

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
COORDENAÇÃO DE ENGENHARIA FLORESTAL  
CÂMPUS DOIS VIZINHOS

LUDIAN CARLOS VITORELO

**AVALIAÇÃO DA ESTRUTURA DE FLORESTAS TROPICAIS A  
PARTIR DA SIMULAÇÃO E INFERÊNCIA BAYESIANA**

TRABALHO DE CONCLUSÃO DE CURSO II

DOIS VIZINHOS

2015

**LUDIAN CARLOS VITORELO**

**AVALIAÇÃO DA ESTRUTURA DE FLORESTAS TROPICAIS A  
PARTIR DA SIMULAÇÃO E INFERÊNCIA BAYESIANA**

Trabalho de Conclusão de Curso apresentado à disciplina de Trabalho de Conclusão de Curso II, do Curso Superior de Engenharia Florestal da Universidade Tecnológica Federal do Paraná – UTFPR, como requisito parcial para obtenção do título de Engenheiro Florestal.

Orientador: Prof. Dr. Edgar de Souza Vismara

**DOIS VIZINHOS**

**2015**

V845a Vitorelo, Ludian Carlos.  
Avaliação da estrutura de florestas tropicais a partir da simulação e inferência bayesiana. / Ludian Carlos Vitorelo – Dois Vizinhos: [s.n], 2015.  
58f.

Orientador: Edgar de Souza Vismara  
Trabalho de Conclusão de Curso (graduação) -  
Universidade Tecnológica Federal do Paraná. Curso de Engenharia Florestal. Dois Vizinhos, 2015.  
Bibliografia p.49-51

1.Floresta tropical. 2.Estrutura diamétrica. 3.Método Bayesiano I.Vismara, Edgar de Souza, orient.  
II.Universidade Tecnológica Federal do Paraná – Dois Vizinhos. III.Título

CDD: 634.9

Ficha catalográfica elaborada por Keli Rodrigues do Amaral CRB: 9/1559

Biblioteca da UTFPR-Dois Vizinhos



---

## TERMO DE APROVAÇÃO

Avaliação da estrutura de florestas tropicais a partir de simulação e inferência Bayesiana

por

Ludian Carlos Vitorelo

Este Trabalho de Conclusão de Curso II foi apresentado em 25 de novembro de 2015 como requisito parcial para a obtenção do título de Bacharel em Engenharia Florestal. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

---

Prof. Dr. Edgar de Souza Vismara  
Orientador

---

Prof<sup>a</sup>. Dra. Veridiana Padoin Weber  
Membro titular (UTFPR)

---

Prof<sup>a</sup>. Msc. Lilian de Souza Vismara  
Membro titular (UTFPR)

- O Termo de Aprovação assinado encontra-se na Coordenação do Curso -

## **AGRADECIMENTOS**

Agradeço a minha mãe Olga e ao meu pai Luide por tudo que fizeram e fazem por mim.

Agradeço ao orientador Edgar de Souza Vismara por todo o conhecimento repassado e que tornou possível a execução deste trabalho.

Agradeço ao meu amigo Matheus e as amigas Dayana e Mariana pela amizade e por toda a ajuda e incentivo durante a graduação.

Agradeço a todos aqueles que me ajudaram na conclusão deste trabalho.

Muito obrigado a todos!

## RESUMO

VITORELO, Ludian Carlos. AVALIAÇÃO DA ESTRUTURA DE FLORESTAS TROPICAIS A PARTIR DA SIMULAÇÃO E INFERÊNCIA BAYESIANA. 58 f. Trabalho de Conclusão de Curso II – Câmpus Dois Vizinhos, Universidade Tecnológica Federal do Paraná. Dois Vizinhos, 2015.

A principal ferramenta de estudo da dinâmica da vegetação arbórea é a distribuição diamétrica. Diferentes métodos são utilizados para o ajuste de funções, entretanto apenas o método Bayesiano considera a experiência do pesquisador no ajuste através da utilização de informações *a priori*. Desta forma, o objetivo deste trabalho foi verificar a importância de diferentes *prioris* no ajuste da distribuição Weibull e a sua influência na estimativa do volume em cenários simulados de uma floresta tropical. O conjunto de dados foi obtido de uma área localizada predominantemente no município de Linhares (ES). A amostragem foi estratificada, sendo cada classe diamétrica considerada como um estrato com 2 árvores, obtendo-se um conjunto com 200 indivíduos. Foram reamostrados 3 cenários com 20 parcelas de 30 árvores cada, no qual a reamostragem é inversa, equiprovável e proporcional ao dap. O conjunto resultante de cada reamostragem foi dividido em 7 classes de diâmetros, com intervalos de 5 cm cada. Para cada cenário foram ajustados por verossimilhança 4 modelos hipsométricos (Polinômio do 1º, Stoffels, Richards e Weibull) e 3 modelos volumétricos (Spurr, Schumacher-Hall e Ogaya) para posterior estimativa do volume por classe diamétrica, cujos critérios de seleção do melhor ajuste foram o AIC e o BIC. No ajuste da distribuição Weibull pelo método Bayesiano foram utilizados 4 níveis de informação para os parâmetros do modelo e somente a verossimilhança. Para o ajuste foi utilizado o algoritmo MWG, 3 cadeias de Markov e 100 mil iterações iniciais. Foram utilizados os critérios de Geweke, BMK, ESS e MCSE para análise da convergência das cadeias de Markov e utilizado o DIC e o fator de Bayes para a seleção do melhor ajuste da função Weibull. A estimativa do número de árvores por hectare se deu através da função de densidade acumulada da distribuição Weibull. O volume por classe diamétrica foi estimado utilizando os parâmetros dos melhores modelos biométricos ajustados. Para o cenário 1, os melhores modelos biométricos ajustados foram o de Richards e Schumacher-Hall e a distribuição Weibull se ajustou melhor pelo critério DIC com a utilização da *priori* 3 ( $a, b$  e  $c \sim \text{Gama}(5; 5)$ ), sendo a diferença entre o volume total amostrado e o volume total estimado de 20,18%. Para o cenário 2, os melhores modelos biométricos ajustados foram o de Richards e Schumacher-Hall e o melhor ajuste da distribuição Weibull, segundo o DIC, foi obtido com a utilização da *priori* 3, sendo a diferença entre o volume total amostrado e o volume total estimado de 25,93%. Para o cenário 3, os melhores modelos biométricos ajustados foram o de Richards e Ogaya e o melhor ajuste da distribuição Weibull, segundo o critério DIC, foi obtido com a utilização da verossimilhança, sendo a diferença entre o volume total amostrado e o volume total estimado de 7%. A utilização do método Bayesiano não foi eficiente para o ajuste da distribuição Weibull em virtude do tamanho da amostra ser muito grande, favorecendo a verossimilhança e também pela própria distribuição dos dados.

**Palavras-chave:** Estrutura diamétrica, Reamostragem, *Prioris*, Modelagem

## ABSTRACT

VITORELO, Ludian Carlos. Evaluation of tropical forest structure from the simulation and bayesian inference. 58 f. Trabalho de Conclusão de Curso II – Câmpus Dois Vizinhos, Universidade Tecnológica Federal do Paraná. Dois Vizinhos, 2015.

The main dynamics study tool of trees is the diameter distribution. Different methods are used for the adjustment functions, but only the Bayesian method considers the experience of the researcher in the setting through the use of information *a priori*. Thus, the aim of this study was to assess the importance of different *prioris* in the setting of Weibull distribution and its influence on the estimated volume in simulated scenarios of a rainforest. The dataset was obtained from an area located predominantly in Linhares (ES). The sample was stratified, each diameter class regarded as a layer 2 tree, yielding a set comprised 200 individuals. They were resampled three scenarios with 20 plots of 30 trees each, in which the resampling is reversed, equiprobable and proportional to dap. The resulting set of each resampling was divided into 7 classes of diameters, each 5 cm intervals. For each scenario were adjusted likelihood 4 hypsometric models (polynomial of the 1st, Stoffels, Richards and Weibull) and three volumetric models (Spurr, Schumacher-Hall and Ogaya) for further estimate of the volume by diameter class, whose the best fit selection criteria were the AIC and BIC. In the setting of Weibull distribution by Bayesian method was used four levels of information for the model parameters and only likelihood. For adjustment we used the MWG algorithm, 3 Markov chains and 100,000 early iterations. The criteria for Geweke, BMK, ESS and MCSE to analyze the convergence of Markov chains and used the DIC and the Bayes factor for the selection of best fit the Weibull function were used. Estimates of the number of trees per hectare was made through the accumulated density of the Weibull distribution function. The volume by diameter class was estimated using the parameters of the best adjusted biometric templates. For scenario 1, the top set biometric templates were to Richards and Schumacher-Hall and the Weibull distribution to best fit the criterion DIC with the use of *priori* 3 ( $a, b \text{ e } c \sim \text{Gama}(5; 5)$ ), and the difference between the total volume sampled and the volume estimated total of 20.18%. For scenario 2, the top set biometric templates were to Richards and Schumacher-Hall, and the best fit of the Weibull distribution, according DIC was obtained with the use of *priori* 3, and the difference between the total volume sampled and the estimated total volume of 25.93%. For scenario 3, the best set biometric templates were to Richards and Ogaya and the best fit of the Weibull distribution, the discretion DIC was obtained using the likelihood, and the difference between the total volume sampled and the estimated total volume 7%. Using the Bayesian method was not effective for adjusting the Weibull distribution because the sample size is very large, favoring the likelihood and also because the distribution of the data

**Keywords:** Diametric sctructure, Resampling, *Prioris*, Modeling

## LISTA DE TABELAS

|           |   |    |
|-----------|---|----|
| TABELA 1  | – Relação dos modelos biométricos de altura e volume com casca ajustados por máxima verossimilhança .....   | 25 |
| TABELA 2  | – Interpretação dos valores de AIC para a seleção de modelos .....  | 25 |
| TABELA 3  | – Interpretação do fator de Bayes .....   | 30 |
| TABELA 4  | – Comparação dos modelos biométricos ajustados por máxima verossimilhança pelos critérios de informação AIC e BIC. ....   | 32 |
| TABELA 5  | – Estimativa dos parâmetros da distribuição Weibull e diagnósticos de convergência das cadeias de Markov para a verossimilhança e <i>prioris</i> do cenário 1 ..... | 34 |
| TABELA 6  | – Comparação do ajuste dos modelos pelo fator de Bayes para o cenário 1, onde o valor indica a superioridade do modelo da linha sobre o modelo da coluna. ....      | 36 |
| TABELA 7  | – Estimativa do número de árvores por hectare por classe diamétrica para cada modelo no cenário 1. ....   | 37 |
| TABELA 8  | – Estimativa do volume por hectare (m <sup>3</sup> /ha) por classe diamétrica para cada modelo no cenário 1 .....   | 38 |
| TABELA 9  | – Comparação dos modelos biométricos ajustados por máxima verossimilhança pelos critérios de informação AIC e BIC. ....   | 38 |
| TABELA 10 | – Diagnósticos de convergência das cadeias de Markov para a verossimilhança e <i>prioris</i> do cenário 2 .....   | 40 |
| TABELA 11 | – Comparação do ajuste dos modelos pelo fator de Bayes para o cenário 2, onde o valor indica a superioridade do modelo da linha sobre o modelo da coluna. ....      | 41 |
| TABELA 12 | – Estimativa do número de árvores por hectare por classe diamétrica para cada modelo no cenário 2. ....   | 43 |
| TABELA 13 | – Estimativa do volume por hectare (m <sup>3</sup> /ha) por classe diamétrica para cada modelo no cenário 2 .....   | 44 |
| TABELA 14 | – Comparação dos modelos biométricos ajustados por máxima verossimilhança pelos critérios de informação AIC e BIC. ....   | 45 |
| TABELA 15 | – Estimativa dos parâmetros da distribuição Weibull e diagnósticos de convergência das cadeias de Markov para a verossimilhança e <i>prioris</i> do cenário 3 ..... | 46 |
| TABELA 16 | – Comparação do ajuste dos modelos pelo fator de Bayes para o cenário 3, onde o valor indica a superioridade do modelo da linha sobre o modelo da coluna. ....      | 48 |
| TABELA 17 | – Estimativa do número de árvores por hectare por classe diamétrica para cada modelo no cenário 3. ....   | 48 |
| TABELA 18 | – Estimativa do volume por hectare (m <sup>3</sup> /ha) por classe diamétrica para cada modelo no cenário 3 .....   | 49 |
| TABELA 19 | – Diagnóstico de Convergência BMK das cadeias de Markov para a verossimilhança e <i>prioris</i> do cenário 1 .....  | 55 |

|  |    |
|--|----|
| TABELA 20 – Diagnóstico de Convergência BMK das cadeias de Markov para a verossimilhança e <i>prioris</i> do cenário 2 ..... | 56 |
| TABELA 21 – Diagnóstico de Convergência BMK das cadeias de Markov para as <i>prioris</i> do cenário 3 .....                  | 57 |
| TABELA 22 – Diagnóstico de Convergência BMK das cadeias de Markov para as <i>prioris</i> do cenário 3 .....                  | 58 |

## SUMÁRIO

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>INTRODUÇÃO</b>  | <b>10</b> |
| 1.1      | OBJETIVO GERAL   | 11        |
| 1.2      | OBJETIVOS ESPECÍFICOS  | 11        |
| <b>2</b> | <b>REFERENCIAL TEÓRICO</b>   | <b>12</b> |
| 2.1      | FLORESTAS TROPICAIS  | 12        |
| 2.2      | ESTRUTURA E MODELAGEM DIAMÉTRICA DE FLORESTAS NATURAIS                     | 14        |
| 2.3      | SIMULAÇÃO  | 16        |
| 2.4      | REAMOSTRAGEM   | 17        |
| 2.5      | INFERÊNCIA BAYESIANA   | 18        |
| 2.5.1    | Teorema de Bayes   | 18        |
| 2.5.2    | Verossimilhança  | 19        |
| 2.5.3    | Métodos de Monte Carlo via cadeias de Markov - MCMC                        | 21        |
| <b>3</b> | <b>METODOLOGIA</b>   | <b>23</b> |
| 3.1      | ÁREA DE ESTUDO E COLETA DE DADOS   | 23        |
| 3.2      | SIMULAÇÃO E REAMOSTRAGEM   | 23        |
| 3.3      | AJUSTE DOS MODELOS BIOMÉTRICOS   | 24        |
| 3.4      | COMPARAÇÃO DO AJUSTE E SELEÇÃO DOS MODELOS                                 | 25        |
| 3.5      | AJUSTE DA DISTRIBUIÇÃO WEIBULL   | 26        |
| 3.6      | VERIFICAÇÃO DA CONVERGÊNCIA DAS CADEIAS DE MARKOV                          | 27        |
| 3.7      | AVALIAÇÃO DO AJUSTE DA DISTRIBUIÇÃO WEIBULL                                | 29        |
| 3.8      | ESTIMATIVA DO NÚMERO DE ÁRVORES E DO VOLUME POR CLASSE<br>DIAMÉTRICA       | 30        |
| <b>4</b> | <b>RESULTADOS E DISCUSSÃO</b>  | <b>32</b> |
| 4.1      | CENÁRIO 1  | 32        |
| 4.1.1    | Comparação do ajuste dos modelos biométricos                               | 32        |
| 4.1.2    | Avaliação da convergência das cadeias de Markov e seleção do melhor modelo | 33        |
| 4.1.3    | Estimação do número de árvores e volume por hectare                        | 36        |
| 4.2      | CENÁRIO 2  | 38        |
| 4.2.1    | Comparação dos modelos biométricos   | 38        |
| 4.2.2    | Avaliação da convergência das cadeias de Markov e seleção do melhor modelo | 39        |
| 4.2.3    | Estimativa do número de árvores e volume por hectare                       | 43        |
| 4.3      | CENÁRIO 3  | 44        |
| 4.3.1    | Comparação dos modelos biométricos   | 44        |
| 4.3.2    | Avaliação da convergência das cadeias de Markov e seleção do melhor modelo | 45        |
| 4.3.3    | Estimação do número de árvores e volume por hectare                        | 48        |
| <b>5</b> | <b>CONCLUSÃO</b>   | <b>50</b> |
|          | <b>REFERÊNCIAS</b>   | <b>51</b> |
| <b>6</b> | <b>APÊNDICE - A</b>  | <b>54</b> |

## 1 INTRODUÇÃO

As florestas tropicais são ecossistemas únicos e complexos que recobrem cerca de 15% da superfície terrestre, no qual o Brasil é país privilegiado pela extensão, tipos e diversidade destas florestas. Encontram-se florestas tropicais em praticamente todas as unidades da federação, abrigando as formações florestais da Amazônia e da Mata Atlântica, as quais constituem imensas reservas de recursos naturais passíveis de exploração.

A degradação que as florestas tropicais, tanto a nível mundial quanto nacional, vem sofrendo ao longo do tempo, devido a ação antrópica desordenada faz com que estudos relacionados a sua dinâmica sejam necessários para entender o seu comportamento e permitir que sejam elaborados planos de conservação e formas de exploração sustentável.

A distribuição diamétrica é um dos parâmetros mais importantes no estudo das florestas inequidâneas. A partir deste parâmetro é possível conhecer o estado regenerativo da floresta, o incremento periódico em diâmetro, a dinâmica de sucessão florestal e o estado regenerativo da floresta, além de ser a base para a extração sustentável a partir do conceito de floresta balanceada.

A distribuição diamétrica pode ser estudada através do ajuste de funções de distribuições de diâmetros, principalmente a partir das funções de densidade de probabilidade (fdp). Entre estas fdp a função Weibull de três parâmetros vem sendo utilizada para esta finalidade devido as suas características de ajuste e flexibilidade em relação a sua forma e grau de inclinação.

A distribuição Weibull pode ser ajustada por vários métodos diferentes, entre os quais o método de máxima verossimilhança e o método bayesiano. O método de máxima verossimilhança baseia-se na informação contida nos dados amostrados para determinar qual o modelo com a maior possibilidade de ter gerado aquela amostra. O método bayesiano além de considerar a informação contida na amostra utiliza o conhecimento a priori do pesquisador sobre o fenômeno pesquisado combinando-os a partir do Teorema de Bayes.

