

**ANÁLISE BAYESIANA PARA MODELOS DE
DEGRADABILIDADE RUMINAL**

TACIANA VILLELA SAVIAN

2008

TACIANA VILLELA SAVIAN

**ANÁLISE BAYESIANA PARA MODELOS DE
DEGRADABILIDADE RUMINAL**

Tese apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de “Doutor”.

Orientador

Prof. Dr. Joel Augusto Muniz

LAVRAS
MINAS GERAIS – BRASIL
2008

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Savian, Taciana Villela.

Análise bayesiana para modelos de degradabilidade ruminal / Taciana Villela Savian. -- Lavras : UFLA, 2008.

81p. : il.

Tese (Doutorado) – Universidade Federal de Lavras, 2008.

Orientador: Joel Augusto Muniz.

Bibliografia.

1. Degradabilidade in situ. 2. Inferência bayesiana. 3. Modelos não-lineares. 4. Métodos MCMC. 5. Plantas forrageiras I. Universidade Federal de Lavras. II. Título.

CDD – 519.542

TACIANA VILLELA SAVIAN

**ANÁLISE BAYESIANA PARA MODELOS DE
DEGRADABILIDADE RUMINAL**

Tese apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de “Doutor”.

APROVADA em 16 de maio de 2008

Profa. Dra. Thelma Sáfadi	UFLA
Prof. Dr. Renato Ribeiro de Lima	UFLA
Prof. Dr. Fabyano Fonseca e Silva	UFV
Dr. Adauto Ferreira Barcelos	EPAMIG/Lavras

Prof. Dr. Joel Augusto Muniz
UFLA
(Orientador)

LAVRAS
MINAS GERAIS – BRASIL

Mãe, Pai, Izaura, Mareska, Cyro, Lara e Theozinho,

"Para estar junto não é preciso estar perto, e sim do lado de dentro".

(Leonardo da Vinci)

Hoje quero agradecer!

**A você, Heber, que estive ao meu lado nas horas que chorei e nas horas que sorri, nas horas que me lamentei e nas horas e que, de uma forma ou de outra, demonstrei total alegria...
Agradecer pelo sorriso diário, sem mágoas nem rancores,
agradecer de alma explosiva...
Hoje quero parar e agradecer porque você fez, faz e fará sempre parte de minha história!**

"Um amigo não racha apenas a gasolina. Racha lembranças, crises de choro, experiências, racha a culpa, racha segredos. Um amigo não empresta apenas o caderno. empresta o verbo, empresta o ombro, empresta o tempo, empresta o calor e a jaqueta. Um amigo não recomenda apenas um disco. Recomenda cautela, recomenda um emprego, recomenda um país. Um amigo não dá carona apenas pra festa. Leva-te pro mundo dele, e topa conhecer o teu. Um amigo não caminha apenas no shopping. Anda em silêncio na dor, entra contigo em campo, sai do fracasso ao teu lado. Um amigo não segura a barra, apenas. Segura a mão, a ausência, segura uma confissão, segura o tranco".

Obrigada, meus amigos!!!

AGRADECIMENTOS

A DEUS, por ter me concedido a graça da vida e por me amparar em todos os momentos.

À Universidade Federal de Lavras e ao Programa de Pós-Graduação Estatística e Experimentação Agropecuária.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig), pelo suporte financeiro concedido para a realização deste trabalho.

Ao professor Joel Augusto Muniz, por ter me honrado com sua orientação.

Ao professor Luiz Henrique de Aquino, pelo incentivo e por ter acreditado em mim.

À professora Thelma Sáfadi, pela co-orientação e ensinamentos.

Aos professores Fabyano Fonseca e Silva, Adauto Ferreira Barcelos, Renato Ribeiro de Lima e João Domingos Scalon, pela valiosa contribuição.

Aos demais professores do Departamento de Ciências Exatas, pelos ensinamentos e por terem me preparado profissionalmente.

Aos colegas do Departamento, por servirem de exemplo.

Aos colegas Ricardo Luis dos Reis e Anderson Castro Soares de Oliveira, pela paciência e por serem verdadeiros anjos da guarda.

À Josi, Joyce, Maria, Vânia, Selminha e Edila, pela atenção e pelos cafezinhos.

Ao Sr. Ubirajara e Sra. Idalci, por me acolherem como a uma filha.

Aos amigos, Felinha, Cassiana, Lucas, Meire, Elter, Néia, Érika, Serginho, Marilza, Robson, Cleube, enfim a todos os amigos pelos dias felizes e por tornarem a minha estada em Lavras um grande prazer.

Aos meus tios, Maria Izaura (minha segunda mãe), Júnior e Simone, Luiz e Miriam, Guido e Cristina, pelo amor, confiança e apoio. Sempre serei grata.

SUMÁRIO

	Página
RESUMO.....	i
ABSTRACT.....	ii
1 INTRODUÇÃO.....	1
2 REFERENCIAL TEÓRICO.....	3
2.1 Modelos não-lineares.....	3
2.2 Modelos empíricos de degradabilidade ruminal.....	5
2.3 Inferência bayesiana.....	9
2.3.1 Teorema de Bayes.....	10
2.3.2 Distribuições <i>a priori</i>	11
2.3.3 Distribuição <i>a posteriori</i> e estimação bayesiana	13
2.4 Métodos de simulação.....	15
2.4.1 Métodos de Monte Carlo via cadeias de Markov.....	15
2.4.1.1 Algoritmo de Metropolis-Hastings.....	16
2.4.1.2 Amostrador de Gibbs.....	17
2.4.2 Diagnósticos de convergência.....	19
2.4.3 Erro de Monte Carlo.....	22
2.4.4 Fator de Bayes.....	23
3 MATERIAL E MÉTODOS.....	27
3.1 Dados simulados.....	28
3.2 Dados reais.....	29
3.3 Implementação das análises.....	31
4 RESULTADOS E DISCUSSÃO.....	33
4.1 Modelo de Orskov e McDonald (1979).....	33

4.1.1. Obtenção das distribuições condicionais completas para os parâmetros do modelo de Orskov e McDonald (1979).....	33
4.2 Modelo de McDonald (1981).....	37
4.2.1. Obtenção das distribuições condicionais completas para os parâmetros do modelo de McDonald (1981).....	38
4.3 Análise dos dados simulados.....	42
4.3.1. Modelo de Orskov e McDonald (1979).....	42
4.3.2. Modelo de McDonald (1981).....	50
4.3.3. Comparação dos modelos.....	57
4.4 Análise dos dados reais.....	58
4.4.1. Modelo de Orskov e McDonald (1979).....	58
4.4.2. Modelo de McDonald (1981).....	65
4.4.3. Comparação dos modelos.....	72
5 CONCLUSÕES.....	74
6 REFERÊNCIAS BIBLIOGRÁFICAS.....	75

RESUMO

SAVIAN, Taciana Villela. **Análise bayesiana para modelos de degradabilidade ruminal**. 2008. 81 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.*

O conhecimento da digestibilidade das forragens, ou seja, a proporção do alimento consumido que é digerido e metabolizado pelo animal, é importante para se alcançar ótima fermentação ruminal e adequado balanceamento de rações. Os modelos utilizados neste tipo de estudo são classificados como modelos de regressão não-linear e apresentam algumas peculiaridades quanto ao procedimento de ajuste. Foi realizado um estudo por meio de dados simulados e dados reais da porcentagem de degradação da fibra em detergente neutro da gramínea coastcross (*Cynodon dactylon* x *Cynodon nlemfuensis*), ao longo do tempo. O ajuste dos modelos de Orskov & McDonald (1979) e McDonald (1981) foi feito por meio da metodologia bayesiana e, posteriormente, os modelos foram comparados pelo Fator de Bayes. A partir dos parâmetros estimados, foram encontradas as degradabilidades efetiva e potencial da gramínea. As amostras das distribuições marginais *a posteriori* dos parâmetros foram obtidas por meio do algoritmo Amostrador de Gibbs e algoritmo Metropolis-Hastings. A cada iteração dos algoritmos foram obtidas, por meio das estimativas dos parâmetros, a degradabilidade efetiva e potencial da gramínea. A metodologia bayesiana mostrou-se bastante eficiente, sendo avaliada e comprovada pelo processo de simulação, que apresentou estimativas bem próximas ao valor paramétrico. As estimativas obtidas para os parâmetros do modelo de Orskov & McDonald (1979) e McDonald (1981) por meio da abordagem bayesiana mostraram-se bastante coerentes com os valores relatados na literatura. O modelo de Orskov & McDonald (1979) foi mais plausível que o modelo de McDonald (1981) na descrição dos dados de degradação.

