curitiba - PR, 2009.

MOREIRA, M. F. B. et al. O uso de métodos bayesianos empíricos para ajustar relações hipsométricas com o modelo de curtis com restrições nos parâmetros. **Revista Brasileira de Biometria**, São Paulo, v. 25, n. 3, p. 23–38, 2007.

MORRIS, C. N. Natural exponential families with quadratic variance functions: statistical theory. **The Annals of Statistics**, v. 11, n. 2, p. 515–529, 1983.

R Development Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2010. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>>.

RAFTERY, A. E. Hypothesis testing and model selection via posterior simulation. In: _____. **Practical Markov Chain Monte Carlo**. New York, 1996. p. 163–187.

- RIBEIRO, A. et al. Estratégias e metodologias de ajuste de modelos hipsométricos em plantios de eucalyptus sp. **Cerne**, Lavras, v. 16, n. 1, p. 22–31, 2010.
- RUFINO, R. F. et al. Ajuste de modelos hipsométricos para um povoamento de eucaliptos conduzido sobre o sistema de rebrota. **Enciclopédia Biosfera**, Goiânia, v. 6, n. 10, p. 1–10, 2010.
- SCHWARZ, G. Estimating the dimension of a model. **Annals of statistics**, v. 6, n. 2, p. 461–464, 1978.
- SCOLFORO, J. R. S. **Biometria florestal 2: técnica de regressão aplicada para estimar: volume, biomassa, relação hipsométrica e múltiplos produtos de madeira**. Lavras-MG: UFLA/FAEPE, 1997. (Fundação de Apoio ao Ensino, Pesquisa e Extensão).
- SEBER, G. A. F.; WILD, C. J. **Nonlinear regression**. New York: Wiley, 1989. 768 p.
- SILVA, G. F. et al. Análise da influência de diferentes tamanhos e composições de amostras no ajuste de uma relação hipsométrica para eucalyptus grandis. **Revista Árvore**, Viçosa-MG, v. 31, n. 4, p. 685–694, 2007.
- SOARES, T. S. et al. Uso de diferentes alternativas para viabilizar a relação hipsométrica no povoamento florestal. **Revista Árvore**, Viçosa, MG, v. 28, n. 6, p. 845–854, 2004.
- SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. **Journal of the Royal Statistician Society B**, v. 64, n. 4, p. 583–639, 2002.

APÊNDICE A – Resultados do método clássico

Neste apêndice são apresentados os gráficos dos modelos clássicos ajustados pelo método de máxima verossimilhança omitidos no capítulo 2. Os resultados são bastante semelhantes aos já apresentados no capítulo da abordagem clássica. Nas Figuras (A.1), (A.2) e (A.3) são apresentados os resultados para a parcela A, nas Figuras (A.4), (A.5) e (A.6) os resultados para a parcela B, nas Figuras (A.7), (A.8) e (A.9) são expostos os resultados para a parcela C e nas Figuras (A.10), (A.11) e (A.12) os resultados para a parcela D, respectivamente para os modelos de Curtis, Stofel e Prodan.

Assim como ocorreu para o modelo de Petterson escolhido pelo critério BIC no capítulo 2, quando não havia coerência biológica nos dados da parcela os resultados não foram satisfatórios, uma vez que o comportamento do modelo ajustado era decrescente (Figuras (A.4), (A.5), (A.6), (A.7), (A.8) e (A.9)). Nestes casos, tornou-se inviável sua utilização, pois não geraram estimativas coerentes para a altura das árvores. Para os demais casos ($\beta_i > 0$) que podem ser vistos nas Figuras (A.1), (A.2), (A.3), (A.10), (A.11) e (A.12), o modelo selecionado pelo critério BIC foi o modelo de Petterson já exposto no capítulo 2.

Figura A.1: Ajuste do modelo de Curtis para a parcela A (A), resíduo ordinário *versus* Dap (B), resíduo estudntizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura A.2: Ajuste do modelo de Stofel para a parcela A (A), resíduo ordinário *versus* Dap (B), resíduo estudentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).