Com o avanço das técnicas de computação e dos recursos computacionais, a utilização

de métodos matemáticos como reamostragem e simulação vem sendo utilizados com o propósito de auxiliar o pesquisador a reduzir o tamanho do conjunto de dados a serem coletados e a inferir conclusões de forma mais rápida utilizando cenários hipotéticos. Estas ferramentas se mostram promissoras por apresentarem resultados aproximados com certo grau de precisão em tempos muito menores do que os demais métodos, além de serem menos dispendiosos em relações aos custos envolvidos.

A utilização combinada de reamostragem, simulação e ajuste bayesiano de modelos pode fornecer estimativas da estrutura de florestas tropicais utilizando pequenos conjuntos de dados, permitindo ao manejador florestal a tomada de decisões ligadas a forma de manejo que será utilizada a fim de manter a estabilidade da floresta para que a mesma se perpetue e continue produtiva.

## 1.1 OBJETIVO GERAL

Verificar a importância de diferentes *prioris* para o ajuste da distribuição Weibull e a sua influência na estimativa do volume em cenários simulados de uma floresta tropical.

## 1.2 OBJETIVOS ESPECÍFICOS

- Simular cenários da estrutura diamétrica de floresta tropical;
- Ajustar a distribuição Weibull por inferência Bayesiana utilizando MCMC e verossimilhança;
- Comparar a convergência das cadeias de Markov e o ajuste da distribuição;
- Estimar o número de árvores e volume por classe diamétrica;
- Inferir sobre as *prioris* utilizadas e sobre a metodologia de ajuste.

## 2 REFERENCIAL TEÓRICO

### 2.1 FLORESTAS TROPICAIS

As florestas tropicais são as formações florestais predominantes nos trópicos úmidos, localizadas entre as latitudes 23,5°N e 23,5°S da linha do Equador, nos continentes da Ásia, Oceania, África e Américas do Sul e Central. Estas regiões não são homogêneas em termos de temperatura e umidade, e desta forma, há muitos locais que não suportam estas formações. Estas florestas são encontradas em áreas relativamente quentes e de temperatura constante desde o nível do mar até altitudes de 3000 metros (HOLL, 2013).

Devido a distribuição particular dos continentes e oceanos, e a circulação de massas de ar e correntes marinhas, há uma grande variação de precipitação, umidade do ar, temperatura, ventos e outras importantes características climatológicas. Tais mudanças no microclima são responsáveis pelo padrão de distribuição e também afetam a composição e estrutura das florestas tropicais (LONGMAN; JENÍK, 1974, p. 8). Entretanto, de forma geral, a precipitação anual nos trópicos úmidos excede 2000 mm e é distribuída durante todo o ano, e em pelo menos 8 meses tem-se mais do que 100 mm de chuva. Em relação a temperatura, a média anual em regiões de florestas tropicais é de 27°C, com médias mensais oscilando entre 24 e 28°C. A combinação de chuva abundante e temperaturas sem variações bruscas permite que as árvores da floresta mantenham-se verdes durante todo o ano, oferecendo suporte a uma rica fauna e flora e a um nível de produção biológica que é a maior que qualquer outro ecossistema natural do mundo (MARTEN, 1984, p. 62).

Segundo Longman e Jeník (1974, p. 13) devido a um número de fatores paleogeográficos associados a diferenças contemporâneas de clima, solos e interferências humanas, as florestas tropicais podem ser divididas em até quatro grandes regiões:

1. Região de florestas tropicais americanas (incluindo partes da América do Sul, Central e Caribe);
2. Região de florestas tropicais africanas (incluindo a Bacia do Congo, costa oeste africana,

planaltos do leste africano e Madagascar);

3. Região de florestas tropicais indo-malaia (incluindo partes da Índia, Myanmar, Península Malaia e ilhas do sudoeste asiático);
4. Região de florestas tropicais da Australásia (incluindo o nordeste australiano, Nova Guiné e ilhas adjacentes do Oceano Pacífico). As ilhas do Havaí representam uma quinta menor região.

Em relação a estrutura, as florestas tropicais são bastantes distintas em relação as demais. As árvores tendem a ocorrer em três estratos, embora os mesmos não sejam distintos. O estrato principal consiste de uma cobertura densa e continua de árvores codominantes com altura entre 30 e 60 metros que compõem o dossel da floresta. Outro estrato é composto por árvores dominantes de grande altura que atravessam o dossel do estrato principal em locais pontuais. O terceiro estrato é composto por árvores dominadas localizadas abaixo do dossel e que estão distribuídas de forma descontínua. As árvores maiores são adaptadas a luz direta, enquanto as do subosque são adaptadas a sombra. O subosque é denso somente em locais onde as árvores maiores caíram e deixaram uma clareira, permitindo que a luz solar chegue para as demais árvores abaixo (MARTEN, 1984, p. 62).

Em se tratando da estrutura horizontal, as florestas tropicais são compostas por um grande número de espécies arbóreas e apresentam, portanto, uma alta diversidade, quando comparada as florestas de clima temperado. Os indivíduos de cada espécie arbórea distribuem-se geralmente de forma casualizada ou de forma agrupada (POGGIANI, 1989). Em estudo em uma floresta tropical úmida realizado por Golley et al. (1978) apud Poggiani (1989), encontrou-se apenas 480 árvores por hectare com diâmetro superior a 10 cm. Destas, 35 ou 40 apresentavam diâmetro superior a 30 cm e apenas 10 ou 12 com diâmetro acima de um metro.

Dois fatores desempenham um papel fundamental no estudo e na compreensão de uma floresta tropical: seu tamanho e escala temporal de seu desenvolvimento. Quando são examinadas as pradarias pode se obter parcelas representativas com alguns decímetros de altura e poucos metros quadrados, Uma parcela amostral adequada para uma floresta tropical mede aproximadamente 40 m de altura por 4 hectares em área (LONGMAN; JENÍK, 1974, p. 7).

Desta forma, não se pode esperar compreender toda a estrutura a partir de um ponto no chão da floresta, e este problema de tamanho e perspectiva fazem toda observação ecológica e qualquer abordagem muito difícil e inconveniente. Além disso, outro problema na investigação florestal é a grande expectativa de vida e idade dos organismos envolvidos, e

além do mais o período de desenvolvimento da comunidade florestal. É impossível para um único pesquisador estudar a dinâmica ou o crescimento de árvores individuais em intervalos de tempo superiores a 150 anos ou mais, sendo esta a expectativa de vida de árvores tropicais emergentes. Necessariamente, a partir de então, estudos experimentais são limitados em muitos casos a estágios juvenis de árvores em desenvolvimento (LONGMAN; JENÍK, 1974, p. 7).

Entretanto a complexidade, tamanho e escalas de tempo inconvenientes criaram condições especiais para a pesquisa em florestas tropicais. A estreita cooperação de muitos especialistas é necessária, e seleção cuidadosa de parcelas amostrais e planejamento engenhoso de experimentos são fundamentais. Em particular, problemas de taxas de crescimento, regeneração natural e sobrevivência de árvores tropicais podem ser resolvidas por colaboração contínua (LONGMAN; JENÍK, 1974, p. 7).

## 2.2 ESTRUTURA E MODELAGEM DIAMÉTRICA DE FLORESTAS NATURAIS

Em florestas naturais onde a idade é desconhecida, a distribuição diamétrica é um parâmetro importante, pois mostra a amplitude dos diâmetros, onde ocorre a maior concentração do número de árvores, mostra a distinção dos tipos florestais e possibilita quando associada a dados de incremento periódico do diâmetro ou a mudança de uma classe diamétrica para a outra a elaboração de tabelas de produção que consideram a dinâmica do povoamento florestal. Além disso, indica como é a regeneração natural das espécies de uma floresta e da própria floresta como um todo, sendo também uma importante medida de estoque. Ainda, a distribuição diamétrica possibilita que seja implementada a técnica de remoção de árvores por classe diamétrica, desde que seja utilizado o conceito de floresta balanceada, ou a análise da estrutura da floresta ou ainda a análise do estoque de regeneração em nível de espécie (SCOLFORO, 2006, p. 195).

Segundo Duan (2013, p. 1), a distribuição de diâmetros é bem conhecida e amplamente utilizada para descrever a estrutura de diâmetros de um povoamento florestal. A quantificação exata das características da árvore permite estudar a interação entre processos físicos e fisiológicos e o crescimento e, ao longo do tempo, permite ao manejador relacionar parâmetros da distribuição a idade ou densidade do povoamento. Além disso, o volume do povoamento pode ser calculado a partir do uso da distribuição de diâmetro, altura das árvores e de modelos de volume. Ressalta-se ainda que a predição do crescimento e produção com base na abordagem da distribuição de diâmetros também tem sido amplamente utilizada.

A distribuição diamétrica de florestas inequiâneas é descrita como um “J” invertido,

podendo ser explicada pela função exponencial negativa, no qual o número de árvores por hectare decresce com o aumento do centro de classe de diâmetro. Desta forma, há um grande número de árvores nas classes inferiores e um número muito menor nas classes superiores (LEAK, 1964; MEYER, 1952; HOUGH, 1932 apud BARROS 1980, p. 7). Em manejo florestal busca-se a definição e sustentação do conceito de floresta balanceada, em que o número de árvores por hectare decresce segundo uma progressão geométrica, assegurando a continuidade do ciclo de colheitas econômico e ecologicamente exequível e que mantenha a capacidade de sustentação das produções futuras (SOUZA; SOARES, 2013, p. 78).

A distribuição diamétrica pode ser estudada a partir do emprego de modelos de distribuição de diâmetros, cuja principal característica é o emprego de uma função de densidade de probabilidade. O ajuste destes modelos necessita da estimação dos parâmetros de alguma distribuição contínua de probabilidades. Para a obtenção das produções futuras é necessário que seja estimado o número de árvores por hectare e contar com equações de altura e volume por árvore individual. Os dados para a construção destes modelos devem ser obtidos a partir de parcelas permanentes ou através da análise de tronco completa (CAMPOS; LEITE, 2013, p. 428).

Várias funções de densidade de probabilidade como a Normal, Lognormal, Gama, Beta, SB de Johnson e Weibull tem sido amplamente utilizadas para descrever a distribuição de diâmetros de povoamentos florestais ao longo dos últimos 30 anos. Além disso, em estudos com distribuições de diâmetros cumulativos, diferentes equações de crescimento teóricas, tais como a Logística, Gompertz, Mitscherlich, Bertalanffy, Schumacher, Korf, Weibull e Richards foram utilizadas para caracterizar a estrutura diamétrica de povoamentos florestais (DUAN et al., 2013, p. 1). Outras distribuições como a Johnson, Logística generalizada, Hiperbólica, Gumbel, Nakagami, Dagum, Pearson e Burr também são utilizadas com uso menos frequente (CAMPOS; LEITE, 2013, p. 429).

Segundo Duan et al. (2013, p. 1) a função Weibull é a função densidade de probabilidade mais utilizada para o ajuste de distribuição diamétrica de árvores. Desde a obtenção da função Weibull de três parâmetros por Weibull (1939) e por conta da relativa simplicidade da fórmula e a sua flexibilidade em adaptar uma variedade de formas e graus de inclinação, esta função tem se provado ser um bom modelo de distribuição. Para a estimativa dos seus parâmetros, diversos métodos têm sido aplicados, como o método dos momentos, método da máxima verossimilhança, método dos percentis e método de regressão não-linear.

## 2.3 SIMULAÇÃO

A simulação é essencialmente uma modalidade experimental de pesquisa que tem como objetivo inferir conclusões através de modelos que teoricamente expressam a realidade. A simulação é, portanto, um processo de representação da realidade por meio de modelos, os quais podem ou não conservar as características físicas e lógicas do sistema representado (SHIMIZU, 1994 apud SILVA, 2001, p. 17).

Segundo Maria (1997, p. 7) o modelo pode ser configurado e experimentado quando, usualmente, isso é impossível, muito caro ou impraticável no sistema que ele representa. O funcionamento do modelo pode ser estudado, e então, propriedades relativas ao comportamento do sistema atual ou subsistema podem ser inferidas. No seu sentido mais amplo, a simulação é uma ferramenta para avaliar o desempenho de um sistema, existente ou proposto, sob diferentes configurações de interesse e para longos períodos de tempo. Ainda, segundo o mesmo autor (p. 7), a simulação é usada antes da alteração ou construção de um novo sistema, para reduzir as chances de falhas, atender as especificações, eliminar gargalos não previstos, prevenir sub ou sobre-utilização de recursos, e para otimizar o desempenho de sistemas.

Os modelos de simulação segundo Andrade (1989) apud Silva (2001, p. 17) podem ser contínuos ou discretos e determinísticos ou estocásticos. Em determinados casos, o modelo pode conter tanto elementos determinísticos quanto estocásticos e, em determinadas situações variáveis estocásticas podem ser tratadas como se fossem conhecidas exatamente. Uma simulação pode ser feita por meio de processos aleatórios, sendo a evolução no decorrer do tempo não previsível, ou por meio dos métodos de Monte Carlo. No primeiro caso, os fenômenos ou eventos da natureza podem ser determinísticos, quando as suas ocorrências são avaliadas com certeza, ou aleatórios, quando suas ocorrências envolvem riscos ou incertezas.

Modelos de simulação florestal tem provado ser uma ferramenta útil para o manejo florestal. Estes modelos são construídos por várias razões para uma grande variedade de usuários, desde gestores a estudantes. Os modelos de simulação possuem diversas aplicações, como predição do volume das árvores, otimização silvicultural apropriada para maximizar o rendimento, entendimento da sucessão e competição florestal, avaliação dos impactos ambientais causados por estresses como poluição do ar, chuva ácida, e mudanças climáticas, avaliação da sustentabilidade de ecossistemas florestais e teste de várias hipóteses sobre a estrutura das árvores. Estes diversos objetivos envolvem processos dinâmicos que variam na escala temporal de minutos a séculos, e são aplicados em escala espaciais que variam de folhas a ecossistemas (PENG; WEN, 2006, p. 109).

Na literatura podem ser encontrados extensos trabalhos de revisão sobre a aplicação de modelos de simulação em manejo florestal como os realizados por Botkin (1993), Vanclay (1995), Shugart e Smith (1996), Newton (1997), Battaglia e Stands (1998), Mäkelä e Lansberg (2000) e Lansberg (2003) (PENG; WEN, 2006, p. 109).

## 2.4 REAMOSTRAGEM

Testes paramétricos clássicos comparam estatísticas observadas a amostras de distribuições teóricas. A reamostragem é uma metodologia revolucionária porque se afasta de distribuições teóricas. Pelo contrário, a inferência é baseada em amostragem repetida dentro da mesma amostra, e por isto que esta escola pode ser chamada de reamostragem. A reamostragem não emergiu sem um contexto. Os métodos de amostragem são tidos como simulações de Monte Carlo, em que pesquisadores “inventam” dados e tiram conclusões baseados em muitas possibilidades de cenários (YU, 2003).

Reamostragem refere-se ao uso de dados observados ou um mecanismo de geração (como uma matriz) para produzir amostras hipotéticas, cujos resultados podem ser analisados. O termo métodos de computação intensiva (*computer-intensive methods*) também são frequentemente usados para se referir a técnicas como estas (SIMON, 1997, p. 2).

Os métodos de reamostragem podem ser genericamente descritos como técnicas de geração que requerem amostragem de variáveis aleatórias em mais de uma etapa. Tipicamente, elas consistem de dois passos, com o primeiro fornecendo um valor amostrado a partir de uma distribuição aproximada e, o segundo é um tipo de mecanismo de correção que possui a função de redirecionar a amostra de modo a torná-la (pelo menos aproximadamente) representativa da distribuição de interesse. O mecanismo de correção é geralmente, mas não necessariamente, estocástico. No que se segue, a densidade de interesse será denotada por  $\pi$  e a densidade auxiliar denotada por  $q$ . Exceto quando indicado, os valores  $x$  gerados podem ser escalares, vetores, matrizes ou arrays<sup>1</sup> de maiores dimensões (GAMERMAM; LOPES, 2006, p. 25).

Atualmente as principais técnicas de reamostragem são os testes de permutação (*permutation test*), validação cruzada (*cross-validation*), *Jackknife* e *Bootstrap*. Embora hoje em dia elas sejam unificadas sob um tema comum, é importante ressaltar que estas técnicas foram desenvolvidas por diferentes pesquisadores em diferentes períodos de tempo e para diferentes finalidades (YU, 2003).

---

<sup>1</sup>Arrays são vetores nos quais são armazenados atributos adicionais que são as dimensões e opcionalmente os nomes destas dimensões. Uma matriz é um array bidimensional.

## 2.5 INFERÊNCIA BAYESIANA

A informação que se tem sobre uma quantidade de interesse  $\theta$  é fundamental na Estatística. O verdadeiro valor de  $\theta$  é desconhecido e a ideia é reduzir esse desconhecimento. Além disso, a intensidade da incerteza a respeito de  $\theta$  pode assumir diferentes graus. Segundo o ponto de vista Bayesiano, os diferentes graus de incerteza são representados por modelos probabilísticos para  $\theta$ . Neste contexto, diferentes pesquisadores têm diferentes graus de incerteza sobre  $\theta$ . Desta forma, não existe distinção entre quantidades observáveis e os parâmetros de um modelo estatístico, sendo todos considerados quantidades aleatórias (EHLERS, 2011, p. 1).

A inferência Bayesiana é fundamentada sobre o Teorema de Bayes, que segundo Paulino et al. (2003, p. 10), é considerado por muitos como um dos poucos resultados da matemática que se propõe a caracterizar a aprendizagem com a experiência, ou seja, a modificação da atitude inicial em relação aos “antecedentes”, “causas”, “hipóteses” ou “estados” depois de ter a informação adicional de que certo acontecimento ou acontecimentos se realizaram.