* Comitê Orientador: Joel Augusto Muniz - UFLA (orientador), Thelma Sáfyadi – UFLA (Co-Orientadora).

ABSTRACT

SAVIAN, Taciana Villela. **Bayesian analysis for ruminal degradability models**. 2008. 81 p. Thesis (Doctorate in Statistics and Agricultural Experimentation) – Federal University of Lavras, Lavras, MG, Brazil.*

The knowledge of forage digestibility, that is, the proportion of feeds consumed that is digested and metabolized by an animal, is important not only to maximize the fermentation in the rumen but also to obtain success in ration balancing. The models used in these kinds of studies are classified as non linear regression models and have some peculiarities in the fitting process. A study was conducted by using both simulated and real data of percentage of neutral detergent fiber degradation of coastcross grass (*Cynodon dactylon* x *Cynodon nlemfuensis*) over the time. First, the models proposed by Orskov & McDonald (1979) and McDonald (1981) were fitted using Bayesian methodology and then these models were compared by using Bayes factor. Effective and potential degradability of grass were calculated from the estimated parameters. The samples of posterior marginal distributions for the parameters were obtained by Gibbs Sampler and Metropolis-Hastings algorithms. It were obtained both the effective and the potential degradability of grass for each iteration of the algorithms by using the parameter estimates. The bayesian approach, evaluated and verified by the simulation process, has proved to be very efficient and presented estimated parameters quite close to the parametric values. The parameters estimates of the models by Orskov & McDonald (1979) and McDonald (1981) using bayesian approach from real data showed a fairly consistent with the values reported in the literature. The model by Orskov & McDonald (1979) was more likely than the model by McDonald (1981) to the description of degradation data.

* Guidance Comitee: Prof. DSc. Joel Augusto Muniz - UFLA (Adviser) and Prof. DSc. Thelma Sáfadi – UFLA.

1 INTRODUÇÃO

As pastagens constituem a base da dieta dos ruminantes, na grande maioria dos sistemas de produção das regiões tropicais. Na composição botânica dessas pastagens, é encontrada uma ampla variação de espécies na sua maioria representadas por gramíneas e leguminosas, que podem ser nativas ou cultivadas, cujas qualidades nutritivas são muito variáveis. Tais variações de qualidade ocorrem não somente entre gêneros, espécies ou cultivares, mas também com as diferentes partes das plantas, o estágio de maturidade, a fertilidade do solo e com as condições locais e estacionais.

A estimativa do valor nutritivo das forrageiras é de grande importância, seja para permitir adequado manejo alimentar de dietas à base de volumosos, ou para orientar no melhoramento e seleção de forrageiras com maior valor nutritivo. Na análise deste último destacam-se a quantificação dos teores de proteína bruta, dos componentes da parede celular e dos coeficientes de digestibilidade da matéria seca. Segundo Reis & Rodrigues (1993), o valor nutritivo se refere ao conjunto formado pela composição química da forragem, sua disponibilidade e a natureza dos produtos da digestão.

A distribuição dos diversos componentes químicos nas plantas varia nos diferentes tecidos e órgãos, em razão de especificidade da organização física das células vegetais. Entretanto, de modo geral, os principais constituintes químicos das plantas forrageiras, podem ser divididos em duas grandes categorias: os que compõem a estrutura da parede celular, que são de mais baixa disponibilidade no processo de digestão, e os contidos no conteúdo celular, de maior disponibilidade. Os componentes do conteúdo celular envolvem substâncias solúveis em água ou levemente solúveis em água, tais como amido, lipídios e algumas proteínas que são digeridas tanto por enzimas dos microrganismos

quanto por aquelas secretadas pelo aparelho digestivo dos animais. Já os componentes da estrutura da parede celular incluem, em sua maior parte, carboidratos estruturais (celulose e hemicelulose) e outras substâncias, como a lignina, cuja digestão é totalmente dependente da atividade enzimática dos microrganismos do trato gastrointestinal dos ruminantes, principalmente no rúmen.

O conhecimento da digestibilidade das forragens, ou seja, a proporção do alimento consumido que é digerido e metabolizado pelo animal é importante para se otimizar a fermentação ruminal com o balanceamento de rações. Existem diferentes métodos para medir a degradabilidade dos alimentos e, entre estes, citam-se o método de determinação *in vivo*, o método *in vitro* e o método *in situ*.

Na área zootécnica, diversos são os fenômenos descritos por modelos de regressão não-linear, como curvas de crescimento, curvas de produção de ovos e curvas de lactação. Esses fenômenos são amplamente discutidos na literatura por meio das abordagens clássica e bayesiana, encontrando-se excelentes resultados. Porém, não são encontrados artigos na área de nutrição animal, mais especificamente em estudos de cinética de degradação ruminal, que são tratados com a abordagem bayesiana.

Diante do exposto, o presente trabalho foi realizado com os objetivos de estudar a cinética de digestão ruminal por meio dos modelos não-lineares de Orskov & McDonald (1979) e McDonald (1981) e de apresentar um procedimento bayesiano para a análise e a comparação entre esses modelos.

2 REFERENCIAL TEÓRICO

2.1 Modelos não-lineares

Avaliar uma possível relação entre uma variável dependente e várias variáveis independentes é uma tarefa comum em análises estatísticas e pode ser feita por meio de modelos de regressão. Com relação à forma da função (parte determinística do modelo), Draper & Smith (1998) classificam os modelos de regressão em: modelos lineares, aqueles que são lineares em relação aos parâmetros; modelos linearizáveis, aqueles que podem ser transformados em lineares por meio de alguma transformação e modelos não-lineares, aqueles em que pelo menos uma das derivadas parciais depende de algum parâmetro do modelo.

Segundo Mazucheli & Achcar (2002), um modelo de regressão não-linear é considerado intrinsecamente linear se ele puder ser reduzido a um modelo linear por meio de uma reparametrização apropriada. Pode-se, ainda, usar o termo intrinsecamente linear para referir-se a modelos que podem ser linearizados via alguma transformação como, por exemplo, a logarítmica. Em geral, na prática, um modelo não-linear é linearizado para facilitar a obtenção de estimativas iniciais dos parâmetros, para o ajuste dos modelos não lineares. O inconveniente de algumas transformações é que, além de o parâmetro perder sua interpretação, pode-se alterar a estrutura e a distribuição do erro, ou seja, se os erros do modelo original satisfizerem as suposições usuais de normalidade, independência e homogeneidade da variância, os erros do novo modelo, em geral, não satisfarão tais suposições (Khuri & Cornell, 1987). Caso não seja possível obter uma reparametrização ou uma transformação apropriada, que reduza o modelo à forma linear, têm-se os chamados modelos intrinsecamente não-lineares.

Souza (1998) considera o modelo de regressão não-linear na forma: $y_i = f(x_i, \theta^0) + \varepsilon_i$, em que: $i = 1, \dots, n$; n é o número de observações; $f(x_i, \theta^0)$ tem forma funcional conhecida; x_i é um vetor k dimensional formado por observações em variáveis exógenas, $\theta^0 \in \Theta$ é um parâmetro p dimensional pertencente ao espaço paramétrico Θ e ε_i é um erro experimental não observável diretamente. O mesmo autor cita que, de forma semelhante aos modelos lineares, o processo de estimação de um parâmetro, em um modelo não-linear, pode ser obtido pela minimização da soma de quadrados dos erros, obtendo-se o sistema de equações normais não-linear, o qual não apresenta uma solução explícita para a estimativa do parâmetro, que é obtida por processos iterativos.

Chiacchio (1993) distingue a caracterização da regressão em função das suposições do vetor de erros da seguinte maneira: modelos ordinários, aqueles cuja estrutura dos erros não viola as pressuposições de independência, normalidade e homogeneidade de variâncias; modelos ponderados, aqueles cuja estrutura dos erros viola a pressuposição de homogeneidade de variâncias e modelos generalizados, aqueles cuja estrutura dos erros viola a pressuposição de independência dos erros e, possivelmente, a de homogeneidade de variâncias. Nesse caso, diz-se que os erros são correlacionados.