Figura A.3: Ajuste do modelo de Prodan para a parcela A (A), resíduo ordinário *versus* Dap (B), resíduo studentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura A.4: Ajuste do modelo de Curtis para a parcela B (A), resíduo ordinário *versus* Dap (B), resíduo estudentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).

Figura A.5: Ajuste do modelo de Stofel para a parcela B (A), resíduo ordinário *versus* Dap (B), resíduo studentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura A.6: Ajuste do modelo de Prodan para a parcela B (A), resíduo ordinário *versus* Dap (B), resíduo estudentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura A.7: Ajuste do modelo de Curtis para a parcela C (A), resíduo ordinário *versus* Dap (B), resíduo estudentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura A.8: Ajuste do modelo de Stofel para a parcela C (A), resíduo ordinário *versus* Dap (B), resíduo estudentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura A.9: Ajuste do modelo de Prodan para a parcela C (A), resíduo ordinário *versus* Dap (B), resíduo studentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).

Figura A.10: Ajuste do modelo de Curtis para a parcela D (A), resíduo ordinário *versus* Dap (B), resíduo studentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

Figura A.11: Ajuste do modelo de Stofel para a parcela D (A), resíduo ordinário *versus* Dap (B), resíduo studentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida Di (F).

Figura A.12: Ajuste do modelo de Prodan para a parcela D (A), resíduo ordinário *versus* Dap (B), resíduo studentizado *versus* Dap (C), gráfico normal qq-plot dos resíduos (D), gráfico de pontos de alavanca (E) e gráfico da medida D_i (F).

APÊNDICE B – Priori empírica para os modelos hipsométricos

B.1 Modelos de Stofel e Curtis

Para os modelos de Stofel e Curtis será adotada uma abordagem empírica para a determinação dos hiperparâmetros β_0 e $\mathbf{P}$. Para isso será considerado os percentis (H_p, Dap_p) e (H_q, Dap_q) , com $p < q$, encontrados em dados históricos coletados anteriormente às observações da amostra D , os quais não ocorreram obrigatoriamente de forma pareada (neste trabalho será considerada nas aplicações $p = 25\%$ e $q = 75\%$, mas outros percentis podem ser considerados sem perda de generalidade). Assumindo que o modelo de Curtis e Stofel (Tabela (1.1)) são válidos para esses percentis, podemos escrever:

$$\beta_{0,0} = y_p + \left(\frac{y_q - y_p}{x_q - x_p} \right) x_p \quad (\text{B.1})$$

$$\beta_{1,0} = - \left(\frac{y_q - y_p}{x_q - x_p} \right) \quad (\text{B.2})$$

onde estamos considerando $y_k = \ln(H_k)$, $x_k = \ln(Dap_k)$ para $k = p, q$ e $\beta_0 = (\beta_{0,0}, \beta_{1,0})'$. A matriz de precisão $\mathbf{P}$ também pode ser encontrada de forma empírica considerando um coeficiente de variação constante $CV = \gamma_j / \beta_{j,0}$, $j = 0, 1$ que expresse a confiança nos hiperparâmetros β_0 . Desta forma temos:

$$P_{j,j} = \frac{1}{\gamma_j^2} = \frac{1}{(CV\beta_{j,0})^2}, \quad j = 0, 1 \quad (\text{B.3})$$

Vamos assumir também que β_0 e β_1 são independentes a priori, de forma que $P_{i,j} = 0$ para $i \neq j$, ($i, j = 0, 1$).

B.2 Modelo de Petterson

Os hiperparâmetros da densidade a priori, dos coeficientes do modelo de Petterson, devem satisfazer as restrições impostas aos parâmetros, tal que a relação entre a altura total (H) e o diâmetro (Dap) sejam biologicamente consistentes, ou seja eles devem ser tal que a relação entre altura total (H) e diâmetros (D) tenha um comportamento côncavo, assegurando assim a priori, um comportamento assintótico consistente do ponto de vista biológico.