### 2.5.1 TEOREMA DE BAYES

O teorema de Bayes pode ser explicado conforme Vismara (2006, p. 83-84):

Considere uma quantidade de interesse desconhecida  $\theta$ . A informação disponível sobre  $\theta$ , resumida probabilisticamente através da distribuição ou função de probabilidade  $p(\theta)$ , pode ser aumentada observando-se uma variável aleatória  $X$  relacionada com  $\theta$ . A distribuição amostral  $p(x|\theta)$  define esta relação. Desta forma, a função de probabilidade conjunta  $p(\theta,x)$ , pode ser escrita como o produto entre a distribuição amostral  $p(x|\theta)$  e a distribuição *a priori*  $p(\theta)$ , respectivamente,  $p(\theta,x)=p(x|\theta)p(\theta)$

Assim, a ideia de que após observar  $X=x$  aumenta-se a quantidade de informação sobre  $\theta$ , cujo teorema de Bayes atua como uma regra de atualização quantificando este aumento e produzindo a densidade *a priori*  $p(x|\theta)$ .

$$p(x|\theta) = \frac{p(\theta,x)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x)}$$

onde:  $p(x)=\sum_{\theta}p(x|\theta)p(\theta)$ , é o somatório sobre todos os valores possíveis de  $\theta$  (ou para o caso contínuo  $\int p(x|\theta)p(\theta),d\theta$ ). Observa-se que  $1/p(x)$ , independente de  $\theta$  e funciona como uma constante normalizadora de  $p(\theta|x)$ .

Para um valor fixo de  $x$ , a função de  $L(\theta;x):=p(x|\theta)$  fornece a verossimilhança de cada um dos possíveis valores de  $\theta$ . A priori e a verossimilhança combinadas levam a densidade *a posteriori* de  $\theta$ ,  $p(\theta|x)$ . Desta forma, escreve-se o teorema de Bayes como

$$p(\theta|x) \propto L(\theta;x)p(\theta)$$

ou seja, a densidade *a posteriori* é proporcional a verossimilhança vezes a densidade *a priori*.

É fácil de verificar se *a priori* ou a verossimilhança serem pequenas *a posteriori* de um conjunto particular de  $\theta$  também será pequena e caso for atribuída o valor 0 para as mesmas, a probabilidade *a posteriori* será 0. A constante normalizadora da posteriori é facilmente recuperada pois

$$p(\theta|x) = kp(\theta)(x|\theta)$$

onde

$$k^{-1} = \int p(\theta)p(x|\theta), d\theta = p(\theta)$$

é chamada de distribuição *preditiva*, ou seja, é a distribuição esperada para a observação  $x$  dado  $\theta$ .

Observando-se  $x_1, x_2, \dots, x_n$  independentes dado  $\theta$  e relacionadas a  $\theta$  através de  $p_i(x_i | \theta)$ , segue que:

$$p(\theta | x_n, x_{n-1}, \dots, x_1) \propto \left[ \prod_{i=1}^n L_i(\theta; x_i) \right] \theta.$$

Ou seja, os conceitos de *priori* e *posteriori* são relativos a observação que está sendo considerada no momento. Por exemplo,  $p(\theta | x_1)$  é a *posteriori* de  $\theta$  em relação a  $X_1$  já observado, mas é a *priori* de  $\theta$  em relação a  $X_2$  ainda não observada. Após observar  $X_2=x_2$  uma nova posteriori (relativa a  $X_1=x_1$  e  $X_2=x_2$ ) é obtida aplicando-se novamente o teorema de Bayes. Portanto, a ordem em que as observações são processadas é irrelevante.

### 2.5.2 VEROSSIMILHANÇA

A função de verossimilhança tem papel fundamental, tanto na inferência clássica quanto na inferência Bayesiana como veículo portador da informação fornecida pela amostra. O princípio da verossimilhança postula que toda a informação fornecida pela amostra está contida na função de verossimilhança (PAULINO et al., 2003, p. 33). O princípio da verossimilhança afirma que ao se realizar uma inferência sobre um parâmetro  $\theta$  só importa aquilo que realmente foi observado e não aquilo que “poderia” ter sido observado (EHLERS, 2011, p. 12).

Um conceito importante de verossimilhança é a Lei da Verossimilhança que pode ser

explicada da seguinte forma:

Dada uma variável aleatória  $X$ , cujo comportamento pode ser explicado por duas hipóteses, sendo que:

- A hipótese A implica que  $X=x$  seria observada com probabilidade  $p_A(x)$
- A hipótese B implica que  $X=x$  seria observado com probabilidade  $p_B(x)$

A Lei da Verossimilhança afirma que a observação  $X=x$  é uma evidência que favorece a hipótese A sobre a hipótese B se e somente se

$$p_A(x) > p_B(x)$$

Mais além, a Lei da verossimilhança implica que a Razão de Verossimilhança

$$\frac{p_A(x)}{p_B(x)}$$

mede a força de evidência da hipótese A sobre a hipótese B (BATISTA, 2009, p. 2).

Segundo Batista (2009, p. 5) existe uma diferença entre probabilidade e verossimilhança; quando numa dada distribuição de probabilidade a observação é fixa e o parâmetro é variável não se tem mais uma função de probabilidade ou densidade de probabilidade e sim uma função de verossimilhança. A função de verossimilhança pode ser aplicada a qualquer distribuição de probabilidade e indica a verossimilhança de uma dada hipótese, no qual é utilizada a notação

$$L\{A|X=x\}$$

Quando a amostra é composta de observações independentes e assumindo-se como verdadeira uma dada hipótese A, a probabilidade de se obter uma amostra  $X_n(X_n = x_1, x_2, \dots, x_n)$ , composta de n observações independentes de  $X(x_1, x_2, \dots, x_n)$  é igual ao produto das probabilidades das observações individuais, dado a hipótese A, ou seja:

$$P(X_n|A) = P(X = x_1|A) \cdot P(X = x_2|A) \cdot \dots \cdot P(X = x_n|A)$$

Este mesmo princípio é aplicado a verossimilhança, sendo assim, a função de verossimilhança de uma amostra composta de observações independentes será o produto das funções de verossimilhança das observações individuais:

$$L\{A|X_n\} = L\{A|X = x_1\} \cdot L\{A|X = x_2\} \cdot \dots \cdot L\{A|X = x_n\}$$

$$L\{A|X_n\} = \prod_{i=1}^n L\{A|X = x_i\}$$

Como o produto das verossimilhanças de cada observação tende a zero quando o tamanho da amostra cresce, utiliza-se a função de log-verossimilhança negativa, que consiste em aplicar a operação logaritmo neperiano e a transformação do sinal:

$$\mathbf{L} = -\log [\{A|X\}]$$

A estimação dos parâmetros de um modelo por verossimilhança é feita através do Método da Máxima Verossimilhança (Maximum Likelihood Estimation). O método consiste em estimar os parâmetros de um modelo utilizando as estimativas que tornam máximo o valor da função de verossimilhança, o que equivale a encontrar o valor para o parâmetro que torne mínimo a função da log-verossimilhança negativa (BATISTA, 2009, p. 9).

### 2.5.3 MÉTODOS DE MONTE CARLO VIA CADEIAS DE MARKOV - MCMC

Os métodos de Monte Carlo via Cadeias de Markov constituem uma alternativa a outros métodos de reamostragem para o ajuste de modelos bayesianos complexos. A ideia chave, assim como nos outros métodos, é a utilização de amostras da distribuição posterior para calcular estimativas amostrais de características desta distribuição através da utilização de processos iterativos a partir da construção de cadeias de Markov (EHLERS, 2011, p. 63).

A utilização de cadeias de Markov são importantes na estatística Bayesiana porque geralmente é possível construir uma cadeia de Markov, ou seja, definir o seu núcleo de transição, de tal maneira que a distribuição alvo é a distribuição a posteriori conjunta de todos os parâmetros desconhecidos no modelo Bayesiano de interesse. Mesmo para modelos multidimensionais em que é impraticável extrair amostras diretamente do conjunto posterior, frequentemente é simples definir um núcleo de transição que amostra condicionalmente, dado a existência de uma amostra do conjunto posterior (COWLES, 2013, p. 124).

Além disso, os métodos de simulação ditos “diretos” não podem ser utilizados em todos os casos principalmente porque se referem a distribuições unidimensionais e focados no cálculo eficaz de integrais específicas e não podem ser utilizados para obter amostras de qualquer distribuição de interesse. Este problema pode ser resolvido utilizando técnicas de simulação baseadas em cadeias de Markov devido as suas propriedades de generalidade e flexibilidade (NTZOUFRAS, 2009, p. 35-36).

Estas características e o avanço massivo dos recursos de computação, tornaram as

técnicas de MCMC populares desde os anos 1990, permitindo que pesquisadores quantitativos pudessem utilizar modelos altamente complicados e estimar as distribuições posteriores correspondentes com precisão, contribuindo grandemente para o desenvolvimento e propagação da teoria Bayesiana (NTZOUFRAS, 2009, p. 35-36).

### 3 METODOLOGIA

#### 3.1 ÁREA DE ESTUDO E COLETA DE DADOS

O conjunto de dados<sup>1</sup> foi obtido em área pertencente à Companhia Vale do Rio Doce, recuperada com espécies nativas, localizada ao norte do Espírito Santo, predominantemente, no município de Linhares, com pequena abrangência nos de Sooretama e Jaguaré. Situam-se geograficamente entre os paralelos 19°06' - 19°18'S e entre os meridianos 39°45' 40°19'W, numa área total de 22 mil hectares.

A escolha dos indivíduos arbóreos seguiu um processo de amostragem estratificada, sendo cada classe de diâmetro definida como um estrato. A amostragem foi uniforme nos estratos, sendo que cada classe de diâmetro foi contemplada com dois indivíduos. A escolha das espécies foi baseada em dois critérios, sendo primeiramente a ampla distribuição geográfica e de forma secundária espécies que possuíam, segundo a literatura, densidades básicas bem distintas. Desta forma, foram coletados 200 indivíduos de dez espécies em dez classes diamétricas. As espécies amostradas foram: *Caesalpinia ferrea*, *Pteropcarpus violaceus*, *Zeyhera tuberculosa*, *Schizolobium parayba*, *Pterygota brasiliensis*, *Cordia trichotoma*, *Astronium graveolens*, *Joannesia princeps*, *Cariniana legalis* e *Tabebuia serritifolia*.

Foram mensurados a altura total (m) e cap (cm) com a árvore ainda em pé. Posteriormente ao abate foram mensurados a altura total (m) e altura comercial (m) e realizada a cubagem rigorosa por Smalian dos galhos e tronco em unidades de decímetro cúbico (dm<sup>3</sup>).

#### 3.2 SIMULAÇÃO E REAMOSTRAGEM

Em termos gerais, bootstrap é um método que usa a amostra fornecida para criar uma nova distribuição, denominada distribuição bootstrap, que se aproxima da distribuição de amostragem para a média da amostra (ou qualquer outra estatística). A ideia por trás

---

<sup>1</sup>A coleta de dados foi realizada por Edgar de Souza Vismara para a execução da sua dissertação de mestrado, o qual forneceu gentilmente os mesmos para a realização deste trabalho.

do bootstrap é que se a amostra for representativa da população original, então a média da distribuição bootstrap será aproximadamente como a média da distribuição amostrada, ou seja, tem a mesma forma e amplitude. Entretanto, a média da distribuição bootstrap será próxima a da distribuição amostral, mas não necessariamente da distribuição original (CHIHARA; HESTERBERG, 2011, p. 100).

A partir dos dados coletados foi realizada a reamostragem de parcelas através da função `sample` presente no pacote base do software R. As árvores foram reamostradas gerando-se 20 parcelas de 30 árvores para cada cenário de estrutura florestal em relação ao diâmetro. A determinação dos diferentes cenários se deu através da inserção do argumento `prob` na função `sample`. A função `prob` controla a probabilidade com que as árvores são amostradas no processo de reamostragem. Neste trabalho, reamostrou-se as árvores com: probabilidade inversamente proporcional ao dap (árvores de menor tamanho tem maior probabilidade de compor a amostra); equiprobabilidade (todas as árvores tem a mesma probabilidade) e diretamente proporcional ao dap (a reamostragem favorece as árvores de maior tamanho). Desta forma, os cenários simulados foram:

1. floresta em estado de regeneração (assimetria a direita);
2. floresta com distribuição diamétrica normal (simetria);
3. floresta com regeneração deficiente (assimétrica a esquerda).

### 3.3 AJUSTE DOS MODELOS BIOMÉTRICOS

Para posterior estimativa do volume por classe diamétrica foi necessário o ajuste de equações hipsométricas e volumétricas. Para isto, quatro modelos de altura e três de volume foram ajustados pelo método da máxima verossimilhança, através da função `mle2` presente no pacote `bbmle` do software R. A relação dos modelos biométricos ajustados é mostrada na tabela 1.

**Tabela 1: Relação dos modelos biométricos de altura e volume com casca ajustados por máxima verossimilhança**

| Modelos                     | Forma   |
|-----------------------------|---|
| Polinômio do 1 <sup>o</sup> | $h_i = \beta_0 + \beta_1 d_i$                   |
| Stoffels                    | $h_i = \beta_0 d_i^{\beta_1}$                   |
| Weibull                     | $h_i = \beta_0 [1 - e(-\beta_1 d_i^{\beta_2})]$ |
| Richards                    | $h_i = \beta_0 [1 - e(-\beta_1 d_i)]^{\beta_2}$ |
| Spurr                       | $v_i = \beta_0 (d_i^2 h_i)^{\beta_1}$           |
| Schumacher-Hall             | $v_i = \beta_0 d_i^{\beta_1} h_i^{\beta_2}$     |
| Ogaya                       | $v_i = d_i^2 (\beta_0 + \beta_1 h_i)$           |

onde:  $h$  é a altura,  $v$  é o volume,  $d$  é diâmetro a 1,30m e  $i$  é o índice da árvore.

### 3.4 COMPARAÇÃO DO AJUSTE E SELEÇÃO DOS MODELOS

Foram utilizados dois critérios de informação para a seleção do melhor modelo: (1) Akaike Information Criterion (AIC) e (2) Bayesian Information Criterion (BIC).

Akaike (1973) propôs o uso da distância de Kullback-Leibler como uma base fundamental para a seleção de modelos, encontrando uma relação entre esta distância relativa esperada e a log-verossimilhança maximizada, definindo aquilo que ele chamou como “*um critério de informação*” que pode ser expressa pela equação (BURNHAM; ANDERSON, 2002, p. 61):

$$AIC = -2\log(L(\hat{\theta}|y)) + 2K$$

onde:  $L(\hat{\theta}|y)$  é a o valor da log-verossimilhança em seu ponto máximo e  $K$  é o número de parâmetros do modelo.

Burnham e Anderson (2002, p. 70-71) recomendam a utilização da tabela 2 para a interpretação dos resultados de AIC na seleção de modelos:

**Tabela 2: Interpretação dos valores de AIC para a seleção de modelos**

| $\Delta_i$ | Nível de suporte empírico do modelo $i$ |
|------------|---|
| 0-2        | Substancial                             |
| 4-7        | Consideravelmente menor                 |
| > 10       | Nenhum                                  |

Onde:  $\Delta_i$  é a diferença entre o AIC do modelo  $i$  e o AIC mínimo.

O BIC é baseado no critério originalmente introduzido por Schwarz (1978), no qual

há a penalização do desvio  $D(\hat{\theta}_m, m)$  igual a  $\log n$  para cada parâmetro estimado do modelo. A estimativa do BIC para um modelo  $m$  é dada por:

$$\text{BIC}(m) = D(\hat{\theta}_m, m) + d_m \log n$$

onde:  $D(\hat{\theta}_m, m)$  é o desvio medido do modelo  $m$ ,  $d_m$  é o número de parâmetros e  $n$  é o tamanho da amostra.

A seleção de modelos através do BIC se dá através da seguinte regra: se a diferença de BIC entre os dois modelos é menor que 2, então não se pode discriminar entre os modelos; se a diferença estiver entre 2 e 6,6, indica que há evidência positiva; se a diferença estiver entre 6,6 e 10, há forte evidência; e se maior que 10, há evidência muito forte em favor do modelo com o menor BIC (NTZOUFRAS, 2009, p. 425).

### 3.5 AJUSTE DA DISTRIBUIÇÃO WEIBULL

A distribuição Weibull foi proposta originalmente por Fisher e Tippett, em 1928, tendo sido desenvolvida independentemente pelo físico sueco Wallodi Weibull, em 1939, no estudo de resistência dos materiais (BATISTA, 1986, p. 14). Segundo o mesmo autor (p. 15), a função de densidade probabilística é apresentada em trabalhos de distribuição de diâmetros como:

$$f(x) = \begin{cases} \frac{c}{b} \left(\frac{x-a}{b}\right)^{c-1} e^{-\left(\frac{x-a}{b}\right)^c}, & \text{para } x \geq a \\ 0, & \text{para } x < a \end{cases}$$

onde:  $x$  é a variável aleatória, e  $a \geq 0$ ,  $b > 0$  e  $c > 0$  os parâmetros da distribuição.

A função de densidade acumulada é dada por:

$$f(x) = 1 - e^{-\left(\frac{x-a}{b}\right)^c}$$

Uma das razões que tornam a distribuição Weibull aplicável a muitas situações é a íntima relação entre os seus parâmetros e a sua forma. O parâmetro  $a$  é chamado de parâmetro de locação, pois controla a posição da curva sob o eixo das abcissas. O parâmetro de escala ( $b$ ) controla a dimensão que a curva assume, dado uma forma constante, quanto maior o valor deste parâmetro mais dispersa é a curva. A grande variedade de formas que a distribuição Weibull pode assumir é dado pelo parâmetro de forma ( $c$ ), quando  $c \leq 1$  a distribuição assume a forma exponencial negativa. Para valores de  $c$  entre 1 e 3,6 a distribuição torna-se positivamente assimétrica, e assume a forma simétrica quando o parâmetro equivale a 3,6. A partir de

3,6 a distribuição torna-se negativamente assimétrica e tendendo a uma forma extremamente leptocúrtica (BATISTA, 1989, p. 16).