Savian & Muniz (2007) avaliaram, por meio da metodologia clássica, o ajuste do modelo de regressão não-linear de Mertens & Loften (1980) aos dados de resíduo potencialmente degradável da fibra em detergente neutro da gramínea coastcross, em duas idades de corte. No ajuste do modelo, aqueles autores assumiram erros correlacionados e concluíram que a consideração de uma estrutura de erros auto-regressivos de segunda ordem melhorou o ajuste do modelo, promovendo estimativas mais precisas para os parâmetros.

Savian et al. (2007) estudaram, também por meio da metodologia clássica, o ajuste do modelo de regressão não-linear proposto por Orskov & McDonald (1979) aos dados de porcentagem de degradação da fibra em detergente neutro da gramínea coastcross avaliada em 12 idades de corte. Nesse trabalho, os autores assumiram erros independentes com distribuição normal de média zero e variâncias homogêneas e concluíram que, nas idades de corte avançadas, as estimativas dos parâmetros foram menos precisas.

2.2 Modelos empíricos de degradabilidade ruminal

A técnica *in situ*, com sacos de náilon e proposta por Mehrez & Orskov (1977), consiste em determinar o desaparecimento de componentes da amostra de alimentos acondicionados em sacos de náilon e incubados no rúmen por períodos variáveis de tempo. Essa técnica foi desenvolvida em meados da década de 1930, na África do Sul, quando Quin et al. (1939) estudaram a degradação de alimentos no rúmen de ovinos canulados (Orskov, 1982).

Segundo Mertens (1993), as primeiras avaliações de processos de digestão, considerando-se os tempos de retenção, eram qualitativas e baseavam-se na interpretação visual de curvas de digestão, sendo de difícil descrição, pois estas curvas mostravam comportamentos não-lineares. O autor relata que Waldo, em 1970, foi o primeiro a sugerir uma inovação conceitual, que serviu como base para uma nova visão dos modelos matemáticos relacionados com a cinética de digestão, com os quais se pretendia uma quantificação real do fato.

Vários modelos foram propostos para descrever a digestão e a passagem de alimentos pelo trato digestivo do ruminante (Mertens & Ely, 1979; France et al., 1982; Baldwin et al., 1987; Mertens, 1993). Nesses modelos, os resíduos de alimento desaparecem do trato digestivo por degradação e absorção ou por passagem para as fezes. A proporção de nutrientes consumidos que se torna

disponível para os ruminantes é o resultado da competição entre digestão e passagem. A taxa e a extensão da digestão são variáveis críticas na descrição do processo digestivo. Esses parâmetros cinéticos são importantes porque eles não só descrevem a digestão, mas também caracterizam as propriedades intrínsecas dos alimentos que limitam sua disponibilidade aos ruminantes (Mertens, 1993).

De acordo com Mehrez & Orskov (1977), os modelos exponenciais são apropriados para o estudo da degradação das forrageiras. Orskov & McDonald (1979) sugeriram um modelo exponencial para auxiliar no estudo da degradação de forrageiras no rúmen. Uma equação baseada no princípio de redução do material durante a incubação no estudo da degradação em função do tempo de incubação no rúmen, foi proposta e é dada por:

$$Deg_i = a + b(1 - e^{-ct_i}) + \varepsilon_i, \quad (1)$$

em que: Deg_i é a degradação do alimento até o tempo i ; t_i é o tempo de incubação das amostras, com $i = 1, \dots, n$; n é o número de coleta das amostras; a é a fração solúvel do componente nutritivo analisado; b é a fração potencialmente degradável do componente nutritivo analisado; c é a taxa de degradação da fração b e ε_i é erro aleatório com distribuição normal de média zero e variância σ^2 .

Entretanto, determinadas frações de alguns alimentos apresentam um período de latência (L), ou “lag time”, em que não se verifica a degradação do substrato. Durante esse período, pode ocorrer hidratação das partículas do alimento, remoção de substâncias inibidoras, eventos ligados à adesão e efetiva colonização das partículas do alimento pelos microrganismos ruminais, de modo que, antes do término desta fase, o alimento permanece inalterado no rúmen, a não ser por ação mecânica. Se o tempo de colonização não for considerado no

modelo, a assíntota é superestimada e a taxa de degradação é subestimada (Stensig et al., 1994).

Desse modo, é justificável a introdução de novos modelos que determinem a contribuição do período de latência (L) sobre a cinética de degradação ruminal (Mertens, 1977; McDonald, 1981; Pereira, 1992). Com base nisso, McDonald (1981) revisou o modelo e propôs a inclusão do tempo de colonização. Esse modelo alternativo foi denominado de modelo revisado para estimação de degradabilidade no rúmen e é dado por:

$$Deg_i = a + b\left(1 - e^{-c(t_i - L)}\right) + \varepsilon_i, \quad \text{para } t_i > L, \quad (2)$$

em que: Deg_i , a , b , c , t_i e ε_i foram definidos anteriormente e L é o *lag time*.

Os parâmetros a e b dos modelos de Orskov & McDonald (1979) e McDonald (1981) são utilizados para calcular a degradabilidade potencial (DP), dada por:

$$DP = a + b,$$

que representa a quantidade de alimento que pode se solubilizar ou degradar dentro do rúmen se o tempo não for um fator limitante. Orskov (1982) cita que, em condições normais, algumas partículas do alimento estão sujeitas à passagem ao abomaso antes de serem completamente degradadas. Logo, a digestão é resultante de duas forças competitivas que atuam simultaneamente, a taxa de passagem (k) e a taxa de degradação da fração potencialmente degradável (c). A inclusão da taxa de passagem no cálculo da degradação origina a degradabilidade efetiva (DE), podendo ser calculada pela expressão:

$$DE = a + \frac{(bc)}{(c + k)}$$

No caso da existência de um tempo de colonização antes do início da degradação da fração potencialmente degradável, McDonald (1981) sugere a

inclusão deste componente na equação. Segundo Orskov (1982), esta modificação resulta em pouca diferença no valor da DE.

O Agricultural and Food Reserch Council (AFRC, 1992) recomenda, para taxa de passagem (k), o valor de 2%/hora, para animais recebendo ração completamente moída e ou em baixo nível alimentar, 5%/hora para vacas leiteiras de baixa produção (menos de 15kg de leite/dia) ou gado de corte recebendo alto nível de dietas mistas (menos que 2 vezes a manutenção) e 8%/hora para vacas leiteiras de alta produção (mais que 15kg de leite/dia), recebendo dietas mistas (mais que 2 vezes a manutenção).

Os modelos de degradabilidade ruminal descritos nas equações (1) e (2) são classificados como modelos de regressão não-linear. Para esta classe de modelos as propriedades dos estimadores dos parâmetros somente são válidas assintoticamente, isto é, quando o tamanho da amostra é suficientemente grande. Em geral, em pequenas amostras, essas propriedades são desconhecidas (Searle, 1971).

Os trabalhos que abordam a estimação de parâmetros pelo método bayesiano demonstram grande sucesso, pelo fato de todos os parâmetros serem utilizados no modelo como variáveis aleatórias. Isso requer um menor número de observações, pois os conceitos probabilísticos envolvidos diminuem a dependência do ajuste do modelo em relação ao número de dados utilizados, uma vez que o conceito de graus de liberdade não é utilizado (Silva et al., 2005, 2006), além de possibilitar a incorporação de conhecimento *a priori* com relação aos parâmetros por meio de experiências anteriores do pesquisador na área ou em publicações sobre o assunto que se deseja estudar.

2.3 Inferência bayesiana

A metodologia bayesiana ficou resguardada durante um grande período, por necessitar de resoluções matemáticas, mais precisamente de integrações, inviáveis de serem feitas algebricamente. Por volta da década de 1960 ela ressurgiu em alguns trabalhos teóricos, como o de Jeffreys (1961), mas somente em 1990, com o trabalho de Gelfan & Smith (1990), que exploraram um recurso de simulação dinâmica, o algoritmo Gibbs Sampler, é que o problema das integrações foi solucionado de maneira alternativa. Atualmente, principalmente devido ao avanço computacional, a metodologia bayesiana é um dos principais assuntos da comunidade científica envolvida com o desenvolvimento e a aplicação de procedimentos estatísticos.

Em linhas gerais, a metodologia bayesiana consiste de informações referentes aos dados amostrais (função de verossimilhança), do conhecimento prévio a respeito dos parâmetros (distribuição *a priori*) e da obtenção da densidade *a posteriori* dos parâmetros por meio do teorema de Bayes. A utilização de informação *a priori* requer a especificação de uma distribuição para o parâmetro, devendo esta representar probabilisticamente o conhecimento que se tem sobre ele, antes da realização do experimento.