Vamos considerar uma abordagem empírica para encontrar os hiperparâmetros (β_0) da densidade a priori dos coeficientes do modelo de Petterson. Para isto vamos considerar um par (H, x) , onde $x = 1/Dap$, então a relação ((1.1)) para esse par de observações é dada por:

$$H^{1/3} = \mu(x; \beta) \quad (\text{B.4})$$

da Tabela (1.1) para o modelo de Petterson podemos escrever $\mu(x; \beta) = 1/g(x)$ com $g(x)$ dada por:

$$g(x) = \beta_0 + \beta_1 x \quad (\text{B.5})$$

então a relação (B.4) pode ser escrita como:

$$H^{1/3} = \frac{1}{g(x)} \quad (\text{B.6})$$

Para que a relação (B.6) seja côncava, a função $g(x)$ deve ser uma função convexa. Para assegurar essa propriedade, considerando que $g(x)$ tem a forma polinomial dada em (B.5) e além disso devemos ter $\beta_i > 0$, $i = 0, 1$, então vamos considerar $g(x)$ como uma aproximação polinomial da função convexa dada por:

$$\tilde{g}(x) = \alpha_0 e^{\alpha_1 x} \quad \text{com} \quad \alpha_0 > 0 \quad \text{e} \quad \alpha_1 > 0 \quad (\text{B.7})$$

considerando que $x = 1/Dap \in (0, 1]$ vamos considerar uma expansão de Taylor de (B.7), dada por:

$$\tilde{g}(x) = \tilde{g}(0) + \left(\frac{d\tilde{g}(0)}{dx} \right) x + o(x^2) \quad (\text{B.8})$$

Calculando-se as derivadas de (B.7) e substituindo-se em (B.6) desprezando-se os termos de ordem maior ou igual a x^2 , temos que:

$$\tilde{g}(x) \approx \alpha_0 + (\alpha_0 \alpha_1) x \quad (\text{B.9})$$

Comparando-se a equação (B.9) com a equação (B.5) temos que:

$$\begin{aligned}\beta_0 &= \frac{\alpha_0 \alpha_1^2}{2} \\ \beta_1 &= \alpha_0 \alpha_1\end{aligned}\tag{B.10}$$

Portanto para encontrar valores de $\beta_i > 0, i = 0, 1$ basta encontrar valores $\alpha_i > 0, i = 0, 1$. Para calcular os parâmetros α_0 e α_1 de $\tilde{g}(x)$ vamos considerar os percentis (H_p, Dap_p) e (H_q, Dap_q) , com $p < q$, encontrados em dados históricos coletados anteriormente às observações da amostra D , os quais não ocorreram obrigatoriamente de forma pareada (neste trabalho consideramos nas aplicações $p = 25\%$ e $q = 75\%$, mas outros percentis podem ser considerados sem perda de generalidade). Assumindo que o modelo de Petterson (B.6) é válido para esses percentis, com $\tilde{g}(x)$ no lugar de $g(x)$, podemos escrever:

$$H_p^{1/3} = \frac{1}{\alpha_0 e^{\alpha_1 x_p}} \quad \text{com} \quad x_p = \frac{1}{Dap_p}\tag{B.11}$$

$$H_q^{1/3} = \frac{1}{\alpha_0 e^{\alpha_1 x_q}} \quad \text{com} \quad x_q = \frac{1}{Dap_q}\tag{B.12}$$

definindo por $y_k = \ln(1/H_k^{1/3})$ $k = p, q$, podemos escrever as equações (B.11) e (B.12) como:

$$y_p = \ln(\alpha_0) + \alpha_1 x_p\tag{B.13}$$

$$y_q = \ln(\alpha_0) + \alpha_1 x_q\tag{B.14}$$

Resolvendo o sistema linear (B.13)-(B.14) para α_0 e α_1 temos:

$$\alpha_0 = \exp \left\{ y_p + \left(\frac{y_q - y_p}{x_q - x_p} \right) x_p \right\}\tag{B.15}$$

$$\alpha_1 = \left(\frac{y_q - y_p}{x_q - x_p} \right)\tag{B.16}$$

A matriz de precisão $\mathbf{P}$ também pode ser encontrada de forma empírica considerando um coeficiente de variação constante $CV = \gamma_i / \beta_{0,i}$, $i = 0, 1$ que expresse a confiança nos hiperparâmetros β_0 . Desta forma temos:

$$P_{ii} = \frac{1}{\gamma_i^2} = \frac{1}{(CV \beta_{0,i})^2}, \quad i = 0, 1, 2\tag{B.17}$$

Vamos assumir também que os coeficientes $\beta_{0,i}$, $i = 0, 1$ são independentes a priori, de forma que $P_{ij} = 0$ para $i \neq j$, ($i, j = 0, 1$).

B.3 Modelo de Prodan

Vamos considerar uma abordagem empírica para encontrar os hiperparâmetros (β_0) da densidade a priori dos coeficientes do modelo de Prodan. Para isto vamos considerar um par (H, x) , onde $x = 1/Dap$, então a relação (1.1) para esse par de observações é dada por:

$$H = \mu(x; \beta) \quad (\text{B.18})$$

da Tabela (1.1) para o modelo de Prodan podemos escrever $\mu(x; \beta) = 1/g(x)$ com $g(x)$ dada por:

$$g(x) = \beta_0 x^2 + \beta_1 x + \beta_2 \quad (\text{B.19})$$

então a relação (B.18) pode ser escrita como:

$$H = \frac{1}{g(x)} \quad (\text{B.20})$$

Para que a relação (B.20) seja côncava, a função $g(x)$ deve ser uma função convexa. Para assegurar essa propriedade, considerando que $g(x)$ tem a forma polinomial dada em (B.19) e além disso devemos ter $\beta_i > 0$, $i = 0, 1, 2$, então vamos considerar $g(x)$ como uma aproximação polinomial da função convexa dada por:

$$\tilde{g}(x) = \alpha_0 e^{\alpha_1 x} \quad \text{com} \quad \alpha_0 > 0 \quad \text{e} \quad \alpha_1 > 0 \quad (\text{B.21})$$

considerando que $x = 1/Dap \in (0, 1]$ vamos considerar uma expansão de Taylor de (B.21), dada por:

$$\tilde{g}(x) = \tilde{g}(0) + \left(\frac{d\tilde{g}(0)}{dx} \right) x + \frac{1}{2} \left(\frac{d^2\tilde{g}(0)}{dx^2} \right) x^2 + O(x^3) \quad (\text{B.22})$$

Calculando-se as derivadas de (B.21) e substituindo-se em (B.20) desprezando-se os termos de ordem maior ou igual a x^3 , temos que:

$$\tilde{g}(x) \approx \alpha_0 + (\alpha_0 \alpha_1) x + \left(\frac{\alpha_0 \alpha_1^2}{2} \right) x^2 \quad (\text{B.23})$$

Comparando-se a equação (B.23) com a equação (B.19) temos que:

$$\begin{aligned} \beta_0 &= \frac{\alpha_0 \alpha_1^2}{2} \\ \beta_1 &= \alpha_0 \alpha_1 \\ \beta_2 &= \alpha_0 \end{aligned} \quad (\text{B.24})$$