Para o ajuste por inferência bayesiana foram determinadas 4 *prioris*, informativas e não informativas, sendo elas:

1. Verossimilhança;
2.  $a, b, c \sim \text{Gama}(1,5; 5)$ , Totalmente informativa para  $a, b, c$ ;
3.  $a, b, c \sim \text{Gama}(5; 1,5)$ , Totalmente informativa para  $a, b, c$ ;
4.  $a, b, c \sim \text{Gama}(5; 5)$ , Totalmente informativa para  $a, b, c$ ;
5.  $a, b, c \sim \text{Gama}(10^{-4}, 10^{-4})$ , Não informativa para  $a, b, c$ .

O ajuste da distribuição se deu através da função LaplacesDemon presente no pacote LaplacesDemon. Fornecido os dados, o modelo e os valores iniciais, a função LaplacesDemon maximiza o logaritmo da densidade *a posteriori* conjunta não normalizada com Monte Carlo via Cadeias de Markov (MCMC) e fornece amostras das distribuições marginais a posteriori, desvio e outras variáveis monitoradas. O algoritmo de otimização utilizado foi *Metropolis-within-Gibbs* (MWG), sendo utilizadas 3 cadeias de Markov e 100000 iterações iniciais.

### 3.6 VERIFICAÇÃO DA CONVERGÊNCIA DAS CADEIAS DE MARKOV

A verificação de convergência das cadeias foi feita através do Diagnóstico de Convergência de Geweke e do Diagnóstico de Convergência BMK.

O Diagnóstico de Convergência de Geweke compara a localização do parâmetro amostrado em dois diferentes intervalos de tempo da cadeia. Se a média dos valores dos parâmetros nos dois intervalos de tempos são próximas uma da outra, então pode se assumir que as duas partes da cadeia tem localização similar no espaço de estado e se assume que as duas amostras fazem parte da mesma distribuição. Normalmente é comparada a última parte da cadeia (dada como convergida), em relação a um intervalo menor no início da cadeia (SAHLIN, 2011, p. 11).

O diagnóstico sugere que se as cadeias convergiram no tempo  $n_0$ , ou seja, se as duas amostras foram extraídas da cadeia com distribuição estacionária as duas médias  $\bar{\theta}_A$  e  $\bar{\theta}_B$  devem ser iguais e a estatística Geweke tem assintoticamente uma distribuição normal padronizada.

Isto é:

$$Z_n = \frac{(\bar{\theta}_A - \bar{\theta}_B)}{\sqrt{\frac{1}{n_A} \bar{S}_\theta^A + \frac{1}{n - n^* + 1} \bar{S}_\theta^B(0)}} \rightarrow N(0, 1) \quad n \rightarrow \infty$$

onde:  $Z_n$  é a estatística Geweke,  $\bar{\theta}_A$  e  $\bar{\theta}_B$  são as médias para dois períodos de tempo diferentes,  $\bar{S}_\theta^A/n_A$  e  $\bar{S}_\theta^B/(n - n^* + 1)$  são as variâncias assintóticas de  $\bar{\theta}_A$  e  $\bar{\theta}_B$ .

A verificação pode ser feita utilizando a função Geweke.Diagnostic do pacote LaplacesDemon, que se baseia na segunda amostra posterior para gerar a estatística Geweke. Caso a cadeia não tenha atingido a convergência a segunda amostra posterior é perdida e não se pode calcular a estatística.

O Diagnóstico de Convergência BMK é baseado na integral de Hellinger (1909) sendo utilizado para quantificar a dissimilaridade entre duas distribuições de probabilidade. Tradicionalmente, a distância de Hellinger é limitada pelo intervalo 0 e 1 ou 0 e  $\sqrt{2}$ . Um valor alto da distância de Hellinger está associada a uma maior dissimilaridade entre as distribuições. A convergência é assumida quando a distância estiver abaixo de um limiar, indicando que as amostras posteriores são similares entre dois lotes consecutivos. Para fins de entendimento, dada uma matriz com 1000 amostras posteriores de três cadeias: beta[1], beta[2] e beta[2], dividida em 10 lotes, sendo cada coluna nomeada como 100, 200, ..., 900. A distância de Hellinger para a cadeia beta[1] em 100 é dada pela distância entre dois lotes: amostras 1-100 e 101:200.

Além deste dois critérios também foi analisado o Tamanho Efetivo da Amostra (Effective Sample Size - ESS) e o Monte Carlo Standart Error (MCSE) como suplemento para a verificação da convergência das cadeias.

De forma prática o ESS responde a pergunta: “*Quanta informação temos atualmente sobre este parâmetro?*”. Quando a distribuição posterior for amostrada com MCMC, geralmente há autocorrelação entre as amostras. Mais autocorrelação é associada com menor perda de informação, porque a informação nas amostras é autocorrelacionada, ou de outra forma, sucessivas amostras não são independentes a partir de amostras anteriores (SAHLIN, 2011, p. 14). A equação do ESS é dada por:

$$ESS(\theta) = \frac{S}{1 + 2 \sum_{k=1}^{\infty} \rho_k(\theta)}$$

onde:  $ESS$  é o Effective Sample Size,  $S$  é o número de amostras posteriores,  $\rho_k$  é a autocorrelação no lag  $k$  e  $\theta$  e o vetor de amostras marginais posteriores.

Os métodos de MCMC, como já abordados, são processos estatísticos cujo objetivo é

a estimação de características da distribuição alvo. Entretanto, de alguma forma é necessário ter o conhecimento da acurácia desta estimativa. O MCSE é um dos métodos que permite esta avaliação. De forma geral, o MCSE é uma medida da inacurácia das amostras de Monte Carlo em relação a expectativa de amostras posteriores, sendo essencialmente um desvio padrão em torno da média da amostras posteriores devido a incerteza associada a utilização de algoritmos MCMC ou processos de Monte Carlo (FLEGAL et al., 2008, p. 1).

O MCSE é considerado satisfatório para cada distribuição alvo se for inferior a 6,27% do desvio padrão da distribuição alvo. Isto permite que a média verdadeira esteja dentro dos 5% da área sobre a distribuição Gaussiana em torno da média estimada (STATISTICAT LLC, 2013, p. 187). Além disso, o MCSE é utilizado como uma regra de parada da simulação dos métodos de MCMC devido as suas propriedades de fácil implementação e por fornecer uma abordagem teoricamente justificada (FLEGAL et al., 2008, p. 10)

### 3.7 AVALIAÇÃO DO AJUSTE DA DISTRIBUIÇÃO WEIBULL

A avaliação do ajuste da distribuição Weibull foi feita analisando o Deviance Information Criterion (*DIC*) e aplicado o fator de Bayes.

O *DIC* é estimado pela expressão:

$$DIC = 2\widehat{D}_{avg}^{pred}(y) - D_{\widehat{\theta}}(y)$$

A expressão  $\widehat{D}_{avg}^{pred}(y)$  denota a discrepância média estimada sobre a distribuição posterior e  $D_{\widehat{\theta}}(y)$  refere-se a discrepância entre os dados e o modelo. O *DIC* também pode ser escrito como  $D + p_D$ , onde  $D$  é os desvio, representado por  $-2\log(y|\Theta)$  e  $p_D$  é um parâmetro que descreve a complexidade do modelo (número efetivo de parâmetros), descrito como a metade da variância de  $D$ , ou,  $var(D)/2$  (GELMAN et al. 2004, p. 180-183). O valor de *DIC* para cada modelo testado está presente no resumo do *output* da função LaplacesDemon.

O fator de Bayes foi introduzido por Harold Jeffreys como uma alternativa bayesiana para o teste de hipóteses da inferência clássica para indicar qual modelo se ajusta melhor aos dados. Segundo Gelman et al. (2004, p. 184) o fator de Bayes avalia a relação entre a verossimilhança marginal de um modelo sobre a verossimilhança marginal de um segundo modelo, ou seja:

$$fator\ de\ Bayes(H_2; H_1) = \frac{p(y|H_2)}{p(y|H_1)} = \frac{\int p(\theta_2|H_2)p(y|\theta_2, H_2)d\theta_2}{\int p(\theta_1|H_1)p(y|\theta_1, H_1)d\theta_1}$$

onde:  $y$  são os dados,  $\theta_i$  o parâmetro e  $H_i$  o modelo.

A interpretação do fator de Bayes é dada pela tabela 3. A verificação do fator de Bayes para os diferentes ajustes da distribuição Weibull foi realizada através da função BayesFactor do pacote LaplacesDemon.

**Tabela 3: Interpretação do fator de Bayes**

| fator de Bayes         | Força de evidência        |
|------------------------|---------------------------|
| $-\infty < B \leq 0,1$ | Evidência forte contra    |
| $0,1 < B \leq (1/3)$   | Evidência moderada contra |
| $(1/3) < B < 1$        | Evidência fraca contra    |
| $1 \leq B < 3$         | Evidência fraca para      |
| $3 \leq B < 10$        | Evidência moderada para   |
| $10 \leq B < \infty$   | Evidência forte para      |

### 3.8 ESTIMATIVA DO NÚMERO DE ÁRVORES E DO VOLUME POR CLASSE DIAMÉTRICA

A partir da função de densidade acumulada da distribuição Weibull foi calculado o número de árvores por classe diamétrica em duas etapas. Na primeira etapa foi calculada a estimativa do número total de árvores dividindo-se o número de árvores observadas  $(1420)^2$  em 1 hectare pela probabilidade de ser maior ou igual a 5cm e menor ou igual a 40cm de dap, ou seja,  $P(5 \leq x \leq 40)$ :

$$\hat{N}_t = \frac{N}{P(5 \leq x \leq 40)} = \frac{N}{\left[ e^{-\left(\frac{5-a}{b}\right)^c} - e^{-\left(\frac{40-a}{b}\right)^c} \right]}$$

onde:  $\hat{N}_t$  é o número de árvores estimadas,  $N$  é o número de árvores observadas,  $a, b$ , e  $c$  são os parâmetros da distribuição.

No segundo passo foi calculado o número de árvores por classe diamétrica, dado pela equação:

$$\hat{N}_c = N_t [P(I_i \leq x \leq S_i)] = \hat{N}_t \left[ e^{-\left(\frac{I_i-a}{b}\right)^c} - e^{-\left(\frac{S_i-a}{b}\right)^c} \right]$$

onde:  $\hat{N}_c$  é o número de árvores estimadas por classe,  $I$  é o limite inferior da classe  $i$ ,  $S$  é limite superior da classe  $i$ ,  $a, b$ , e  $c$  são os parâmetros<sup>3</sup> da distribuição.

<sup>2</sup>Número de árvores baseado em estudo realizado em um fragmento de Mata Atlântica localizado no município de Santa Maria de Jetibá, Espírito Santo (IMANÁ-ENCINAS et al., 2012, p. 570).

<sup>3</sup>Foi utilizada a moda como parâmetro para a estimação do número de árvores por hectare

Com base no modelo hipsométrico e de volume que melhor se ajustaram aos dados e utilizando o dap do centro de classe e a estimativa do número de árvores foi calculado o volume por classe diamétrica e o volume para cada cenário.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 CENÁRIO 1

#### 4.1.1 COMPARAÇÃO DO AJUSTE DOS MODELOS BIOMÉTRICOS

A relação dos modelos ajustados, as estimativas da média dos parâmetros e dos critérios de informação AIC e BIC é mostrado na tabela 4, onde os melhores resultados estão destacados em negrito.

**Tabela 4: Comparação dos modelos biométricos ajustados por máxima verossimilhança pelos critérios de informação AIC e BIC.**

| Modelo          | $\beta_0$        | $\beta_1$        | $\beta_2$ | AIC              | BIC              |
|-----------------|------------------|------------------|-----------|------------------|------------------|
| Polinômio do 1º | 3,429376         | 0,546018         | -         | 2683,786         | 2696,977         |
| Stoffels        | 1,782927         | 0,705067         | -         | 2608,090         | 2621,281         |
| Weibull         | 23,535857        | 0,030490         | 1,191465  | 2535,637         | 2553,225         |
| Richards        | 23,965387        | 0,064349         | 1,331715  | <b>2535,570</b>  | <b>2553,158</b>  |
| Spurr           | $1,178801e^{-5}$ | 1,122086         | -         | -1671,635        | -1658,445        |
| Schumacher-Hall | $4,122768e^{-5}$ | 2,082740         | 0,905142  | <b>-1812,787</b> | <b>-1795,199</b> |
| Ogaya           | $1,135770e^{-6}$ | $4,122236e^{-5}$ | -         | -1808,741        | -1795,550        |

Onde:  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  são os coeficientes do modelo, AIC e BIC são o Akaike Information Criterion e Bayesian Information Criterion, respectivamente.

Fonte: O Autor (2015).

Analisando-se os critérios de informação AIC e BIC obtidos com o ajuste dos modelos hipsométricos verifica-se que o modelo de Richards pode ser considerado aquele que melhor se ajustou aos dados. Entretanto, as diferenças dos critérios de informação entre o modelo de Richards e o modelo Weibull é inferior a 2, ou seja, não pode ser feita qualquer distinção entre os dois ajustes em termos de qualidade. O Polinômio do 1º e o modelo de Stoffels apresentaram os piores ajustes, com valores de AIC e BIC superiores a 10, de forma que quando comparados ao modelo de Richards e Weibull não oferecem nenhum suporte empírico para a utilização. Desta forma, para a estimativa da altura por classe diamétrica deu-se preferência ao modelo de

Richards devido ao seu menor valor.

Em relação aos modelos de volume, os menores valores de AIC e BIC encontrados estão relacionados ao modelo de Schumacher-Hall. O modelo de Spurr apresenta para ambos os critérios os maiores valores e devido à diferença superior a 10 em relação ao melhor ajuste não oferece nenhum suporte para a sua utilização. Observa-se que segundo o BIC não pode ser feita qualquer distinção entre os modelos de Schumacher-Hall e o modelo de Ogaya, dada a diferença para este critério ser inferior a 2. Entretanto, a diferença entre os valores para o AIC é superior a 2 a favor do modelo de Schumacher, de modo que este oferece um nível empírico substancial para a sua utilização. Desta forma, para a estimativa do volume por classe diamétrica deu-se preferência ao modelo de Schumacher-Hall

#### 4.1.2 AVALIAÇÃO DA CONVERGÊNCIA DAS CADEIAS DE MARKOV E SELEÇÃO DO MELHOR MODELO

Em relação ao número de iterações necessárias para a convergência das cadeias de Markov, a verossimilhança e todas as *prioris* testadas atingiram a estacionaridade com as 100 mil iterações iniciais propostas. O resumo dos diagnósticos de convergência é mostrado na tabela 5.

Em relação ao Z-score observa-se que todos os valores do desvio padrão estão entre o intervalo de  $\pm 1,96$ , indicando que as cadeias de Markov da verossimilhança e de todas as *prioris* testadas convergiram para a distribuição posterior e que houve similaridade entre cadeias consecutivas.

Analisando-se o MCSE observa-se que todos os valores estão abaixo de 6,27%, valor de referência utilizado pelo pacote LaplacesDemons como regra de parada da atualização das cadeias de Markov. Em relação ao ESS todos os valores estão bem distantes do valor de referência (100) utilizado pelo pacote como regra de parada das atualizações, indicando que há pouca perda de informação devido a maior autocorrelação entre as amostras.

Em relação ao diagnóstico de convergência BMK verifica-se que o valor máximo do teste para a verossimilhança e as *prioris* testadas encontra-se abaixo do valor de referência de 0,5, indicando que as cadeias são similares entre si e, portanto, são estacionárias e convergiram para a distribuição posterior. A tabela 19 do Apêndice A traz o resultado completo deste diagnóstico para o cenário 1.