Broemeling (1989) e Gelman et al. (1997) relatam que a característica essencial da inferência bayesiana está no uso explícito de probabilidades para quantificar as incertezas que se tem sobre o parâmetro. Portanto, a inferência bayesiana é baseada no conceito de probabilidade subjetiva, que mede o grau de confiança que alguém deposita no acontecimento de um determinado evento do espaço amostral.

2.3.1 Teorema de Bayes

Segundo Paulino et al. (2003), a idéia de definir probabilidade como grau de confiança em uma dada proposição que não se sabia ser verdadeira ou falsa foi introduzida por Bernoulli, em 1713. A idéia de coerência em um sistema de graus de credibilidade parece ter sido introduzida por Ramsey, em 1926.

O teorema de Bayes, apresentado na expressão (3), é uma proposição sobre probabilidades condicionadas, dado por:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_i P(B|A_i)P(A_i)}. \quad (3)$$

Uma interpretação importante consiste em considerar, relativamente à dada situação, os A_i , com $i = 1, \dots, m$ como “antecedentes” a que o pesquisador atribui credibilidade ou probabilidades *a priori* $P(A_i)$, de natureza subjetiva. Depois da informação adicional, que consiste em saber que o evento B se realizou (este evento pode ser a observação de um conjunto de dados, ou seja, uma amostra), o pesquisador revê as suas probabilidades *a priori* por meio da fórmula de Bayes e passa a atribuir aos A_i as probabilidades *a posteriori* $P(A_i|B)$, $i = 1, \dots, m$.

O teorema de Bayes é um dos resultados da matemática que se propõem a caracterizar a aprendizagem com a experiência, isto é, modificar a atitude inicial em relação aos “antecedentes”, depois de ter a informação adicional de que certo evento se realizou (Paulino et al., 2003).

A informação de que dispomos sobre um parâmetro (θ), resumida probabilisticamente por meio da distribuição *a priori*, $p(\theta)$, pode ser aumentada observando-se uma variável aleatória X relacionada com θ . Esta

relação é definida pela distribuição amostral $p(x|\theta)$. Para um valor fixo de x , a função $L(\theta; x) = p(x|\theta)$ fornece a plausibilidade ou a verossimilhança de cada um dos possíveis valores de θ . O teorema de Bayes é a regra de atualização utilizada para quantificar o aumento de informação com respeito a θ , após observar $X = x$, dada pela seguinte expressão (Ehlers, 2007):

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

Note que $1/p(x)$, que não depende de θ , funciona como uma constante normalizadora de $p(\theta|x)$. Assim, as duas fontes de informação, distribuição *a priori* de θ , $p(\theta)$, e função de verossimilhança, $L(\theta; x) = p(x|\theta)$, são combinadas, levando à distribuição *a posteriori* de θ , $p(\theta|x)$. Uma forma usual do teorema de Bayes é dada por:

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

em que: \propto é o símbolo de proporcionalidade.

2.3.2 Distribuições *a priori*

A utilização da informação *a priori*, em inferência bayesiana, requer a especificação de uma distribuição *a priori* para a quantidade de interesse θ . Esta distribuição deve representar, probabilisticamente, o conhecimento que se tem sobre θ , antes da realização do experimento (Ehlers, 2007). A questão de como obter e quantificar essa informação, de natureza subjetiva, de modo a transformá-la em uma distribuição que possa ser utilizada para prosseguir com a metodologia bayesiana, é um tema tratado abundantemente na literatura científica (Paulino et al., 2003).

Esses mesmos autores consideram que, em algumas situações em que não existam informações *a priori* palpáveis (freqüentemente chamado “estado de ignorância *a priori*”) ou em que o conhecimento *a priori* é pouco significativo relativamente à informação amostral (ou “estado de conhecimento vago” ou “difuso”), o pesquisador é conduzido a adotar distribuições *a priori* minimamente informativas, também chamadas de distribuições *a priori* não-informativas.

Segundo Box & Tiao (1992), a *a priori* não-informativa não representa, necessariamente, o desconhecimento do pesquisador sobre o experimento, sendo utilizada também de forma a viabilizar a inferência *a posteriori*.

O argumento primeiramente invocado para gerar distribuições não informativas foi o Princípio da Razão Insuficiente de Bayes-Laplace (Paulino et al., 2003). De acordo com este princípio, na ausência de razão suficiente para privilegiar uma possibilidade em detrimento de outras, decorrente da escassez de informação *a priori*, deve-se adotar a equiprobabilidade.

Respondendo ao problema de a distribuição *a priori* não conseguir traduzir fielmente uma informação *a priori*, surgiu a metodologia Bayes hierárquica, que modela a incerteza nos hiperparâmetros por meio de novas distribuições, induzindo assim uma decomposição da distribuição *a priori* em dois ou mais níveis (Paulino et al., 2003).

A distribuição *a priori* de θ depende dos valores dos hiperparâmetros ϕ e podemos escrever $p(\theta|\phi)$ em vez de $p(\theta)$. Além disso, em vez de fixar valores para os hiperparâmetros, é possível especificar uma distribuição *a priori* $p(\phi)$, completando, assim, o segundo estágio na hierarquia. Teoricamente, não há limitações quanto ao número de estágios, mas, devido às complexidades resultantes, as *prioris* hierárquicas são especificadas, em geral, em 2 ou 3 estágios (Ehlers, 2007).

2.3.3 Distribuição *a posteriori* e estimação bayesiana

Para se inferir com relação a qualquer elemento de um vetor de parâmetros, a distribuição *a posteriori* conjunta, $p(\theta_i, \theta_{-i} | y)$, deve ser integrada em relação a todos os outros elementos que o constituem (Rosa, 1998). Assim, se o interesse do pesquisador se concentra sobre determinado elemento de θ (vetor paramétrico), por exemplo, θ_i , tem-se a necessidade da obtenção da distribuição marginal *a posteriori*, $p(\theta_i | y)$, dada pela expressão:

$$p(\theta_i | y) = \int \dots \int p(\theta_i, \theta_{-i} | y) d\theta_{-i},$$

em que: $\theta_{-i} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p)$ é o conjunto complementar de parâmetros para θ_i e p é o número de parâmetros do modelo. Porém, muitas vezes é impossível obter expressões explícitas para as integrais envolvidas na obtenção dessa distribuição e a inferência exata somente será possível se estas integrais puderem ser calculadas analiticamente. Caso contrário, os métodos aproximados deverão ser utilizados.

A distribuição marginal *a posteriori* de um parâmetro (θ) contém toda a informação probabilística a respeito deste parâmetro. No entanto, algumas vezes, é necessário resumir a informação contida nesta distribuição por meio de alguns poucos valores numéricos. Um caso simples é a estimação pontual de θ , em que se resume a distribuição marginal *a posteriori* por meio de um único número, $\hat{\theta}$. É importante também associar alguma informação sobre o quão precisa é a especificação deste número. As medidas de incerteza mais usuais são: a variância e o coeficiente de variação para a média *a posteriori*, a medida de informação observada de Fisher para a moda *a posteriori* e a distância entre quartis para a mediana *a posteriori* (Ehlers, 2007).

A própria idéia de estimação de parâmetros por pontos conduz, no cenário bayesiano, a tomar como estimativas os pontos críticos da distribuição *a posteriori*. A escolha das estimativas bayesianas de θ depende, naturalmente, da forma da distribuição *a posteriori*, $p(\theta|x)$, bem como dos objetivos do seu uso.

As estimativas mais usadas são:

- Moda *a posteriori*: $\hat{\theta}$ tal que, $p(\theta|x) = \max_{\theta \in \Theta} p(\theta|x) = \max_{\theta \in \Theta} \{p(\theta)l(\theta|x)\}$;
- Média *a posteriori*: $\hat{\theta} = E[\theta|x]$ com, $E[\theta_i|x] = \int_{\Theta} \theta_i p(\theta|x) d\theta$, $i = 1, \dots, k$; k

é o número de parâmetros;

- Vetor das medianas *a posteriori*: $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_k)$ tal que, $P\{\theta_i \geq \hat{\theta}_i | x\} \geq 1/2$ e $P\{\theta_i \leq \hat{\theta}_i | x\} \leq 1/2$, $i = 1, \dots, k$; k é o número de parâmetros.

Um resumo de $p(\theta|x)$ mais informativo do que qualquer estimativa pontual é obtido de uma região do espaço paramétrico, Θ , que contenha uma parte substancial da massa probabilística *a posteriori* (Paulino et al., 2003).