Portanto para encontrar valores de $\beta_i > 0$, $i = 0, 1, 2$ basta encontrar valores $\alpha_i > 0$, $i = 0, 1$. Para calcular os parâmetros α_0 e α_1 de $\tilde{g}(x)$ vamos considerar os percentis (H_p, Dap_p)

e (H_q, Dap_q) , com $p < q$, encontrados em dados históricos coletados anteriormente às observações da amostra D , os quais não ocorreram obrigatoriamente de forma pareada (neste trabalho consideramos nas aplicações $p = 25\%$ e $q = 75\%$, mas outros percentis podem ser considerados sem perda de generalidade). Assumindo que o modelo de Prodan (B.20) é válido para esses percentis, com $\tilde{g}(x)$ no lugar de $g(x)$, podemos escrever:

$$H_p = \frac{1}{\alpha_0 e^{\alpha_1 x_p}} \quad \text{com} \quad x_p = \frac{1}{Dap_p} \quad (\text{B.25})$$

$$H_q = \frac{1}{\alpha_0 e^{\alpha_1 x_q}} \quad \text{com} \quad x_q = \frac{1}{Dap_q} \quad (\text{B.26})$$

definindo por $y_k = \ln(1/H_k)$ $k = p, q$, podemos escrever as equações (B.25) e (B.26) como:

$$y_p = \ln(\alpha_0) + \alpha_1 x_p \quad (\text{B.27})$$

$$y_q = \ln(\alpha_0) + \alpha_1 x_q \quad (\text{B.28})$$

Resolvendo o sistema linear (B.27)-(B.28) para α_0 e α_1 temos:

$$\alpha_0 = \exp \left\{ y_p + \left(\frac{y_q - y_p}{x_q - x_p} \right) x_p \right\} \quad (\text{B.29})$$

$$\alpha_1 = \left(\frac{y_q - y_p}{x_q - x_p} \right) \quad (\text{B.30})$$

A matriz de precisão $\mathbf{P}$ também pode ser encontrada de forma empírica considerando um coeficiente de variação constante $CV = \gamma_i / \beta_i$, $i = 0, 1, 2$ que expresse a confiança nos hiperparâmetros β_0 . Desta forma temos:

$$P_{ii} = \frac{1}{\gamma_i^2} = \frac{1}{(CV\beta_i)^2}, \quad i = 0, 1, 2 \quad (\text{B.31})$$

Vamos assumir também que os coeficientes β_i , $i = 0, 1, 2$ são independentes a priori, de forma que $P_{ij} = 0$ para $i \neq j$, ($i, j = 0, 1, 2$).

APÊNDICE C – Resultados do método bayesiano

Como foi discutido no capítulo 3, as estimativas dos parâmetros bayesianos foram obtidas através de simulação de Monte Carlo. Gera-se uma cadeia de Markov a qual é utilizada para se obter as estimativas bayesianas. Durante o processo, verifica-se a convergência da cadeia graficamente e por critério Geweke (1992). Devido ao grande número de gráficos de convergência gerados para as diversas situações aqui apresentadas, optamos por apresentar os demais gráficos de convergência ausentes no texto apenas para as duas situações mais promissoras, a abordagem bayesiana com densidade *a priori* empírica e a abordagem bayesiana com densidade *a priori* potência.

Vale ressaltar que para as outras duas densidades de probabilidade *a priori*, de Jeffreys e normal *flat* os gráficos de convergência são semelhantes aos já apresentados no capítulo 3, nas Figuras (3.4), (3.5), (3.9) e (3.10).

C.1 Resultados do método bayesiano utilizando densidade de probabilidade *a priori* empírica

Os gráficos de convergência para os demais casos não apresentados no capítulo 3 seção 3.1.3 são apresentados nas Figuras (C.1), (C.2), (C.3), (C.4), (C.5), (C.6), (C.7), (C.8), (C.9), (C.10), (C.11), (C.12), (C.13), (C.14), (C.15), (C.16), (C.17), (C.18), (C.19), (C.20), (C.21) e (C.22).

Figura C.1: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição *a priori* empírica.

Figura C.2: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela A, com distribuição *a priori* empírica.

Figura C.3: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição *a priori* empírica.

Figura C.4: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela A, com distribuição *a priori* empírica.

Figura C.5: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela A, com distribuição *a priori* empírica.

Figura C.6: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela A, com distribuição *a priori* empírica.