**Tabela 5: Estimativa dos parâmetros da distribuição Weibull e diagnósticos de convergência das cadeias de Markov para a verossimilhança e *prioris* do cenário 1**

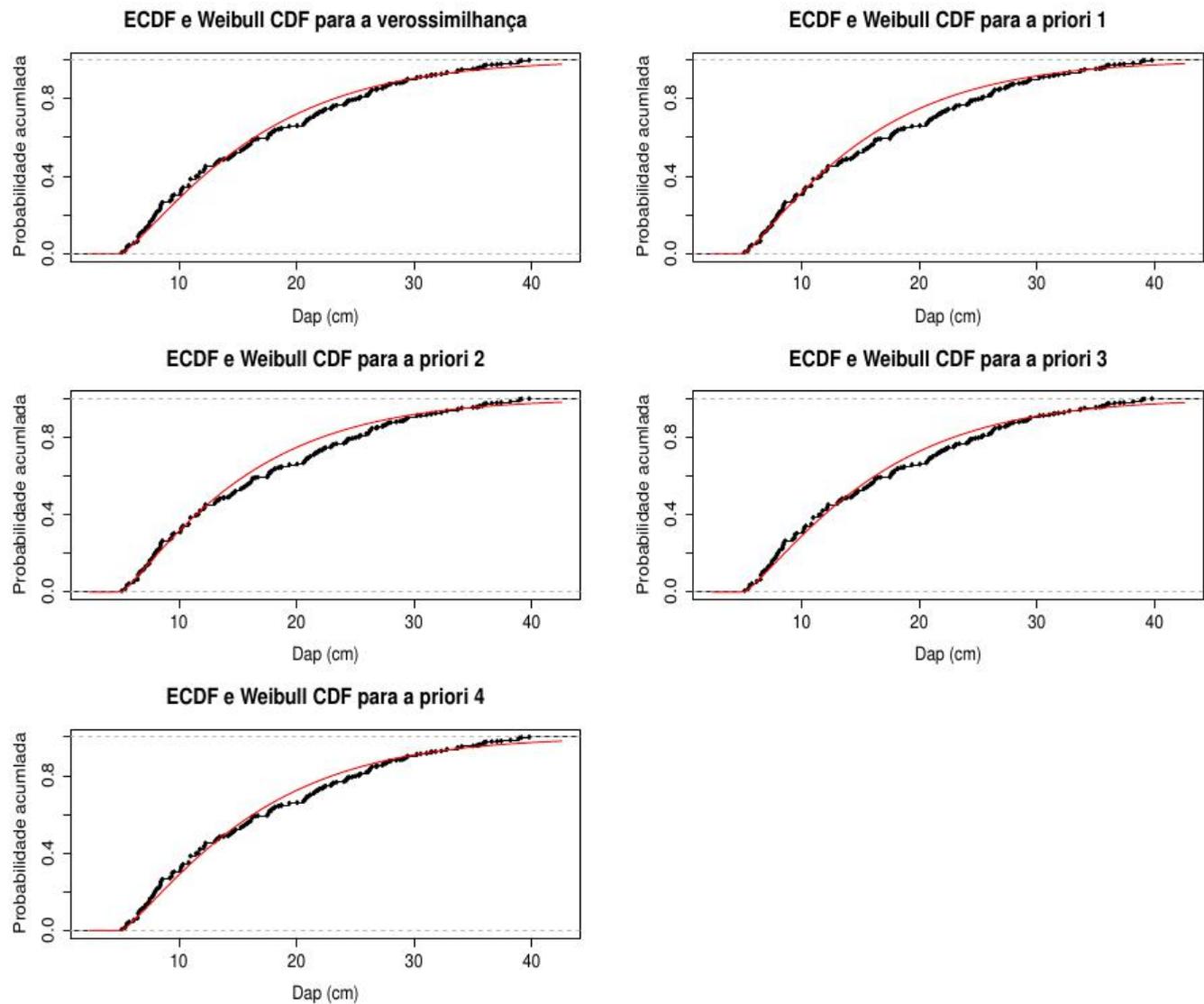
| Modelo          | Parâmetro | Moda      | Diagnóstico |          |          |                    |          |
|-----------------|-----------|-----------|-------------|----------|----------|--------------------|----------|
|                 |           |           | Z-score     | MCSE     | ESS      | BMK <sub>max</sub> | DIC      |
| Verossimilhança | <i>a</i>  | 5,127587  | -1,178153   | 0,002535 | 1106,254 | 0,139905           | 4104,028 |
|                 | <i>b</i>  | 12,134303 | 0,695472    | 0,009022 | 3759,597 | 0,067936           |          |
|                 | <i>c</i>  | 1,200853  | 0,101526    | 0,001859 | 1019,046 | 0,201006           |          |
| <i>Priori 1</i> | <i>a</i>  | 5,137953  | 0,192114    | 0,001729 | 880,469  | 0,186554           | 4115,419 |
|                 | <i>b</i>  | 11,238996 | -1,892134   | 0,009074 | 3326,194 | 0,076360           |          |
|                 | <i>c</i>  | 1,157938  | -1,680640   | 0,001703 | 1116,136 | 0,135766           |          |
| <i>Priori 2</i> | <i>a</i>  | 5,130127  | -0,967181   | 0,001810 | 835,957  | 0,143914           | 4112,990 |
|                 | <i>b</i>  | 11,279424 | 1,233270    | 0,009393 | 3062,846 | 0,083917           |          |
|                 | <i>c</i>  | 1,158232  | 1,629061    | 0,001738 | 1145,777 | 0,138677           |          |
| <i>Priori 3</i> | <i>a</i>  | 5,12972   | 0,226320    | 0,002233 | 1159,201 | 0,135153           | 4103,973 |
|                 | <i>b</i>  | 11,94408  | -0,066268   | 0,007962 | 4084,879 | 0,074360           |          |
|                 | <i>c</i>  | 1,19600   | -0,276517   | 0,001681 | 1320,460 | 0,144694           |          |
| <i>Priori 4</i> | <i>a</i>  | 5,123537  | -0,873678   | 0,002431 | 1172,45  | 0,139173           | 4104,228 |
|                 | <i>b</i>  | 12,038118 | 1,111792    | 0,009531 | 3671,719 | 0,067333           |          |
|                 | <i>c</i>  | 1,198221  | -0,110300   | 0,001874 | 1082,606 | 0,150725           |          |

Onde: *a*, *b*, *c* são os parâmetros do modelo, Z-score é o Diagnóstico de Convergência Geweke, MCSE é o Monte Carlo Standart Error, ESS é o Effective Sample Size, BMK<sub>max</sub> é o valor máximo do diagnóstico BMK e DIC é o Deviance Information Criterion.

Fonte: O Autor (2015).

Quando se compara o valor de DIC, critério que indica o melhor ajuste do modelo aos dados, verifica-se que a distribuição Weibull ajustou-se melhor quando utilizada a *priori* número 3, como pode ser observado através do valor de 4103,973. Entretanto, há uma diferença pouco considerável entre a *priori* 3, a verossimilhança e a *priori* 4. A distribuição Weibull ajustada com as *prioris* 1 e 2 apresentou as maiores discrepâncias em relação ao melhor ajuste, evidenciado pelos maiores valores de DIC.

As condições descritas anteriormente podem ser melhor observadas quando plotadas as distribuições cumulativas empíricas (ECDF) e a distribuição cumulativa Weibull (CDF Weibull) para os ajustes (Figura 1). Observa-se que quanto mais distantes as CDF's (linhas vermelhas) estão da ECDF (pontos interligados) pior é a qualidade de ajuste. Fazendo-se uma análise visual verifica-se que a verossimilhança, a *priori* 3 e *priori* 4 geram as CDF's que mais ficam próximas da ECDF, enquanto que as *priori* 1 e *priori* 2 são as mais distantes.



**Figura 1:** Comparação entre a distribuição cumulativa empírica e distribuição cumulativa Weibull obtida com o ajuste por verossimilhança e inferência Bayesiana para o cenário 1.

Fonte: O Autor (2015).

A tabela 6 mostra os resultados para o fator de Bayes para o cenário 1, onde o modelo da linha é comparado com o modelo da coluna. Comparando-se os valores encontrados com a tabela interpretativa (tabela 3) observa-se que segundo este critério praticamente não há distinção entre os ajustes da distribuição Weibull. Somente quando comparado o ajuste por verossimilhança com a *priori* 1 pode-se dizer que a mesma apresenta um nível de informação substancial para a sua utilização. Os demais modelos que apresentam valor maior que 1 apresentam poucos indícios para serem considerados melhores ajustados. Os valores menores que 1 refletem a mesma condição anterior, porém de modo inverso, ou seja, quando é comparado um modelo com um ajuste fraco com um modelo considerado superior. Desta forma, a verossimilhança se ajustou melhor aos dados mas com pouca superioridade em relação as

*prioris* 1, 2, 3 e 4.

**Tabela 6: Comparação do ajuste dos modelos pelo fator de Bayes para o cenário 1, onde o valor indica a superioridade do modelo da linha sobre o modelo da coluna.**

|                 | Verossimilhança | <i>Priori</i> 1 | <i>Priori</i> 2 | <i>Priori</i> 3 | <i>Priori</i> 4 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Verossimilhança | 1,000           | 3,207           | 2,051           | 1,180           | 1,250           |
| <i>Priori</i> 1 | 0,312           | 1,000           | 0,639           | 0,368           | 0,390           |
| <i>Priori</i> 2 | 0,488           | 1,564           | 1,000           | 0,575           | 0,609           |
| <i>Priori</i> 3 | 0,848           | 2,719           | 1,739           | 1,000           | 1,060           |
| <i>Priori</i> 4 | 0,800           | 2,566           | 1,641           | 0,944           | 1,000           |

Fonte: O Autor (2015).

Quando comparado os dois critérios de seleção de modelos verifica-se que houve uma inversão em relação ao melhor modelo ajustado. Segundo o DIC, o melhor modelo é obtido com a utilização da *priori* 3, enquanto que pelo fator de Bayes o melhor ajuste se dá somente utilizando a verossimilhança. A explicação para isto se deve ao fato de que geralmente o DIC fornece um valor menor quando utilizada uma *priori* informativa, como é o caso da *priori* 3, em comparação a uma *priori* vaga ou a completa ausência de informação a *priori*, como quando utilizada a verossimilhança.

Desta forma, a utilização das *prioris* totalmente informativas não foi eficiente no ajuste do modelo de Weibull, pois os melhores resultados indicados pelo fator de Bayes foram obtidos com a utilização da verossimilhança e da *priori* 4 (não informativa). Mesmo considerando-se o DIC como melhor critério de seleção de modelos a diferença entre os melhores ajustes favorece fracamente a *priori* 3, sendo a diferença numérica tão pequena que pode ser desconsiderada.

#### 4.1.3 ESTIMAÇÃO DO NÚMERO DE ÁRVORES E VOLUME POR HECTARE

A estimativa do número de árvores por hectare por classe diamétrica para o cenário 1 é exibida na tabela 7, onde fica evidente a proximidade entre o número de árvores estimadas pela verossimilhança e pelas *prioris* 3 e 4. As maiores concentrações de árvores nas classes de 5 a 10 cm e de 10 a 15 cm foram estimadas com as *prioris* 1 e 2 que obtiveram os piores resultados segundo os critérios de seleção de modelos. De forma oposta, os melhores ajustes estimaram o maior número de árvores a partir da classe de 15 a 20 cm de diâmetro. Verifica-se também pouca semelhança entre as estimativas e o valor amostrado, onde chega-se a ter quase o dobro (97,38%) de árvores para a classe de 35 a 40 cm em relação a estimativa por verossimilhança.

**Tabela 7: Estimativa do número de árvores por hectare por classe diamétrica para cada modelo no cenário 1.**

| Classe (cm) | Amostrado* | Verossimilhança | <i>Priori 1</i> | <i>Priori 2</i> | <i>Priori 3</i> | <i>Priori 4</i> |
|-------------|------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| [5-10)      | 430,73     | 411,04          | 453,12          | 452,74          | 418,25          | 415,38          |
| [10-15)     | 310,03     | 378,36          | 382,63          | 381,81          | 380,18          | 379,05          |
| [15-20)     | 194,06     | 263,12          | 252,36          | 252,32          | 261,74          | 262,23          |
| [20-25)     | 194,06     | 168,84          | 155,91          | 156,20          | 166,48          | 167,52          |
| [25-30)     | 151,47     | 103,12          | 92,54           | 92,91           | 100,86          | 101,90          |
| [30-35)     | 71,00      | 60,74           | 53,36           | 53,69           | 58,97           | 59,80           |
| [35-40)     | 68,63      | 34,77           | 30,08           | 30,33           | 33,52           | 34,12           |
| Total       | 1420       | 1420            | 1420            | 1420            | 1420            | 1420            |

\* refere-se a uma extrapolação para 1 hectare com base na amostra utilizada para o ajuste.

Fonte: O Autor (2015).

As estimativas de volume por classe diamétrica são mostradas na tabela 8, no qual os piores ajustes foram aqueles que apresentaram os menores volumes. Isto se deve ao fato de que nos piores ajustes há uma maior concentração de árvores nas primeiras classes diamétricas que pouco influenciam no volume total.

Comparando-se a menor estimativa de volume (*priori 1*) com a maior estimativa (verossimilhança), verifica-se uma diferença de 23,87 m<sup>3</sup>/ha, ou seja, de 8,88%. Quando são comparados as estimativas do volume entre os dois melhores modelos ajustados (*priori 3* e verossimilhança) observa-se uma diferença de 5,25 m<sup>3</sup>/ha, ou 1,83%, que pode ser considerada pouco expressiva.

Comparando-se o volume total amostrado com a estimativa por verossimilhança verifica-se uma diferença de 59,26 m<sup>3</sup>/ha (20,18%). Tal fato ocorre devido ao valor amostrado do número de árvores nas classes de 20 a 25 cm e posteriores ser maior do que as estimativas obtidas com os ajustes, sendo estas classes as que influenciam fortemente no volume total.

**Tabela 8: Estimativa do volume por hectare ( $m^3/ha$ ) por classe diamétrica para cada modelo no cenário 1**

| Classe (cm) | Amostrado* | Verossimilhança | Priori 1 | Priori 2 | Priori 3 | Priori 4 |
|-------------|------------|-----------------|----------|----------|----------|----------|
| [5,10)      | 6,72       | 6,28            | 6,92     | 6,91     | 6,39     | 6,34     |
| [10,15)     | 18,98      | 26,06           | 26,35    | 26,29    | 26,18    | 26,10    |
| [15,20)     | 31,32      | 46,54           | 44,63    | 44,62    | 46,29    | 46,38    |
| [20,25)     | 68,78      | 58,53           | 54,04    | 54,14    | 57,71    | 58,07    |
| [25,30)     | 82,40      | 59,87           | 53,73    | 53,94    | 58,56    | 59,16    |
| [30,35)     | 61,48      | 53,36           | 46,88    | 47,17    | 51,80    | 52,54    |
| [35,40)     | 83,31      | 43,09           | 37,27    | 37,58    | 41,53    | 42,28    |
| Total       | 352,98     | 293,72          | 269,82   | 270,67   | 288,46   | 290,87   |

\* refere-se a uma extrapolação para 1 hectare com base na amostra utilizada para o ajuste.

Fonte: O Autor (2015).

## 4.2 CENÁRIO 2

### 4.2.1 COMPARAÇÃO DOS MODELOS BIOMÉTRICOS

A relação dos modelos ajustados, as estimativas da média dos parâmetros e dos critérios de informação AIC e BIC é mostrado na tabela 9, onde os melhores resultados estão destacados em negrito.

**Tabela 9: Comparação dos modelos biométricos ajustados por máxima verossimilhança pelos critérios de informação AIC e BIC.**

| Modelo          | $\beta_0$         | $\beta_1$        | $\beta_2$  | AIC              | BIC              |
|-----------------|-------------------|------------------|------------|------------------|------------------|
| Polinômio do 1º | 4,473831          | 0,483468         | -          | 2683,786         | 2696,977         |
| Stoffels        | 2,095836          | 0,647751         | -          | 2608,090         | 2621,281         |
| Weibull         | 24,434815         | 0,03765343       | 1,08635980 | 2535,637         | 2553,225         |
| Richards        | 24,224489         | 0,056767         | 1,189368   | <b>2535,570</b>  | <b>2553,158</b>  |
| Spurr           | $3,058728e^{-5}$  | 1,032535         | -          | -1398,359        | -1385,168        |
| Schumacher-Hall | $3,425434e^{-5}$  | 2,164534         | 0,879061   | <b>-1405,646</b> | <b>-1388,058</b> |
| Ogaya           | $-1,320957e^{-5}$ | $4,299118e^{-5}$ | -          | -1390,281        | -1377,090        |

Onde:  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  são os coeficientes do modelo, AIC e BIC são o Akaike Information Criterion e Bayesian Information Criterion, respectivamente.

Fonte: O Autor (2015).

Em relação ao ajuste hipsométrico, os menores valores para os critérios de informação AIC e BIC estão relacionados ao modelo de Richards. Entretanto, a diferença entre os valores para os critérios para os modelos de Richards e Weibull é inferior a 2, de modo que não pode-se

fazer distinção entre os dois ajustes em termos de qualidade. O Polinômio do 1º e o modelo de Stoffels forneceram valores muito superiores aos modelos de Richards e Weibull e desta forma não fornecem qualquer suporte empírico para a sua utilização. Para a estimativa do número de árvores por hectare para o cenário 2 foi utilizado o modelo de Richards apenas pela pequena superioridade numérica dos critérios de informação a seu favor.

Quando analisado os resultados dos critérios de informação para o ajuste volumétrico observa-se que o melhor ajuste foi obtido com o modelo de Schumacher-Hall, sendo que a diferença para ambos os critérios é superior a 2 em favor deste modelo quando comparados aos modelos de Spurr e Ogaya. Desta forma para a estimativa do volume por classe diamétrica para o cenário 2 foi utilizado o modelo de Schumacher-Hall

#### 4.2.2 AVALIAÇÃO DA CONVERGÊNCIA DAS CADEIAS DE MARKOV E SELEÇÃO DO MELHOR MODELO

Em relação ao número de iterações necessárias para a convergência das cadeias, a verossimilhança e todas as *prioris* atingiram a estacionaridade com as 100 mil iterações iniciais propostas. O resumo dos diagnósticos utilizados para testar a convergência das cadeias de Markov é mostrado na tabela 10.

Analisando o diagnóstico de Geweke, verifica-se que o teste não foi satisfatório para a *priori* 3, onde os valores para todos os parâmetros da distribuição Weibull ficaram fora do intervalo utilizado como referência ( $\pm 1,96$ ), indicando que não houve convergência para a distribuição posterior. Para os demais ajustes todos os valores estão dentro do intervalo, garantindo a convergência.

Em relação a *priori* 3 é feita uma ressalva; o pacote LaplacesDemons avalia cinco critérios na função Consort, sendo eles: o algoritmo, a taxa de aceitação do algoritmo, MCSE, ESS e a estacionaridade, quando pelo menos um destes cinco critérios não é satisfatório a função faz uma sugestão do que pode ser feito (função Demonic Suggestion), como não foi feita nenhuma sugestão, justamente por todos os valores serem aceitáveis, acredita-se que houve a convergência. Além disso, como o critério de Geweke utiliza duas amostras para retornar o resultado, nas posições de 10 e 50% do número total de cadeias, acredita-se que nestas posições houve grande dissimilaridade entre as mesmas.