Ehlers (2007) define C como um intervalo de credibilidade de $100(1-\alpha)\%$ para θ se $P(\theta \in C) \geq 1-\alpha$. Assim, quanto menor for o tamanho do intervalo, mais concentrada é a distribuição do parâmetro, ou seja, o tamanho do intervalo informa sobre a dispersão de θ .

É possível construir uma infinidade de intervalos utilizando a definição acima, mas é de principal interesse aquele com menor comprimento possível. Os intervalos de comprimento mínimo são obtidos tomando-se os valores de θ com maior densidade *a posteriori*, denominados intervalos de credibilidade de máxima densidade *a posteriori*, ou intervalos HPD.

Segundo Ehlers (2007), um intervalo de credibilidade C de $100(1-\alpha)\%$ para θ é de máxima densidade *a posteriori* se $C = \{\theta \in \Theta : p(\theta|y) \geq k(\alpha)\}$, em que $k(\alpha)$ é a maior constante tal que $P(\theta \in C) \geq 1 - \alpha$.

Usando esta definição, todos os pontos dentro do intervalo de máxima densidade *a posteriori* (intervalo HPD) terão densidade maior do que qualquer ponto fora do intervalo.

2.4 Métodos de simulação

Os métodos aproximados, baseados em simulação estocástica, consistem no tratamento de problemas reais a partir de reproduções em ambientes controlados pelo pesquisador. Alguns problemas apresentam componentes aleatórios, os quais não podem ser descritos de forma exata e, sim, baseados em informações probabilísticas. Para estes problemas, o processo de simulação é estocástico, ou seja, baseado em distribuições de probabilidades.

Segundo Ehlers (2007), existem métodos de simulação não iterativos em que as amostras são obtidas da distribuição *a posteriori* em um único passo. São eles os métodos de Monte Carlo (ordinário e via função de importância) e os métodos de reamostragem (método da rejeição e método da reamostragem ponderada). Os métodos que fazem uso de técnicas de simulação iterativa são os métodos de Monte Carlo via cadeias de Markov (amostrador de Gibbs e o algoritmo de Metropolis-Hastings, por exemplo).

2.4.1 Métodos de Monte Carlo via cadeias de Markov

A idéia básica desses métodos é a de transformar o problema estático, em consideração, num problema de natureza dinâmica, construindo, para tal,

efeito um processo estocástico temporal, artificial, que seja fácil de simular e que convirja para a distribuição marginal (Paulino et al., 2003).

Metropolis et al. (1953) desenvolveram um algoritmo para esse efeito, o qual foi, mais tarde, generalizado por Hastings (1970), sendo atualmente conhecido na literatura por algoritmo de Metropolis-Hastings.

Um caso particular destes métodos é o método de amostragem Gibbs, o qual foi introduzido por Geman & Geman (1984) para simular distribuições multivariadas, de natureza bastante complexa, que surgem em problemas de reconstrução de imagens. Gelfand & Smith (1990) mostraram como o algoritmo pode ser usado para simular distribuições marginais *a posteriori* e, conseqüentemente, a sua importância na resolução de problemas em inferência bayesiana.

2.4.1.1 Algoritmo de Metropolis-Hastings

O algoritmo de Metropolis-Hastings usa a idéia de gerar um valor de uma distribuição auxiliar, o qual é aceito com uma dada probabilidade. Este mecanismo de correção garante a convergência da cadeia para a distribuição de equilíbrio que, neste caso, é a distribuição marginal *a posteriori* (Ehlers, 2007).

Suponha que a cadeia esteja em um estado m e que θ é o valor amostrado, um valor θ' é gerado de uma distribuição proposta $q(\cdot|\theta)$. O novo valor θ' é aceito com probabilidade:

$$\alpha(\theta, \theta') = \min \left(1, \frac{\pi(\theta')q(\theta|\theta')}{\pi(\theta)q(\theta'|\theta)} \right),$$

em que π é a distribuição de interesse.

Em termos práticos, o algoritmo de Metropolis-Hastings pode ser especificado pelos seguintes passos:

- 1) inicialize o contador de iterações $m=0$ e especifique um valor inicial $\theta^{(0)}$;
- 2) gere um novo valor θ' da distribuição $q(\cdot|\theta)$;
- 3) calcule a probabilidade de aceitação $\alpha(\theta, \theta')$ e gere $u \sim U(0,1)$;
- 4) se $u < \alpha(\theta, \theta')$, então, aceite o novo valor, θ' , e faça $\theta^{(m+1)} = \theta'$, caso contrário rejeite e faça $\theta^{(m+1)} = \theta$;
- 5) incremente o contador de m para $m+1$ e volte ao passo 2.

Embora a distribuição proposta possa ser escolhida arbitrariamente na prática devem-se tomar alguns cuidados para garantir a eficiência do algoritmo. Em aplicações bayesianas a distribuição de interesse é a própria distribuição marginal *a posteriori*.

Um caso particular é chamado algoritmo de Metropolis e considera apenas distribuições propostas do tipo simétricas, ou seja, $q(\theta'|\theta) = q(\theta|\theta')$, para todos os valores de θ e θ' . Neste caso, a probabilidade de aceitação se reduz para:

$$\alpha(\theta, \theta') = \min \left(1, \frac{\pi(\theta')}{\pi(\theta)} \right).$$

2.4.1.2 Amostrador de Gibbs

O algoritmo Gibbs, ou método de amostragem Gibbs, é baseado no fato de que se a distribuição *a posteriori* conjunta $p(\theta|y)$ for positiva em $\Theta_1 \times \dots \times \Theta_k$, com Θ_i suporte da distribuição de θ_i para $i = 1, \dots, k$, em que k é o número de parâmetros, então, ela é unicamente determinada pelas distribuições condicionais completas $p(\theta_i|y, \theta(-i))$, $i = 1, \dots, k$, em que $\theta = (\theta_1, \dots, \theta_k)$ e

$\theta(-i) = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$, ou seja, é o vetor θ sem a i -ésima componente (Paulino et al., 2003). Esses autores descrevem o algoritmo da seguinte maneira:

Seja $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$ um valor arbitrário inicial para o vetor θ .

Procede-se iterativamente da seguinte forma:

- 1) obtém-se $\theta_1^{(1)}$ de $p(\theta_1 | x, \theta_2^{(0)}, \dots, \theta_k^{(0)})$,
 obtém-se $\theta_2^{(1)}$ de $p(\theta_2 | x, \theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)})$,
 obtém-se $\theta_3^{(1)}$ de $p(\theta_3 | x, \theta_1^{(1)}, \theta_2^{(1)}, \theta_4^{(0)}, \dots, \theta_k^{(0)})$,
 ...
 obtém-se $\theta_k^{(1)}$ de $p(\theta_k | x, \theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}, \dots, \theta_{k-1}^{(1)})$.

Completa-se, assim uma iteração do esquema e uma transição de θ^0 para $\theta^1 = (\theta_1^{(1)}, \dots, \theta_k^{(1)})$.

- 2) O esquema anterior é repetido com θ^1 , obtido anteriormente, como vetor inicial para obter um novo vetor θ^2 e, assim, haver uma transição de θ^1 para $\theta^2 = (\theta_1^{(2)}, \dots, \theta_k^{(2)})$.
- 3) Repete-se m vezes este ciclo de geração de observações aleatórias de cada uma das distribuições condicionais, produzindo, assim, $\theta^0, \dots, \theta^m$.

A sucessão $\theta^0, \dots, \theta^m$ é uma realização de uma cadeia de Markov com espaço de estados Θ . Resultados teóricos permitem concluir que, quando $t \rightarrow \infty$, $(\theta_1^{(m)}, \dots, \theta_k^{(m)})$, tende em distribuição para um vetor aleatório cuja função densidade de probabilidade conjunta é $p(\theta | y)$.

De acordo com o exposto, o método de amostragem Gibbs permite, por meio de um processo iterativo, gerar realizações de uma cadeia de Markov, de

tal modo que, à medida que o número de iterações aumenta, a cadeia se aproxima da sua condição de equilíbrio. No entanto, sucessivas realizações de uma mesma cadeia ao longo do tempo não constituem uma amostra aleatória da distribuição que se pretende gerar, pois os vetores θ^m que vão sendo gerados são correlacionados.