Figura C.7: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição *a priori* empírica.

Figura C.8: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela B, com distribuição *a priori* empírica.

Figura C.9: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição *a priori* empírica.

Figura C.10: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela B, com distribuição *a priori* empírica.

Figura C.11: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela B, com distribuição *a priori* empírica.

Figura C.12: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela B, com distribuição *a priori* empírica.

Figura C.13: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela B, com distribuição *a priori* empírica.

Figura C.14: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela B, com distribuição *a priori* empírica.

Figura C.15: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição *a priori* empírica.

Figura C.16: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela C, com distribuição *a priori* empírica.

Figura C.17: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição *a priori* empírica.

Figura C.18: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela C, com distribuição *a priori* empírica.

Figura C.19: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição *a priori* empírica.

Figura C.20: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela C, com distribuição *a priori* empírica.

Figura C.21: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela C, com distribuição *a priori* empírica.

Figura C.22: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela C, com distribuição *a priori* empírica.

C.2 Resultados do método bayesiano utilizando densidade de probabilidade *a priori* potência

Os gráficos de convergência para os demais casos não apresentados no capítulo 3 seção 3.1.4 são apresentados nas Figuras (C.23), (C.24), (C.25), (C.26), (C.27), (C.28), (C.29), (C.30), (C.31), (C.32), (C.33), (C.34), (C.35), (C.36), (C.37), (C.38), (C.39), (C.40), (C.41), (C.42), (C.43) e (C.44).

Figura C.23: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 ($A1$ e $A2$) e β_1 ($B1$ e $B2$), parcela A, com distribuição *a priori* potência.

Figura C.24: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela A, com distribuição *a priori* potência.

Figura C.25: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela A, com distribuição *a priori* potência.

Figura C.26: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela A, com distribuição *a priori* potência.

Figura C.27: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A_1 e A_2) e β_1 (B_1 e B_2), parcela A, com distribuição *a priori* potência.

Figura C.28: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela A, com distribuição *a priori* potência.

Figura C.29: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 ($A1$ e $A2$) e β_1 ($B1$ e $B2$), parcela B, com distribuição *a priori* potência.

Figura C.30: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela B, com distribuição *a priori* potência.

Figura C.31: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 ($A1$ e $A2$) e β_1 ($B1$ e $B2$), parcela B, com distribuição *a priori* potência.

Figura C.32: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela B, com distribuição *a priori* potência.

Figura C.33: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 (A_1 e A_2) e β_1 (B_1 e B_2), parcela B, com distribuição *a priori* potência.

Figura C.34: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela B, com distribuição *a priori* potência.

Figura C.35: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela B, com distribuição *a priori* potência.

Figura C.36: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela B, com distribuição *a priori* potência.

Figura C.37: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, para os parâmetros β_0 ($A1$ e $A2$) e β_1 ($B1$ e $B2$), parcela C, com distribuição *a priori* potência.

Figura C.38: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Curtis, parcela C, com distribuição *a priori* potência.

Figura C.39: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, para os parâmetros β_0 (A1 e A2) e β_1 (B1 e B2), parcela C, com distribuição *a priori* potência.

Figura C.40: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Stofel, parcela C, com distribuição *a priori* potência.

Figura C.41: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, para os parâmetros β_0 ($A1$ e $A2$) e β_1 ($B1$ e $B2$), parcela C, com distribuição *a priori* potência.

Figura C.42: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Petterson, parcela C, com distribuição *a priori* potência.

Figura C.43: Histograma e trajetória das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, para os parâmetros β_0 (A1 e A2), β_1 (B1 e B2) e β_2 (B1 e B2), parcela C, com distribuição *a priori* potência.

Figura C.44: Trajetória (A1) e histograma (A2) das amostras geradas com o algoritmo de simulação MCMC após ser verificada a convergência para o modelo de Prodan, parcela C, com distribuição *a priori* potência.