Em relação ao MCSE, todos os valores encontram-se abaixo de 6,27%, indicando a estacionaridade das cadeias. Em relação ao ESS todos os valores encontram-se acima de 100, indicando que houve pouca perda de informação devido a alta correlação entre as cadeias amostradas. Avaliando-se o DIC como critério de seleção, percebe-se que a distribuição Weibull

**Tabela 10: Diagnósticos de convergência das cadeias de Markov para a verossimilhança e *prioris* do cenário 2**

| Modelo          | Parâmetro | Moda      | Diagnóstico |          |          |                    | DIC      |
|-----------------|-----------|-----------|-------------|----------|----------|--------------------|----------|
|                 |           |           | Z-score     | MCSE     | ESS      | BMK <sub>max</sub> |          |
| Verossimilhança | <i>a</i>  | 4,466203  | 1,244930    | 0,028787 | 560,646  | 0,124139           | 4360,224 |
|                 | <i>b</i>  | 17,826529 | -1,251609   | 0,043052 | 665,798  | 0,121922           |          |
|                 | <i>c</i>  | 1,641949  | -1,144146   | 0,006383 | 495,008  | 0,170466           |          |
| <i>Priori 1</i> | <i>a</i>  | 5,059776  | -1,074617   | 0,005624 | 1524,769 | 0,112639           | 4377,848 |
|                 | <i>b</i>  | 15,849232 | 0,630412    | 0,013712 | 1991,041 | 0,089078           |          |
|                 | <i>c</i>  | 1,456377  | 0,661448    | 0,002794 | 1031,292 | 0,142032           |          |
| <i>Priori 2</i> | <i>a</i>  | 5,014276  | 0,716661    | 0,006224 | 1466,085 | 0,088534           | 4375,274 |
|                 | <i>b</i>  | 16,036666 | -1,277256   | 0,014559 | 2060,689 | 0,093611           |          |
|                 | <i>c</i>  | 1,473118  | -1,277276   | 0,002759 | 1027,663 | 0,168399           |          |
| <i>Priori 3</i> | <i>a</i>  | 4,74470   | 3,076200    | 0,018415 | 739,093  | 0,111959           | 4359,898 |
|                 | <i>b</i>  | 17,20920  | -3,190852   | 0,030485 | 938,248  | 0,101162           |          |
|                 | <i>c</i>  | 1,56593   | -2,782085   | 0,004928 | 591,081  | 0,151881           |          |
| <i>Priori 4</i> | <i>a</i>  | 4,528690  | 1,643215    | 0,031104 | 481,611  | 0,118869           | 4363,219 |
|                 | <i>b</i>  | 17,747614 | -1,563610   | 0,039625 | 717,922  | 0,165407           |          |
|                 | <i>c</i>  | 1,615018  | -1,627609   | 0,005163 | 829,717  | 0,167954           |          |

Onde: *a*, *b*, *c* são os parâmetros do modelo, Z-score é o Diagnóstico de Convergência Geweke, MCSE é o Monte Carlo Standart Error, ESS é o Effective Sample Size, BMK<sub>max</sub> é o valor máximo do diagnóstico BMK e DIC é o Deviance Information Criterion.

Fonte: O Autor (2015).

combinada com a *priori 3* é aquela que apresenta o menor valor, seguida pela verossimilhança. Os piores ajustes estão relacionados ao uso das *prioris 1* e *2*.

Em relação ao diagnóstico BMK, observa-se que os valores máximos para os parâmetros da distribuição Weibull estão abaixo de 0,5 e indicam convergência de todas as cadeias, incluindo a *priori 3*. A tabela 20 do Apêndice A traz o resultado completo deste diagnóstico para o cenário 2.

A análise pelo fator de Bayes é mostrada na tabela 11. Interpretando-se os resultados observa-se que a melhor combinação para o cenário 2 é quando utilizado a *priori 1*, pois esta apresentou os melhores resultados sobre todas as demais (valores acima de 1). O maior contraste é quando a *priori 1* é comparada com a *2*, onde o valor para o fator de Bayes é 17,13, ou seja, uma evidência muito forte para a utilização desta combinação. Também quando comparada com a *priori 3*, a *priori 1* apresenta uma evidência considerada substancial para a sua utilização. Em

relação a verossimilhança e a *priori* 4, a *priori* 1 apresenta um desempenho um pouco melhor, mas o nível de evidência para a sua utilização é fraco.

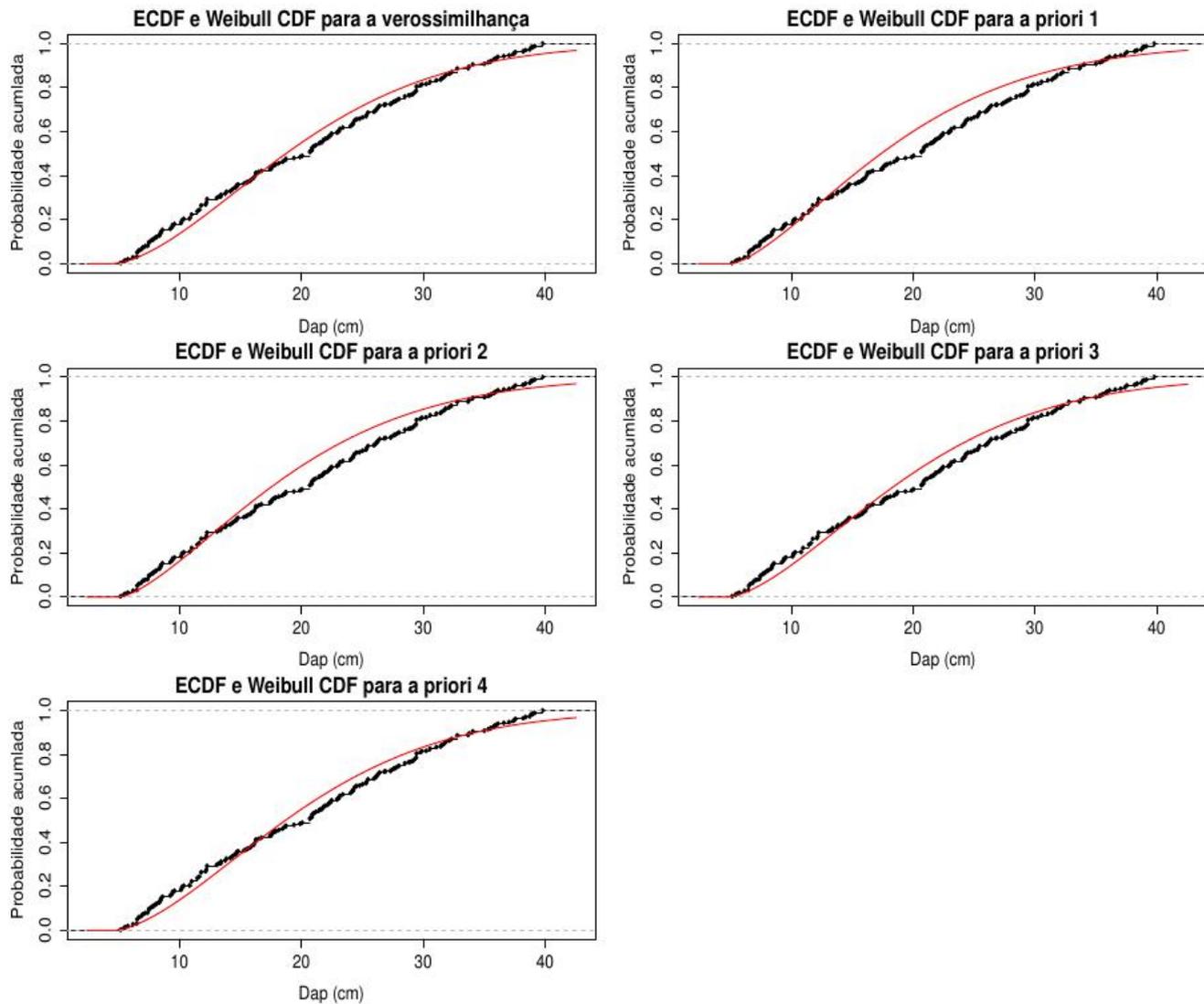
**Tabela 11: Comparação do ajuste dos modelos pelo fator de Bayes para o cenário 2, onde o valor indica a superioridade do modelo da linha sobre o modelo da coluna.**

|                 | Verossimilhança | <i>Priori</i> 1 | <i>Priori</i> 2 | <i>Priori</i> 3 | <i>Priori</i> 4 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Verossimilhança | 1,000           | 0,551           | 9,501           | 1,760           | 0,999           |
| <i>Priori</i> 1 | 1,813           | 1,000           | 17,228          | 3,192           | 1,811           |
| <i>Priori</i> 2 | 0,105           | 0,058           | 1,000           | 0,185           | 0,105           |
| <i>Priori</i> 3 | 0,568           | 0,313           | 5,398           | 1,000           | 0,567           |
| <i>Priori</i> 4 | 1,001           | 0,552           | 9,512           | 1,762           | 1,000           |

Fonte: O Autor (2015).

Comparando-se os dois métodos de seleção de modelos verifica-se uma divergência em relação ao melhor ajuste. Segundo o DIC, o melhor ajuste foi obtido com a *priori* número 3, seguida pela verossimilhança e pela *priori* 4, enquanto que o pior ajuste foi obtido com a *priori* 1, seguida pela 2. Já para o fator de Bayes, o melhor ajuste é dado pela *priori* 1, seguido pela *priori* 4, a verossimilhança e a *priori* 3, enquanto que o pior ajuste foi obtido com a utilização da *priori* 2.

Comparado-se o critério de informação DIC com o fator de Bayes e fazendo-se uma análise gráfica da distribuição cumulativa empírica e das distribuições empíricas da função Weibull ajustadas com a utilização das *prioris* e pela verossimilhança (figura 2) observa-se que uma maior plausibilidade do melhor ajuste quando utilizado o DIC. As menores distâncias entre as curvas acumuladas ajustadas (CDF) e a curva acumulada empírica (CDF) estão relacionadas a *priori* 3 e a verossimilhança, seguida pela *priori* 4 e as maiores distâncias estão relacionadas as *prioris* 3 e 4.



**Figura 2:** Comparação entre a distribuição cumulativa empírica e distribuição cumulativa Weibull obtida com o ajuste por verossimilhança e inferência Bayesiana para o cenário 2.

Fonte: O Autor (2015).

Entretanto, segundo Millar (2009, p. 968) a comparação de modelos usando DIC e fator de Bayes geralmente não levam as mesmas conclusões. O fator de Bayes quantifica o quão bem o modelo *a priori* prediz os dados observados, enquanto que o DIC é projetado para medir o quão bem o modelo *a posteriori* pode prever futuras observações. Além disso, segundo Kass and Raftery (1995, p. 782) o fator de Bayes pode ser sensível a *priori* utilizada. Como o modelo e os dados são os mesmos e o que varia são os níveis de informação inicial, espera-se que as *prioris* informativas utilizadas aumentem a verossimilhança marginal, especialmente no caso da *priori* 3.

#### 4.2.3 ESTIMATIVA DO NÚMERO DE ÁRVORES E VOLUME POR HECTARE

As estimativas para o número de árvores por hectare para cada classe diamétrica e por nível de informação utilizado no ajuste é mostrada na tabela 12.

**Tabela 12: Estimativa do número de árvores por hectare por classe diamétrica para cada modelo no cenário 2.**

| Classe (cm) | Amostrado* | Verossimilhança | <i>Priori 1</i> | <i>Priori 2</i> | <i>Priori 3</i> | <i>Priori 4</i> |
|-------------|------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| [5-10)      | 255,60     | 198,57          | 247,73          | 242,90          | 213,39          | 203,25          |
| [10-15)     | 255,60     | 309,81          | 341,62          | 338,56          | 319,69          | 311,36          |
| [15-20)     | 177,50     | 306,77          | 300,91          | 301,70          | 304,41          | 304,81          |
| [20-25)     | 246,13     | 249,23          | 225,71          | 228,04          | 241,30          | 246,52          |
| [25-30)     | 222,47     | 176,96          | 152,67          | 154,94          | 169,36          | 175,24          |
| [30-35)     | 127,80     | 112,90          | 95,44           | 97,04           | 108,11          | 112,54          |
| [35-40)     | 134,90     | 65,74           | 55,92           | 56,82           | 63,73           | 66,30           |
| Total       | 1420       | 1420            | 1420            | 1420            | 1420            | 1420            |

\* refere-se a uma extrapolação para 1 hectare com base na amostra utilizada para o ajuste.

Fonte: O Autor (2015).

Observa-se que os piores ajustes (considerando o DIC) foram aqueles que estimaram o maior número de árvores por hectare nas duas primeiras classes diamétricas. A diferença entre a estimativa do número de árvores por hectare e por classe diamétrica entre o melhor ajuste (utilizando a *priori 3*) e o pior (utilizando a *priori 1*) é de 34,34; 21,93; 3,50; 15,59; 16,69; 12,67 e 7,81, respectivamente. Já a diferença entre os dois melhores ajustes, obtidos com a *priori 3* e a verossimilhança é de 14,82; 9,88; 2,36; 7,93; 7,61; 4,79; 2,02 árvores por hectare por classe diamétrica.

Novamente observa-se que houve pouca similaridade entre os valores amostrados nas classes diamétricas e os valores estimados. Os melhores ajustes permitem que sejam estimadas mais árvores nas classes de 10 a 15 e 20 a 25 cm enquanto que os valores amostrados se mostram maiores nas demais, destacando-se as classes posteriores a de 20 a 25 cm de diâmetro.

A estimativa do volume por hectare é mostrada na tabela 13. Comparando-se os valores quanto ao nível de informação verifica-se que os piores ajustes foram aqueles que obtiveram as menores somas de volume, justamente por concentrarem mais árvores nas primeiras classes diamétricas que pouco influenciam no volume total. A variação entre o volume total entre o melhor ajuste e o pior é de 34,59 m<sup>3</sup>/ha, enquanto que entre os dois melhores é de 13,37 m<sup>3</sup>/ha. A variação percentual entre o melhor e o pior ajuste é de 8,72% e entre os dois melhores é de 3,1%.

Entretanto, quando comparado o valor amostrado com o melhor ajuste (*priori 3*) a diferença no volume total é de 111,83 m<sup>3</sup>/ha (25,93%), sendo que a diferença de volume na última classe diamétrica é de 89,79 m<sup>3</sup>/ha (111,64%), mais que o dobro da melhor estimativa.

**Tabela 13: Estimativa do volume por hectare (m<sup>3</sup>/ha) por classe diamétrica para cada modelo no cenário 2**

| Classe (cm) | Amostrado* | Verossimilhança | <i>Priori 1</i> | <i>Priori 2</i> | <i>Priori 3</i> | <i>Priori 4</i> |
|-------------|------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| [5,10)      | 4,06       | 2,90            | 3,62            | 3,55            | 3,12            | 2,97            |
| [10,15)     | 16,67      | 20,40           | 22,49           | 22,29           | 21,05           | 20,50           |
| [15,20)     | 28,47      | 52,35           | 51,35           | 51,49           | 51,95           | 52,02           |
| [20,25)     | 85,82      | 84,45           | 76,48           | 77,27           | 81,76           | 83,53           |
| [25,30)     | 126,65     | 101,86          | 87,86           | 89,17           | 97,47           | 100,85          |
| [30,35)     | 111,19     | 99,70           | 84,28           | 85,69           | 95,47           | 99,38           |
| [35,40)     | 170,22     | 82,97           | 70,58           | 71,72           | 80,43           | 83,67           |
| Total       | 543,08     | 444,62          | 396,66          | 401,17          | 431,25          | 442,92          |

\* refere-se a uma extrapolação para 1 hectare com base na amostra utilizada para o ajuste.

Fonte: O Autor (2015).

### 4.3 CENÁRIO 3

#### 4.3.1 COMPARAÇÃO DOS MODELOS BIOMÉTRICOS

A relação dos modelos ajustados, as estimativas da média dos parâmetros e dos critérios de informação AIC e BIC é mostrado na tabela 4, onde os melhores resultados estão destacados em negrito.

O melhor ajuste para os modelos hipsométricos foi obtido com o modelo de Richards, pois este apresentou os menores valores para ambos os critérios de seleção de modelos. Novamente, o modelo Weibull apresentou resultados muito próximo do modelo de Richards, ou seja, a diferença de ajuste entre os dois modelos para os dois critérios foi inferior a 2. O Polinômio do 1<sup>o</sup> e o modelo de Stoffels apresentaram para ambos os critérios diferenças maiores que 10 em relação ao melhor modelo, de forma que podem ser desconsiderados na estimativa hipsométrica. Por uma questão numérica, deu-se preferência ao modelo de Richards para a estimativa da altura por classe diamétrica.

Para os modelos volumétricos o melhor resultado foi obtido com o modelo de Spurr, o qual apresentou os menores valores para os critérios AIC e BIC de seleção de modelos. Segundo o AIC, a diferença entre os modelos de Spurr e Schumacher-Hall é inferior a 2, porém quando analisado os valores para o BIC há uma indicação de que o modelo de Spurr se ajustou melhor

**Tabela 14: Comparação dos modelos biométricos ajustados por máxima verossimilhança pelos critérios de informação AIC e BIC.**

| Modelo          | $\beta_0$         | $\beta_1$        | $\beta_2$ | AIC              | BIC              |
|-----------------|-------------------|------------------|-----------|------------------|------------------|
| Polinômio do 1° | 6,254161          | 0,413287         | -         | 2769,256         | 2782,447         |
| Stoffels        | 2,648479          | 0,575898         | -         | 2716,782         | 2729,972         |
| Weibull         | 24,622433         | 0,043105         | 1,038741  | 2690,218         | 2707,806         |
| Richards        | 24,243265         | 0,054814         | 1,1320314 | <b>2689,790</b>  | <b>2707,378</b>  |
| Spurr           | $3,246789e^{-5}$  | 1,025906         | -         | <b>-1298,158</b> | <b>-1385,168</b> |
| Schumacher-Hall | $3,005624e^{-5}$  | 2,086908         | 1,011307  | -1297,504        | -1279,916        |
| Ogaya           | $-5,088841e^{-5}$ | $4,476073e^{-5}$ | -         | -1292,842        | -1279,651        |

Onde:  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  são os coeficientes do modelo, AIC e BIC são o Akaike Information Criterion e Bayesian Information Criterion, respectivamente.

Fonte: O Autor (2015).

aos dados devido à diferença entre os dois ser superior a 2. O modelo de Ogaya obteve valores muito próximo ao de Schumacher-Hall quando avaliado o critério BIC, entretanto quando analisado o AIC a diferença é muito maior, mas de qualquer modo este modelo não oferece suporte empírico para a sua utilização. Desta forma, para a estimativa do volume por classe diamétrica utilizou-se o modelo de Spurr.

#### 4.3.2 AVALIAÇÃO DA CONVERGÊNCIA DAS CADEIAS DE MARKOV E SELEÇÃO DO MELHOR MODELO

A convergência das cadeias de Markov para todas as *prioris* testadas foi obtida com as 100 mil iterações iniciais propostas. Entretanto, para que a estacionaridade fosse atingida utilizando a verossimilhança foram necessárias 5 milhões de iterações. O resumo dos diagnósticos de convergência das cadeias de Markov para a verossimilhança e para as *prioris* testadas é mostrada na tabela 15.