Uma idéia é utilizar uma única e longa cadeia a partir de um estado inicial, designado por $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$ e com comprimento $m=l+k^*n$, em que: m é o número de iterações, l é o número inicial de iterações necessárias para que a cadeia atinja o estado de equilíbrio; n é o número das iterações que vão ser usadas na aplicação do método de Monte Carlo e k^* é o espaçamento entre iterações sucessivas delineado para eliminar a autocorrelação entre consecutivas iterações.

Para fazer inferências, extrai-se, da realização original da cadeia, um subconjunto de dimensão q , contendo as observações $\theta^{l+k^*}, \theta^{l+2k^*}, \dots, \theta^{l+qk^*}$. Desse modo obtém-se uma amostra com componentes $\theta_{(j)} = \theta^{l+jk^*}$, com base na qual serão realizadas as inferências (Paulino et al., 2003).

2.4.2 Diagnósticos de convergência

Como descrito anteriormente, a idéia dos métodos de Monte Carlo via cadeias de Markov é a de transformar o problema estático, em consideração, num problema de natureza dinâmica, construindo, para isto, um processo estocástico temporal, artificial, que seja fácil de simular e que convirja para a distribuição original (Paulino et al., 2003). A questão é: como monitorar a convergência das cadeias?

Alguns métodos são utilizados para diagnosticar esta convergência, sendo eles classificados como métodos informais e formais. Como método de

monitoramento informal tem-se o método proposto por Gelfand & Smith (1990) e, como métodos de monitoramento formal, têm-se, dentre os mais utilizados, aqueles propostos por Gelman & Rubin (1992), Raftery & Lewis (1992) e Geweke (1992) e o de Heidelberg & Welch (1983).

Gelfand & Smith (1990) sugeriram, como proposta inicial, uma verificação informal da convergência baseada em representações gráficas. Os autores recomendam processar um número de iterações que se acredita serem suficientes (N). Com estes resultados, um histograma é gerado. Em seguida, mais Z iterações são realizadas e, em seguida, cria-se um novo histograma ($N+Z$). A comparação do gráfico anterior com o último é realizada. Visualmente, verifica-se se há diferenças; caso não existam, assume-se que a seqüência estabilizou e o processo convergiu. A verificação informal da convergência pode ser realizada, desde que seja feita de forma combinada com os métodos de monitoramento formal.

Gelman & Rubin (1992) propuseram um método que se baseia no uso de componentes de variância de seqüências múltiplas da cadeia, simuladas a partir de pontos dispersos. O método preconiza que sejam simuladas $m \geq 2$ seqüências, cada uma de comprimento $2n$ e com pontos iniciais simulados de uma distribuição sobredispersa com relação à distribuição de equilíbrio. As primeiras n iterações, de cada cadeia, são descartadas (*burn in*) para evitar a influência dos valores iniciais. O método resulta em um fator (R), chamado fator de redução de escala potencial e pode ser interpretado como um fator de diagnóstico de convergência. Quando a estimativa de R for grande, sugere-se que qualquer das estimativas das variâncias pode ser reduzida com a utilização de mais iterações. Se a estimativa de R aproxima-se de um, pode-se concluir que cada m conjunto de n iterações se aproxima da distribuição de interesse.

O critério proposto por Raftery & Lewis (1992) é um método que estima quantas iterações são necessárias para que o amostrador de Gibbs apresente

convergência à distribuição estacionária, propondo quantidades necessárias para isso. O método fornece as estimativas do *burn-in*, o número total de iterações que devem ser computadas e a distância mínima de uma iteração à outra (*thin*), para se obter a subamostra aproximadamente independente. A regra de decisão do critério de Raftery & Lewis (1992) é baseada no fator de dependência, que é responsável pelo acréscimo multiplicativo ao número de iterações necessárias para se alcançar a convergência. De acordo com os autores, se o fator de dependência for maior que cinco, pode-se concluir que a cadeia não atingiu convergência.

O critério proposto por Geweke (1992) consiste em indicar a convergência da média *a posteriori* da quantidade amostrada, $g(\theta)$, em uma única cadeia, isto é, os valores $g(\theta^{(i)})$ são calculados a cada iteração da simulação, formando uma série temporal. Em um primeiro passo, divide-se a cadeia de tamanho N em duas seqüências. A primeira seqüência possui N_a primeiras iterações e a segunda, N_b restantes iterações. O autor sugere que as médias devem ser estimadas após o descarte de algumas iterações iniciais e que sejam considerados $N_a = 0,1N$ e $N_b = 0,5N$. Em um segundo passo, estima-se as médias \hat{g}_a e \hat{g}_b e as variâncias assintóticas $\hat{S}_{g_a}^2(0)$ e $\hat{S}_{g_b}^2(0)$, respectivamente. As variâncias são determinadas pela estimação da densidade espectral. Se a diferença padronizada entre as médias for muito grande, existe indicação de ausência de convergência. A regra de decisão do critério de Geweke é baseada na estimativa de um valor p . Se esse valor for menor que um nível de significância adotado pelo pesquisador, conclui-se que a cadeia não atingiu convergência.

O critério de Heidelberger & Welch (1993) utiliza testes estatísticos para avaliar a hipótese nula de estacionariedade da amostra gerada. Se a hipótese nula

for rejeitada para um dado valor de descarte das observações iniciais, o teste é repetido depois de descartados os 10% valores iniciais da amostra. Se a hipótese é novamente rejeitada, mais 10% dos valores iniciais são descartados, e assim sucessivamente, até serem descartados os 50% valores iniciais. Se a hipótese for novamente rejeitada, isso indica que é necessário um número maior de iterações. Caso contrário, o número de iterações descartadas é indicado como o tamanho do *burn-in*.

Uma avaliação detalhada dos critérios de convergência para o método de Monte Carlo via Cadeias de Markov pode ser encontrado em Nogueira et al. (2004). Os autores concluíram que o seguinte procedimento deveria ser utilizado para que se obtivesse uma avaliação mais precisa da convergência: aplicar o critério de Raftery e Lewis em uma amostra piloto e determinar o tamanho ideal da seqüência; monitorar a convergência das seqüências nas proximidades do tamanho ideal, indicado pelo critério de Raftery e Lewis, por meio dos critérios de Gelman e Rubin e Geweke e determinar o tamanho de *burn-in*, pelo critério de Heidelberger e Welch.

2.4.3 Erro de Monte Carlo

No contexto de inferência bayesiana, o método de Monte Carlo é utilizado para o cálculo de integrais definidas, utilizando-se aproximações.

Paulino et al. (2003), sugerem considerar o problema de aproximar uma integral da forma:

$$\int g(\theta)h(\theta|x)d\theta = E[g(\theta)|x], \quad (4)$$

em que: θ e x podem ser vetores, cuja existência se admite. Muitas quantidades a posteriori de interesse podem ser expressas pela equação acima, para algum tipo de função $g(\theta)$ integrável. É o caso de médias *a posteriori*, θ_i , de

componentes de um vetor de parâmetros θ . Se for possível simular uma amostra aleatória $(\theta_1, \dots, \theta_m)$ da densidade a posteriori $h(\theta|x)$, o Método de Monte Carlo aproxima a integral (4) pela média empírica, dada por:

$$\hat{E}[g(\theta)|x] = \frac{1}{m} \sum_{i=1}^m g(\theta_i),$$

a qual pela Lei Forte dos Grandes Números, converge quase certamente para $E[g(\theta)|x]$. A precisão desta aproximação pode ser medida pelo erro padrão de Monte Carlo, dado por:

$$\frac{1}{\sqrt{m(m-1)}} \left\{ \sum_{i=1}^m \left[g(\theta_i) - \frac{1}{m} \sum_{i=1}^m g(\theta_i) \right]^2 \right\}^{1/2},$$

isto é, a aproximação pode ser tão acurada quanto se deseja, bastando aumentar o valor de m (número de iterações).

Este método consiste em gerar vários números aleatórios uniformemente distribuídos no intervalo de interesse, avaliando-se o valor da função a ser integrada em cada um destes pontos e, finalmente, calculando-se a média desses valores obtidos para se obter a aproximação da integral definida. Como o método de Monte Carlo é uma aproximação, não se pode afirmar que o valor encontrado por ele seja o valor exato da integral. Entretanto, quanto mais pontos forem gerados, maior será a confiabilidade na aproximação obtida.

2.4.4 Fator de Bayes

Existem diversos critérios, descritos na literatura, para a seleção de modelos. Alguns deles são: Fator de Bayes, Pseudofator de Bayes, Critério de Informação de Akaike, Critério de Informação Bayesiano e Critério de Informação Deviance. Estes três últimos são baseados em aproximações ao

Fator de Bayes e têm como objetivo considerar, no processo de seleção, a complexidade dos modelos (Paulino et al., 2003).

Suponha-se que existam m modelos diferentes, M_i , para $i = 1, \dots, m$, que tentam explicar y . Os modelos M_i dependem de parâmetros θ_i . Portanto, a densidade *a posteriori* para os parâmetros calculados sob M_i pode ser expressa como:

$$p(\theta^i | y, M_i) = \frac{p(y | \theta^i, M_i) p(\theta^i | M_i)}{p(y | M_i)},$$

em que: M_i é o i -ésimo modelo a ser ajustado, y é o conjunto de dados e θ^i é o vetor de parâmetros relacionado ao i -ésimo modelo, M_i .

A lógica da inferência bayesiana sugere que seja utilizado o teorema de Bayes para estimar uma probabilidade sobre o que não se conhece (qual o modelo é mais consistente com os dados), condicional ao que se conhece (os dados). Tal probabilidade é denominada probabilidade *a posteriori* do modelo.

Utilizando-se o teorema de Bayes, tem-se que:

$$p(M_i | y) = \frac{p(y | M_i) p(M_i)}{p(y)},$$

em que: $p(y | M_i)$ é a verossimilhança marginal do i -ésimo modelo, $p(M_i)$ é a probabilidade *a priori* do modelo e $p(y)$ é a distribuição marginal dos dados.

Segundo Paulino et al. (2003), um problema com o fator de Bayes é que, para muitos modelos de interesse, pelo menos parte da especificação *a priori* é vaga, de modo que a distribuição *a priori*, $p(\theta)$, é imprópria. Dado que o cálculo direto do denominador, da expressão acima, é complexo na maioria das situações, tornou-se prática comum comparar dois modelos M_i e M_j , utilizando a razão de chances *a posteriori* (RCP), isto é:

$$RCP_{ij} = \frac{p(M_i|y)}{p(M_j|y)} = \frac{p(y|M_i)p(M_i)}{p(y|M_j)p(M_j)}.$$

Em diversas situações, o pesquisador pode estar interessado em avaliar a consistência dos modelos com os dados, fornecendo a mesma probabilidade *a priori* para cada modelo. Neste caso, a razão de chances *a posteriori* entre os modelos torna-se simplesmente a razão das verossimilhanças marginais e recebe a denominação de fator de Bayes, que é definido por:

$$FB_{ij} = \frac{p(y|M_i)}{p(y|M_j)}.$$

Kass & Raftery (1995) e Raftery (1995) citam que a verossimilhança marginal pode ser obtida da seguinte forma:

$$p(y|M_i) = \int_{\Theta} l(y|\theta, M_i) p(\theta|M_i) d\theta, \quad (5)$$

em que: $l(y|\theta, M_i)$ e $p(\theta|M_i)$ são, respectivamente, a função de verossimilhança e distribuição *a priori* correspondente ao modelo M_i .

A resolução da integral apresentada na equação (5) geralmente não é analítica, mas segundo Raftery (1995) uma possível solução é considerar valores de $\theta^{(k)}$ gerados via algoritmos MCMC, pois assim é possível obter uma estimativa de $p(y|M_i)$, a qual é obtida por:

$$\hat{p}(y|M_i) = \frac{1}{Q} \sum_{q=1}^Q l(y|\theta^{(q)}, M_i),$$

em que o índice q representa cada iteração dos algoritmos MCMC.

Em relação à interpretação do fator de Bayes, pode-se dizer que esta é simples e direta, pois se FB_{12} é maior que 1, tem-se a indicação que o modelo disposto no numerador é mais plausível ou verossímil do que o modelo disposto

no denominador, caso contrário, se FB_{12} é menor que 1, o modelo no denominador é mais plausível que o modelo disposto no numerador.

Uma interpretação para o fator de Bayes é dada por Jeffreys (1961), a qual contém categorias de evidências em relação aos modelos M_i e M_j , podendo ser utilizada na escolha do melhor modelo.

3 MATERIAL E MÉTODOS

Na análise foram considerados os ajustes dos modelos de Orskov & McDonald (1979) e McDonald (1981).

Para a obtenção da distribuição *a posteriori* conjunta dos parâmetros do modelo de Orskov & McDonald (1979), dado por:

$$Deg_i = a + b(1 - \exp(-ct_i)) + \varepsilon_i, \text{ com } i = 1, \dots, n,$$

em que: n é o número de coletas amostradas e os demais termos já foram descritos na seção 2.2, assumiram-se *prioris* uniformes para os parâmetros a , b e c .

Supondo que os $\varepsilon_i = y_i - a - b[1 - \exp(-ct_i)]$ são independentes, identicamente distribuídos (*i.i.d.*) e seguem uma distribuição normal, ou seja, $\varepsilon_i \sim N(0, \sigma_e^2)$, assumiu-se como *priori* para a precisão, τ , o inverso da variância, σ_e^2 , uma distribuição Gama com hiperparâmetros α e β , isto é, $p(\tau|\alpha, \beta) \propto \tau^{\alpha-1} \exp(-\beta\tau)$. A verossimilhança para o modelo é dada por:

$$L(y|a, b, c, \tau, \alpha, \beta) \propto (\tau)^{n/2} \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-ct_i))\}^2\right],$$

em que $y = \{y_1, \dots, y_n\}$.

Por intermédio do teorema de Bayes, a distribuição *a posteriori* conjunta pode ser descrita da seguinte forma:

$$p(a, b, c, \tau|y, \alpha, \beta) \propto L(y|a, b, c, \tau, \alpha, \beta) p(a) p(b) p(c) p(\tau|\alpha, \beta).$$

Para o modelo de McDonald (1981), dado por:

$$y_i = a + b\left[1 - e^{-c(t_i-L)}\right] + \varepsilon_i, \text{ para } t > L, \text{ com } i = L, \dots, n,$$

em que: n também é o número de coletas amostradas e os demais termos já foram descritos na seção 2.2, assumiram-se *prioris* uniformes para os parâmetros a , b e c . Para o parâmetro L , considerou-se como *priori* uma função indicadora do tipo $p(L) \propto I_{(L>0)}$, uma vez que a unidade de medida deste parâmetro é dada em horas.

Sob a mesma suposição de que os erros, $\varepsilon_i = y_i - a - b[1 - \exp(-c(t_i - L))]$, são *i.i.d.*, assumiram-se, novamente, como *priori* para a precisão, τ , uma distribuição Gama com hiperparâmetros α e β , isto é, $p(\tau|\alpha, \beta) \propto \tau^{\alpha-1} \exp(-\beta\tau)$. A verossimilhança para o modelo é dada por:

$$L(y|a, b, c, L, \tau, \alpha, \beta) \propto \tau^{n/2} \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-c(t_i - L)))\}^2\right],$$

em que $y = \{y_1, \dots, y_n\}$.

Por intermédio do teorema de Bayes, a distribuição *a posteriori* conjunta pode ser descrita da seguinte forma:

$$p(a, b, c, L, \tau|y, \alpha, \beta) \propto L(y|a, b, c, L, \tau, \alpha, \beta) p(a) p(b) p(c) p(L) p(\tau|\alpha, \beta).$$

Como um primeiro passo, antes do ajuste dos modelos aos dados reais, foi feito um estudo com dados simulados, no intuito de avaliar o recurso computacional empregado e testar a metodologia utilizada. Os códigos implementados para as análises se encontram disponíveis no site www.dex.ufla.br/~muniz/

3.1 Dados simulados

Foram simulados 120 valores de porcentagem de degradação da fibra em detergente neutro (FDN), correspondendo a uma coleta por hora, a partir de

ambos os modelos. Os valores assumidos como verdadeiros para os parâmetros do modelo de Orskov & McDonald (1979) foram: a (fração solúvel do componente nutritivo) igual a 6; b (fração potencialmente degradável do componente nutritivo) igual a 40, c (taxa de degradação da fração b) igual a 0,04 e ε_i seguindo uma distribuição normal com média zero e variância 2, ou seja, $\varepsilon \sim N(0,2)$. Para o modelo de McDonald (1981), utilizaram-se os mesmos valores paramétricos para a , b , c e mesma distribuição para os erros, $\varepsilon \sim N(0,2)$. Para o parâmetro L , foi utilizado o valor paramétrico de 3 horas. Os valores paramétricos foram escolhidos com base em dados literários do comportamento da degradação da gramínea coastcross.