**Tabela 15: Estimativa dos parâmetros da distribuição Weibull e diagnósticos de convergência das cadeias de Markov para a verossimilhança e *prioris* do cenário 3**

| Modelo          | Parâmetro | Moda      | Diagnóstico |          |          |                    |          |
|-----------------|-----------|-----------|-------------|----------|----------|--------------------|----------|
|                 |           |           | Z-score     | MCSE     | ESS      | BMK <sub>max</sub> | DIC      |
| Verossimilhança | <i>a</i>  | 4,451803  | -0,029245   | 0,022577 | 589,933  | 0,269290*          | 4319,179 |
|                 | <i>b</i>  | 37,828364 | 0,032380    | 0,173160 | 583,446  | 0,277246*          |          |
|                 | <i>c</i>  | -9,357965 | 0,038952    | 0,169503 | 573,763  | 0,277206*          |          |
| <i>Priori 1</i> | <i>a</i>  | 4,554811  | 0,205987    | 0,019363 | 1186,003 | 0,074419           | 4378,483 |
|                 | <i>b</i>  | 22,332258 | 0,004472    | 0,027647 | 1135,460 | 0,078278           |          |
|                 | <i>c</i>  | 2,346236  | -0,110268   | 0,005247 | 927,263  | 0,094535           |          |
| <i>Priori 2</i> | <i>a</i>  | 4,263561  | 0,837729    | 0,020423 | 1188,936 | 0,113217           | 4373,155 |
|                 | <i>b</i>  | 22,671313 | -1,000695   | 0,029818 | 1170,595 | 0,110086           |          |
|                 | <i>c</i>  | 2,418844  | -1,021187   | 0,005310 | 1005,430 | 0,138876           |          |
| <i>Priori 3</i> | <i>a</i>  | 1,162720  | 0,124488    | 0,048133 | 633,794  | 0,153773           | 4333,916 |
|                 | <i>b</i>  | 26,434329 | -0,012305   | 0,057159 | 634,891  | 0,151084           |          |
|                 | <i>c</i>  | 2,976831  | -0,124904   | 0,008919 | 690,871  | 0,144151           |          |
| <i>Priori 4</i> | <i>a</i>  | 0,084936  | 0,109343    | 0,023918 | 1107,703 | 0,104201           | 4336,162 |
|                 | <i>b</i>  | 27,604404 | -0,494364   | 0,029518 | 1300,438 | 0,104850           |          |
|                 | <i>c</i>  | 3,103530  | -0,279775   | 0,005204 | 1401,757 | 0,118604           |          |

Onde: *a*, *b*, *c* são os parâmetros do modelo, Z-score é o Diagnóstico de Convergência Geweke, MCSE é o Monte Carlo Standart Error, ESS é o Effective Sample Size, BMK<sub>max</sub> é o valor máximo do diagnóstico BMK e DIC é o Deviance Information Criterion.

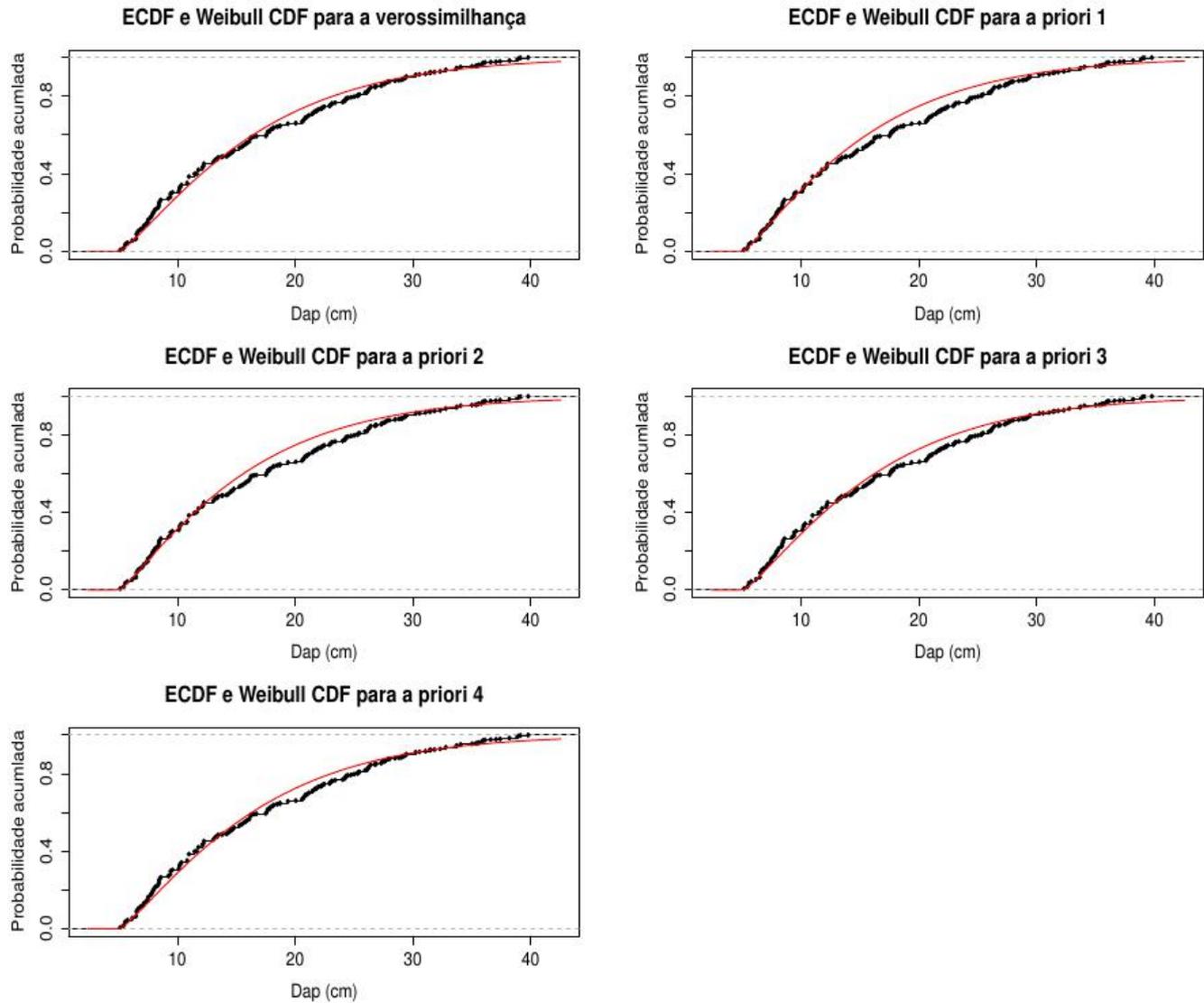
O símbolo \* refere-se aos lotes de 50000, 100000, 150000, 200000, 250000, 300000, 350000, 400000, 450000.

Fonte: O Autor (2015).

Analisando-se o Z-score observa-se que todos valores estão dentro do intervalo de  $\pm 1,96$ , indicando que as cadeias de Markov da verossimilhança e de todas as *prioris* testadas convergiram para a distribuição posterior e que houve similaridade das cadeias. Em relação ao MCSE observa-se que todos os valores estão abaixo de 6,27%, indicando que houve estacionaridade das cadeias de Markov.

Em relação ao ESS verifica-se que todos os valores são superiores a 100, indicando que houve pouca perda de informação devido a alta correlação entre as cadeias amostradas. Em relação ao valor máximo do diagnóstico BMK, todos os valores encontram-se abaixo de 0,5, indicando convergência para a distribuição posterior. A tabela 21 e 22 do Apêndice A trazem o resultado completo deste diagnóstico para o cenário 2.

Segundo o DIC o melhor ajuste foi obtido com a verossimilhança, com um valor bem inferior ao segundo melhor ajuste que foi obtido com a *priori* 3. Como o DIC possui uma interpretação análoga ao AIC e a diferença entre estes dois ajustes é de 14,737 pode-se afirmar que o ajuste por verossimilhança é muito superior aos demais e que estes não oferecem nenhum suporte para a sua utilização. Analisando-se as curvas da distribuição acumulada (figura 3), pode-se notar uma menor distância entre a curva teórica e a acumulada obtida com o ajuste.



**Figura 3:** Comparação entre a distribuição cumulativa empírica e distribuição cumulativa Weibull obtida com o ajuste por verossimilhança e inferência Bayesiana para o cenário 3.

Fonte: O Autor (2015).

A tabela 16 mostra todos os valores para o fator de Bayes encontrados com o ajuste da distribuição Weibull no cenário 3. Segundo este critério de seleção do melhor ajuste, verifica-se que os maiores valores foram calculados quando a verossimilhança foi comparada com as demais *prioris*. Todos os valores gerados ficaram acima de 7600 quando foram comparados os

ajustes da verossimilhança em relação as *prioris*. Foram encontrados valores extremos como é o caso das *prioris* 1 e 2, com valores nas casas da dezena e centena de milhões a favor da verossimilhança. O pior ajuste foi obtido com a utilização da *priori* 2, o qual apresenta um ajuste considerado fraco até mesmo perto do segundo pior ajuste, obtido com a *priori* 1. Observa-se também uma estratificação em 3 classes conforme o valor do fator de Bayes: 1 - verossimilhança; 2 - *prioris* 1 e 2; 3 - *prioris* 3 e 4. A força de evidência entre a classe 2 é apenas moderada a favor da *priori* 1 e para a classe 3 é considerada fraca a favor da *priori* 4.

**Tabela 16: Comparação do ajuste dos modelos pelo fator de Bayes para o cenário 3, onde o valor indica a superioridade do modelo da linha sobre o modelo da coluna.**

|                 | Verossimilhança         | <i>Priori</i> 1        | <i>Priori</i> 2        | <i>Priori</i> 3         | <i>Priori</i> 4         |
|-----------------|-------------------------|------------------------|------------------------|-------------------------|-------------------------|
| Verossimilhança | 1,000                   | 9,903042e <sup>7</sup> | 8,516093e <sup>8</sup> | 9300,457                | 7635,436                |
| <i>Priori</i> 1 | 1,009791e <sup>-8</sup> | 1,000000               | 8,599                  | 9,391515e <sup>-4</sup> | 7,710192e <sup>-4</sup> |
| <i>Priori</i> 2 | 1,174247e <sup>-9</sup> | 0,116                  | 1,000                  | 1,092104e <sup>-4</sup> | 8,965890e <sup>-5</sup> |
| <i>Priori</i> 3 | 1,075216e <sup>-5</sup> | 1064,791               | 9156,640               | 1,000                   | 0,821                   |
| <i>Priori</i> 4 | 1,309683e <sup>-5</sup> | 1296,985               | 11153,38               | 1,218                   | 1,000                   |

Fonte: O Autor (2015).

#### 4.3.3 ESTIMAÇÃO DO NÚMERO DE ÁRVORES E VOLUME POR HECTARE

As estimativas do número de árvores por hectare para cada nível de informação utilizada no ajuste são mostradas na tabela 17.

**Tabela 17: Estimativa do número de árvores por hectare por classe diamétrica para cada modelo no cenário 3.**

| Classe (cm) | Amostrado* | Verossimilhança | <i>Priori</i> 1 | <i>Priori</i> 2 | <i>Priori</i> 3 | <i>Priori</i> 4 |
|-------------|------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| [5,10)      | 73,37      | 54,03           | 53,50           | 52,47           | 51,23           | 53,86           |
| [10,15)     | 149,10     | 122,78          | 178,21          | 173,16          | 145,77          | 144,30          |
| [15,20)     | 189,33     | 217,01          | 282,88          | 280,04          | 253,30          | 248,87          |
| [20,25)     | 260,33     | 302,93          | 318,84          | 320,78          | 319,99          | 316,90          |
| [25,30)     | 295,33     | 326,20          | 279,31          | 283,55          | 306,39          | 307,26          |
| [30,35)     | 186,97     | 257,76          | 195,97          | 198,71          | 222,56          | 225,72          |
| [35,40)     | 265,07     | 139,29          | 111,28          | 111,29          | 120,77          | 123,09          |
| Total       | 1420       | 1420            | 1420            | 1420            | 1420            | 1420            |

\* refere-se a uma extrapolação para 1 hectare com base na amostra utilizada para o ajuste.

Fonte: O Autor (2015).

Observa-se que todos os ajustes apresentam os valores muito próximos na primeira classe diamétrica, porém a partir da segunda até a classe dos 25 a 30 cm a verossimilhança

apresenta as menores estimativas. As estimativas obtidas com as *prioris* 1 e 2 são muito semelhantes, sendo que a maior diferença entre as duas é inferior a 5 árvores (classe de 10 a 15 cm de diâmetro). O mesmo ocorre com as estimativas obtidas com as *prioris* 3 e 4, onde a diferença máxima é inferior a 5 árvores (classe de 15 a 20 cm diâmetro). Este comportamento é corroborado com os valores do fator de Bayes e a separação em 3 classes como descrito anteriormente. Novamente houve pouca semelhança entre as estimativas e o valor amostrado.

Em relação a estimativa de volume por hectare (Tabela 18), verifica-se que a maior estimativa de volume total foi obtida com o ajuste por verossimilhança (734,54 m<sup>3</sup>/ha). A estratificação das classes pelo número de árvores fica mais evidente quando comparados os volumes totais. A diferença entre as estimativas entre as *prioris* 1 e 2 é de 4,65 m<sup>3</sup>/ha (0,727%) e entre as *prioris* 3 e 4 de 4,25 m<sup>3</sup>/ha (0,622%). Quando comparada a diferença entre as classes 1 e 2 (menor estimativa) obtêm-se uma divergência de 94,94 m<sup>3</sup>/ha (14,34%). A diferença entre as classes 1 e 3 (menor estimativa) é de 54,4 m<sup>3</sup>/ha (7,52%) e entre as classes 2 e 3 (menores estimativas para ambas) é de 43,54 m<sup>3</sup>/ha (6,81%).

A diferença entre o volume total amostrado e o volume total da melhor estimativa (verossimilhança) é de 51,45 m<sup>3</sup>/ha (7%). O que contribuiu para o volume total da amostra ser superior foi o maior número de árvores na classe de 35 a 40 cm, o que equivale a uma diferença de volume de 152,85 m<sup>3</sup>/ha (88,70%).

**Tabela 18: Estimativa do volume por hectare (m<sup>3</sup>/ha) por classe diamétrica para cada modelo no cenário 3**

| Classe (cm) | Amostrado* | Verossimilhança | Priori 1 | Priori 2 | Priori 3 | Priori 4 |
|-------------|------------|-----------------|----------|----------|----------|----------|
| [5,10)      | 1,23       | 0,82            | 0,81     | 0,79     | 0,77     | 0,81     |
| [10,15)     | 9,82       | 8,28            | 12,02    | 11,68    | 9,83     | 9,73     |
| [15,20)     | 31,88      | 37,60           | 49,01    | 48,52    | 43,89    | 43,12    |
| [20,25)     | 87,10      | 103,28          | 108,70   | 109,36   | 109,09   | 108,04   |
| [25,30)     | 164,98     | 187,22          | 160,31   | 162,74   | 175,85   | 176,35   |
| [30,35)     | 165,79     | 225,03          | 171,08   | 173,48   | 194,30   | 197,05   |
| [35,40)     | 325,19     | 172,33          | 137,68   | 137,68   | 149,41   | 152,28   |
| Total       | 785,99     | 734,54          | 639,60   | 644,25   | 683,14   | 687,39   |

\* refere-se a uma extrapolação para 1 hectare com base na amostra utilizada para o ajuste.

Fonte: O Autor (2015).

## 5 CONCLUSÃO

A utilização de reamostragem gerou os cenários esperados. Entretanto, foi difícil obter um bom ajuste da distribuição Weibull.

Considerando o DIC como critério de seleção de modelos, o melhor ajuste para os cenários 1 e 2 foi obtido com a utilização da *priori* 3 e para o cenário 3 foi obtido com a verossimilhança. A utilização das *prioris* com os parâmetros de forma e escala simétricos permite um ajuste mais eficiente da distribuição Weibull.

A diferença entre os melhores ajuste utilizando as *prioris* e a verossimilhança não foi considerada substancial para a utilização dos mesmos, sendo apenas uma pequena vantagem numérica. A grande quantidade de dados tornou o efeito da verossimilhança muito pronunciado e isto foi um fator determinante no ajuste. Além disso, a seleção do melhor ajuste pelo fator de Bayes e DIC levou a divergências.

O número de árvores por classe diamétrica divergiu entre os valores amostrados e as estimativas para todos os cenários, principalmente nas últimas classes diamétricas, as quais influenciam fortemente o volume. Os melhores ajustes subestimam o volume total por hectare em 20,18%, 25,93% e 7% para os cenários 1, 2 e 3, respectivamente.

De forma geral, a inferência Bayesiana e a utilização dos Métodos de Monte Carlos via Cadeias de Markov não se mostraram eficientes no ajuste da distribuição Weibull. O tamanho da amostra e a distribuição dos dados não favoreceram um bom ajuste da função.

## REFERÊNCIAS

- BARROS, Paulo Luis C. de. **Estudo das Distribuições Diamétricas da Floresta do Planalto do Tapajós**. 123p. Dissertação (Mestrado) — Universidade Federal do Paraná, Curitiba, 1980.
- BATISTA, João L. F. **Verossimilhança e Máxima Verossimilhança**. 2009. Acesso em: 10 nov. 2014. Disponível em: <[cmq.esalq.usp.br/BIE5781/lib/exe/fetch.php?media=03-funcao...pdf](http://cmq.esalq.usp.br/BIE5781/lib/exe/fetch.php?media=03-funcao...pdf)>.
- BURNAM, Kenneth P.; ANDERSON, David R. **Model selection and multimodel inference: a practical information-theoretic approach**. 2. ed. New York: Springer. 2002. 488 p.
- CAMPOS, João C. C.; LEITE, Helio G. **Mensuração Florestal**. 4. ed. Viçosa: UFV, 2013. 605 p.
- CHIHARA, Laura; HESTERBERG, Tim. **Mathematical Statistics with Resampling and R**. 1. ed. New Jersey, USA: John Wiley & Sons. 2012. 418 p.
- COWLES, Mary k. **Applied Bayesian Statistics with R and OpenBUGS Examples**. 1 ed. New York, USA: Springer. 2013. 232 p.
- DUAN, Ai-guo et al. Stand diameter distribution modelling and prediction based on Richards function. **Plos One**, v. 8, n. 4, p.1–12, 2013.
- EHLERS, Ricardo S. Inferência Bayesiana. 2011. Acesso em: 10 nov. 2014. Disponível em: <[www.icmc.usp.br/ehlers/bayes/bayes.pdf](http://www.icmc.usp.br/ehlers/bayes/bayes.pdf)>.
- FERREIRA, Rinaldo L. C.; SOUZA, Agostinho. L. de; REGAZZI, A. J. Dinâmica da estrutura de uma floresta secundária de transição. I - estrutura paramétrica. **Revista Árvore**, v. 22, n. 1, p.21–27, 1998.
- FLEGAL, James M. **Monte Carlo Standard Errors for Markov Chain Monte Carlo**. 181 p. Dissertação (Doutorado) - Universidade de Minnesota, Minnesota, USA, 2008.
- GAMERMAM, Dani; LOPES, Hedibert F. **Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference**. 2.ed. Boca Raton: Chapman & Hall\CRC. 2006. 344 p.
- GELMAN, Andrew; CARLIN, John B.; STERN, Hal S.; RUBIN, Donald B. **Bayesian data analysis**. 2. ed. Boca Raton: Chapman & Hall\CRC. 2004. 668 p.
- HOLL, Karen D. Restoring Tropical Forest. 2013. **Nature Education Knowledge**. Acesso em 10 nov. 2014. Disponível em: <<http://www.nature.com/scitable/knowledge/library/restoring-tropical-forest-97756726>>.

HURVICH, Clifford M.; TSAI, Chih-Ling. Regression and time series model selection in small samples. **Biometrika**, v. 76, n. 2, p.297–307, jun 1989.

IMAÑA-ENCIMAS, José; PAULA, José E. de; CONCEIÇÃO, Cláudio de A. Florística, volume e biomassa lenhosa de um fragmento de Mata Atlântica no município de Santa Maria de Jetibá, Espírito Santo. **Floresta**. v. 42, n. 3, p. 565 - 576, jul./set. 2012.

KASS, Robert E.; RAFTERY, Adrian E. Bayes Factors. **Journal of the Americal Statistical Association**. v. 90, n. 430, p. 773-795, jun. 1995.

LONGMAN, Kenneth A.; JENÍK, Jan. **Tropical Forest and its Environment**. 1. ed. Norfolk, Inglaterra: Longman Group Limited, 1974. 196p.

MARIA, Anu. Introduction to modeling and simulation. In: **Proceedings of the 29th Conference on Winter Simulation**. Washington, DC, USA: IEEE Computer Society, 1997. (WSC '97), p.7–13. Disponível em: <<http://dx.doi.org/10.1145/268437.268440>>.

MARTEN, Gerald. G. The tropical rain forest as an ecosystem. In: RAMBO, A. T.; SAJISE, P. E. (Ed.). **An Introduction to Human Ecology Research on Agricultural Systems in Southeast Asia**. University of the Phillipines, 1984. p.61–74.

MILLAR, Russel B. Comparison of Hierarchical Bayesian Models for Overdispersed Count Data using DIC and Bayes' Factors. **Biometrics**. n. 65, p. 962-969, set. 2009.

NTZOUFRAS, Ioannis. **Bayesian Modeling Using WinBUGS**. Wiley Series in Computational Statistics. New Jersey, USA: John Wiley & Sons, 2009. 590p.

PAULINO, Carlos D. M.; TURKMAN, Maria A. A.; MURTEIRA, Bento. **Estatística Bayesiana**. Fundação Calouste Gulbenkian, 2003. 446 p.

PENG, Changui; WEN, Xuezhi. Forest simulation models. In: SHAO, Guofan; REYNOLDS, Keith M. (Ed.). **Computer Applications in Sustainable Forest Management**. Springer Netherlands, 2006. v. 11, p.101–125.

POGGIANI, F. Estrutura, Funcionamento e Classificação das Florestas: Implicações Ecológicas das Florestas Plantadas. **Documentos florestais**, v. 3, p.1–14, set 1989.

SAHLIN, Kristoffer. **Estimating convergence of Markov chain Monte Carlo simulations**. 36 p. Tese (Doutorado em Estatística Matemática) - Universidade de Estocolmo, Estocolmo, Suécia. 2011.

SCOLFORO, José R. S. **Biometria florestal: modelos de crescimento e produção florestal**. Lavras: UFLA, 2006.

SILVA, Gilson F. da. **Problemas no uso de programação matemática e simulação em regulação florestal**. 89p. Tese (Doutorado) — Universidade Federal de Viçosa, Viçosa, 2001.

SIMON, Julian L. **Resampling: The New Statistics**. 1997. 209 p.

SOUZA, Agostinho L. de; SOARES, Carlos P. B. **Florestas Nativas: estrutura, dinâmica e manejo**. Viçosa-MG: UFV, 322p, 2013.

VISMARA, Lilian de S. **Aplicação da Inferência Bayesiana para a Simulação da Dinâmica de Produção de Sementes de Plantas Daninhas**. 223p. Dissertação (Mestrado) — Universidade de São Paulo, 2006.

YU, Chong H. **Resampling Methods: Concepts, Applications, and Justification 2003**. Practical Assessment, Research & Evaluation. Acesso em 10 nov. 2014. Disponível em: <<http://pareonline.net/getvn.asp?v=8&n=19>>.

## **6 APÊNDICE - A**

A inserção destas tabelas neste apêndice e não no decorrer do texto é motivada pelo fato de que apenas os maiores valores do diagnóstico de convergência BMK são considerados de interesse e não é realizada qualquer análise mais aprofundada sobre as relações entre os valores de cada um dos 9 lotes do diagnóstico, além do tamanho considerável das mesmas. Desta forma, as tabelas ficam como uma informação adicional para conferência a quem interessar.

**Tabela 19: Diagnóstico de Convergência BMK das cadeias de Markov para a verossimilhança e priors do cenário 1**

| Modelo          | Parâmetro | Lotes      |            |            |            |            |            |            |            |            |  |
|-----------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--|
|                 |           | 1000       | 2000       | 3000       | 4000       | 5000       | 6000       | 7000       | 8000       | 9000       |  |
| Verossimilhança | <i>a</i>  | 0,09804274 | 0,08420153 | 0,09774082 | 0,09640754 | 0,09701184 | 0,09032869 | 0,08421475 | 0,09174659 | 0,13990478 |  |
|                 | <i>b</i>  | 0,04175000 | 0,06947439 | 0,05856569 | 0,06793605 | 0,04817055 | 0,03571840 | 0,04099298 | 0,05457338 | 0,05245058 |  |
|                 | <i>c</i>  | 0,07482320 | 0,12017845 | 0,20100638 | 0,13136544 | 0,07317947 | 0,09615094 | 0,07065458 | 0,10729225 | 0,08723484 |  |
| <i>Priori 1</i> | <i>a</i>  | 0,12018122 | 0,09169410 | 0,15096530 | 0,18655413 | 0,12771398 | 0,10793268 | 0,10066303 | 0,11706463 | 0,12209389 |  |
|                 | <i>b</i>  | 0,05831114 | 0,05449081 | 0,05150722 | 0,07635962 | 0,05974711 | 0,03540554 | 0,05174625 | 0,05171774 | 0,04852985 |  |
|                 | <i>c</i>  | 0,13576564 | 0,12290779 | 0,06564628 | 0,11692938 | 0,11619552 | 0,09375094 | 0,12936105 | 0,11221476 | 0,09972887 |  |
| <i>Priori 2</i> | <i>a</i>  | 0,11175804 | 0,10375592 | 0,10078131 | 0,07869774 | 0,16064796 | 0,13274172 | 0,13924837 | 0,14391400 | 0,10471152 |  |
|                 | <i>b</i>  | 0,03303133 | 0,04438548 | 0,03723828 | 0,03570889 | 0,03517911 | 0,06179092 | 0,08391743 | 0,05315987 | 0,07065519 |  |
|                 | <i>c</i>  | 0,11630222 | 0,12509315 | 0,09427514 | 0,09920625 | 0,08346245 | 0,12526043 | 0,11538518 | 0,13867689 | 0,13557298 |  |
| <i>Priori 3</i> | <i>a</i>  | 0,12331398 | 0,13515295 | 0,11950793 | 0,12701440 | 0,09922347 | 0,09938203 | 0,06274304 | 0,12724128 | 0,09207599 |  |
|                 | <i>b</i>  | 0,04014526 | 0,04873773 | 0,07023061 | 0,07435988 | 0,05017543 | 0,05516616 | 0,03342943 | 0,04487848 | 0,06244680 |  |
|                 | <i>c</i>  | 0,08805402 | 0,09276543 | 0,14469428 | 0,10724328 | 0,08951238 | 0,11551510 | 0,11000268 | 0,11096732 | 0,11681204 |  |
| <i>Priori 4</i> | <i>a</i>  | 0,11255461 | 0,08878755 | 0,08921162 | 0,09808295 | 0,09602169 | 0,07742643 | 0,1391726  | 0,13458437 | 0,12480397 |  |
|                 | <i>b</i>  | 0,06733318 | 0,04593994 | 0,05799584 | 0,05492319 | 0,04271256 | 0,05852942 | 0,0387457  | 0,04148288 | 0,04579155 |  |
|                 | <i>c</i>  | 0,10503217 | 0,15072463 | 0,13977415 | 0,14235821 | 0,12293555 | 0,11421775 | 0,1041291  | 0,07878208 | 0,09434326 |  |

Onde: *a*, *b*, *c* são os coeficientes do modelo, 1000, 2000, ..., 9000 são os lotes utilizados como amostra.

Fonte: O Autor (2015).

**Tabela 20: Diagnóstico de Convergência BMK das cadeias de Markov para a verossimilhança e priors do cenário 2**

| Modelo          | Parâmetro | Lotes      |            |            |            |            |            |            |            |            |
|-----------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
|                 |           | 1000       | 2000       | 3000       | 4000       | 5000       | 6000       | 7000       | 8000       | 9000       |
| Verossimilhança | <i>a</i>  | 0,05130344 | 0,1113981  | 0,1266639  | 0,10406538 | 0,1241391  | 0,06367841 | 0,07844392 | 0,06512445 | 0,08674395 |
|                 | <i>b</i>  | 0,06773256 | 0,1219224  | 0,1090516  | 0,07761685 | 0,1082991  | 0,07099060 | 0,08933938 | 0,07589394 | 0,07349330 |
|                 | <i>c</i>  | 0,09057833 | 0,1542239  | 0,1622706  | 0,11728027 | 0,1704660  | 0,09684974 | 0,10341959 | 0,07899588 | 0,09681322 |
| <i>Priori 1</i> | <i>a</i>  | 0,08034115 | 0,09676742 | 0,08397779 | 0,05835241 | 0,08909711 | 0,05157700 | 0,11263928 | 0,07513527 | 0,05757823 |
|                 | <i>b</i>  | 0,06261903 | 0,04702739 | 0,06401192 | 0,05825585 | 0,07033994 | 0,05485485 | 0,08907795 | 0,06430923 | 0,05714375 |
|                 | <i>c</i>  | 0,09520310 | 0,14203237 | 0,08867858 | 0,06651735 | 0,09055269 | 0,06212051 | 0,13162951 | 0,11486322 | 0,10080514 |
| <i>Priori 2</i> | <i>a</i>  | 0,08192311 | 0,08329113 | 0,05020303 | 0,04349600 | 0,04035519 | 0,06006013 | 0,08853376 | 0,05468311 | 0,05034341 |
|                 | <i>b</i>  | 0,09361086 | 0,07088369 | 0,04577307 | 0,05477820 | 0,05207243 | 0,03457160 | 0,06686114 | 0,04993328 | 0,03077332 |
|                 | <i>c</i>  | 0,16839912 | 0,13308763 | 0,09669027 | 0,07148648 | 0,05275365 | 0,07349872 | 0,12226693 | 0,09215467 | 0,11147856 |
| <i>Priori 3</i> | <i>a</i>  | 0,1119588  | 0,07437306 | 0,06418430 | 0,06354901 | 0,09153480 | 0,08757760 | 0,08171573 | 0,04224073 | 0,08023070 |
|                 | <i>b</i>  | 0,1011617  | 0,05098340 | 0,05654148 | 0,04292068 | 0,06709008 | 0,06586714 | 0,06488991 | 0,04057622 | 0,07471487 |
|                 | <i>c</i>  | 0,1269167  | 0,08533099 | 0,12721657 | 0,09497297 | 0,14384276 | 0,15188087 | 0,12172779 | 0,06377129 | 0,11784357 |
| <i>Priori 4</i> | <i>a</i>  | 0,04843092 | 0,05076321 | 0,04250388 | 0,04504981 | 0,02960181 | 0,05179411 | 0,09413494 | 0,1188685  | 0,09067387 |
|                 | <i>b</i>  | 0,03572881 | 0,04336301 | 0,03919614 | 0,03682032 | 0,02481303 | 0,04137742 | 0,09157180 | 0,1654073  | 0,12069550 |
|                 | <i>c</i>  | 0,05759209 | 0,06695374 | 0,07065853 | 0,06656215 | 0,04973102 | 0,05755311 | 0,10277411 | 0,1679537  | 0,12854676 |

Onde: *a*, *b*, *c* são os coeficientes do modelo, 1000, 2000, ..., 9000 são os lotes utilizados como amostra.

Fonte: O Autor (2015).

Tabela 21: Diagnóstico de Convergência BMK das cadeias de Markov para as *prioris* do cenário 3

| Modelo          | Parâmetro | Lotes      |            |            |            |            |            |            |            |            |
|-----------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
|                 |           | 1000       | 2000       | 3000       | 4000       | 5000       | 6000       | 7000       | 8000       | 9000       |
| <i>Priori 1</i> | <i>a</i>  | 0,06752340 | 0,07056044 | 0,04927891 | 0,06391668 | 0,06379360 | 0,03380474 | 0,07441860 | 0,07092452 | 0,05360398 |
|                 | <i>b</i>  | 0,06766094 | 0,06589171 | 0,04409821 | 0,03702003 | 0,06444374 | 0,05289773 | 0,07827797 | 0,07169843 | 0,03569974 |
|                 | <i>c</i>  | 0,08383105 | 0,06841883 | 0,08873661 | 0,10131818 | 0,08234854 | 0,07746007 | 0,09453470 | 0,07343553 | 0,08131018 |
| <i>Priori 2</i> | <i>a</i>  | 0,08673249 | 0,05002717 | 0,03367437 | 0,07922064 | 0,08043986 | 0,09004248 | 0,1119579  | 0,11321705 | 0,05785399 |
|                 | <i>b</i>  | 0,09095750 | 0,07041305 | 0,06084653 | 0,07344957 | 0,07622152 | 0,08582277 | 0,1100859  | 0,09545449 | 0,06551835 |
|                 | <i>c</i>  | 0,09029998 | 0,08752560 | 0,11294686 | 0,10562922 | 0,07874386 | 0,10146341 | 0,1145880  | 0,13887635 | 0,07260834 |
| <i>Priori 3</i> | <i>a</i>  | 0,07289547 | 0,05565084 | 0,04753980 | 0,1537733  | 0,09830921 | 0,06106925 | 0,06693662 | 0,07481740 | 0,05127566 |
|                 | <i>b</i>  | 0,10357208 | 0,06039165 | 0,05704341 | 0,1510844  | 0,09773528 | 0,07271335 | 0,07287395 | 0,07709969 | 0,05846917 |
|                 | <i>c</i>  | 0,09423530 | 0,07445379 | 0,05407374 | 0,1441508  | 0,09261066 | 0,08872494 | 0,09693875 | 0,09542226 | 0,07377442 |
| <i>Priori 4</i> | <i>a</i>  | 0,10357584 | 0,1042008  | 0,06065875 | 0,06120083 | 0,09262542 | 0,05253531 | 0,09300414 | 0,05035487 | 0,04286285 |
|                 | <i>b</i>  | 0,08011021 | 0,1048498  | 0,06709054 | 0,04124901 | 0,08896212 | 0,05917495 | 0,07698561 | 0,07111254 | 0,05240816 |
|                 | <i>c</i>  | 0,10662828 | 0,1186042  | 0,11043095 | 0,08823727 | 0,10241699 | 0,07803606 | 0,09375017 | 0,07406554 | 0,06327247 |

Onde: *a*, *b*, *c* são os coeficientes do modelo, 1000, 2000, ..., 9000 são os lotes utilizados como amostra.

Fonte: O Autor (2015).

**Tabela 22: Diagnóstico de Convergência BMK das cadeias de Markov para as *prioris* do cenário 3**

| Modelo                 | Parâmetro | Lotes      |            |            |           |           |           |           |           |           |  |  |  |
|------------------------|-----------|------------|------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|--|--|--|
|                        |           | 50000      | 100000     | 150000     | 200000    | 250000    | 300000    | 350000    | 400000    | 450000    |  |  |  |
| <i>Verossimilhança</i> | <i>a</i>  | 0,09608816 | 0,07173625 | 0,05411281 | 0,1459058 | 0,2692902 | 0,2004604 | 0,1405185 | 0,1050818 | 0,1277646 |  |  |  |
|                        | <i>b</i>  | 0,10250128 | 0,07877446 | 0,06333088 | 0,1509134 | 0,2772458 | 0,2073476 | 0,1484740 | 0,1131185 | 0,1345001 |  |  |  |
|                        | <i>c</i>  | 0,10205694 | 0,07825657 | 0,06298647 | 0,1509195 | 0,2772056 | 0,2073093 | 0,1482659 | 0,1127730 | 0,1344192 |  |  |  |

Onde: *a*, *b*, *c* são os coeficientes do modelo, 50000, 100000, ..., 450000 são os lotes utilizados como amostra.

Fonte: O Autor (2015).