3.2 Dados reais

Para ilustrar a metodologia de ajuste dos modelos, foram utilizados dados parciais de um experimento de degradabilidade conduzido por Reis (2000). O experimento foi instalado segundo um delineamento em blocos casualizados com três repetições (animais) e três tratamentos (Coastcross, Tifton 85 e Tifton 68). Foi utilizada a técnica de degradabilidade *in situ* por meio do uso de sacos de náilon incubados no rúmen, que mediram 10 x 20 cm, com porosidade aproximada de 65 μm , fechados a quente em máquina seladora (Teixeira et al., 1988). As amostras das forrageiras foram colocadas nos sacos, em quantidade de matéria seca (MS) suficientes para manter a relação proposta por Nocek (1988), em torno de 10 a 20 mg MS/cm² de área superficial dos sacos. Na avaliação dos parâmetros cinéticos da degradação ruminal *in situ* foram utilizadas três vacas da raça Jersey, fistuladas no rúmen e com peso médio de 310 Kg.

A porcentagem de degradação da fibra em detergente neutro (FDN) da gramínea coastcross (*Cynodon dactylon* x *Cynodon nlemfuensis*), com idade de corte de 120 dias foi utilizadas para demonstrar o procedimento de ajuste dos modelos de Orskov & McDonald (1979) e McDonald (1981). O perfil de degradação da gramínea foi avaliado em nove tempos de incubação (0, 3, 6, 12, 24, 48, 72, 96 e 120 horas) e pode ser visualizado na Figura 1.

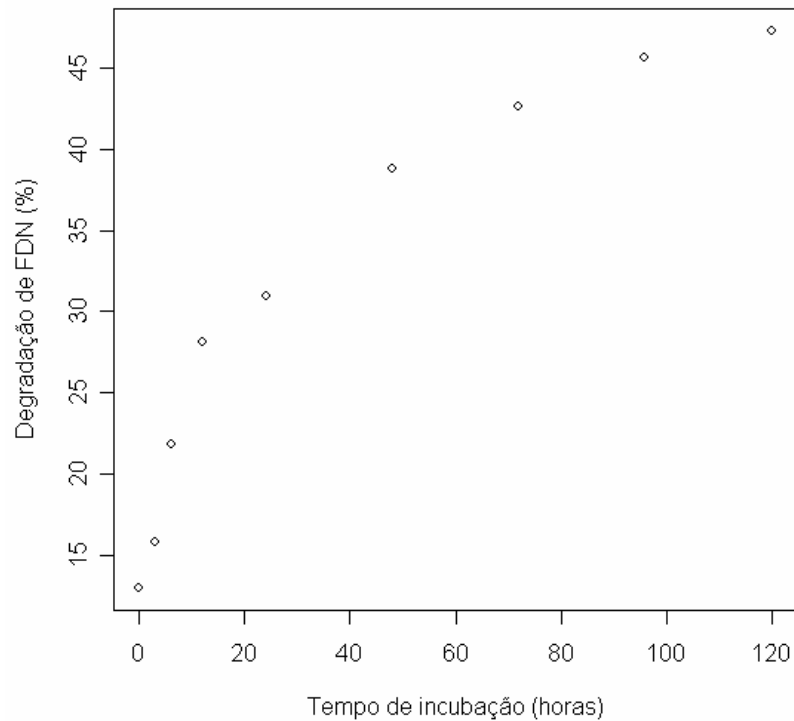


FIGURA 1. Porcentagem de degradação de fibra em detergente neutro, FDN, da gramínea coastcross, em função dos tempos de incubação, em horas.

3.3 Implementação das análises

O algoritmo para o amostrador de Gibbs e o Metropolis-Hastings foi implementado utilizando-se a linguagem R (R Development Core Team, 2007). Considerou-se, nas análises dos dados simulados, para ambos os modelos em estudo, um número fixo de 200.000 iterações, valor este sugerido pelo critério de Raftery & Lewis (1992) na análise de uma amostra piloto com 10.000 iterações, para ambos os modelos. Para o ajuste do modelo de Orskov & McDonald (1979) aos dados simulados, foi utilizado o descarte (*burn-in*) das primeiras 20.000 iterações para retirar o efeito dos valores iniciais e um espaçamento entre pontos amostrados (*thin*) de 40 observações. Para o ajuste do modelo de McDonald (1981) aos dados simulados foram descartadas (*burn-in*) as 10.000 primeiras iterações e espaçamento entre pontos amostrados (*thin*) de 30 observações. Para esses valores estipulados foi verificada a condição de convergência das cadeias por meio dos critérios de Geweke (1992), Gelman & Rubin (1992), Heidelberger & Welch (1993) e Raftery & Lewis (1992), que estão disponíveis no pacote *BOA* (“Bayesian Output Analysis”) do software R.

Para análise dos dados reais e ajuste do modelo de Orskov & McDonald (1979), também foram consideradas 200.000 iterações, descarte das primeiras 20.000 iterações e salto entre pontos amostrados de 40 observações. Para análise dos dados reais e ajuste do modelo de McDonald (1981), foram necessárias 650.000 iterações para que a convergência fosse alcançada, descarte das primeiras 650 iterações e salto entre pontos amostrados de 100 observações. Os mesmos critérios citados no ajuste aos dados simulados foram utilizados para detectar a convergência das cadeias.

Amostras das distribuições marginais *a posteriori*, para as degradabilidades efetiva e potencial, que são funções dos parâmetros do modelo de Orskov & McDonald (1979) e McDonald (1981), foram obtidas pela

substituição de a , b e c pelas suas respectivas estimativas, a cada iteração dos algoritmos, obtendo-se, assim, ao final do processo, um conjunto de valores representativos dessas distribuições. A degradabilidade efetiva foi calculada considerando-se uma taxa de passagem (k) de 5%/hora, adequada, segundo dados de literatura, para animais de médio consumo.

A comparação do modelo de Orskov & McDonald (1979) e modelo de McDonald (1981) utilizou-se o fator de Bayes (FB) utilizando os valores gerados pelos métodos MCMC para obter as estimativas da verossimilhança marginal. A interpretação do fator de Bayes seguiu a interpretação sugerida por Jeffreys (1961) e que está apresentada na Tabela 1.

TABELA 1 - Interpretação do fator de Bayes, segundo Jeffreys (1961).

Valores de $FB_{(M_i, M_j)}$	Conclusão
$FB_{(M_i, M_j)} < 1$	Evidência a favor de M_j
$1 \leq FB_{(M_i, M_j)} < 3,2$	Evidência muito fraca a favor de M_i
$3,2 \leq FB_{(M_i, M_j)} < 10$	Evidência fraca a favor de M_i
$10 \leq FB_{(M_i, M_j)} < 100$	Evidência forte a favor de M_i
$FB_{(M_i, M_j)} \geq 100$	Evidência muito forte a favor de M_i

4 RESULTADOS E DISCUSSÃO

4.1 Modelo de Orskov e McDonald (1979)

De posse da verossimilhança obtida para o modelo Orskov & McDonald (1979), das distribuições *a priori* consideradas e por intermédio do teorema de Bayes, foi obtida a distribuição *a posteriori* conjunta dos parâmetros do modelo, dada por:

$$p(a, b, c, \tau | y, \alpha, \beta) \propto \tau^{n/2} \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-ct_i))\}^2 \right] \tau^{\alpha-1} \exp(-\beta\tau).$$

A partir da distribuição *a posteriori* conjunta dos parâmetros foram obtidas as distribuições condicionais completas para cada parâmetro, nas quais foi aplicada a amostragem Gibbs e o algoritmo Metropolis-Hastings, com o intuito de obter aproximações das distribuições marginais dos parâmetros.

4.1.1 Obtenção das distribuições condicionais completas para os parâmetros do modelo de Orskov e McDonald (1979)

A distribuição condicional completa para o parâmetro a do modelo de Orskov & McDonald (1979), que contempla a fração solúvel do componente nutritivo analisado, segundo a distribuição *a posteriori* conjunta dos parâmetros, é dada por:

$$p(a|b, c, \tau, y, \alpha, \beta) = \frac{p(a, b, c, \tau | y, \alpha, \beta)}{\int p(a, b, c, \tau | y, \alpha, \beta) da}. \quad (6)$$

A expressão (6) pode ser reescrita como:

$$p(a|b, c, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-ct_i))\}^2 \right]. \quad (7)$$

Na expressão (7), fazendo $y_i^* = y_i - b(1 - \exp(-ct_i))$, obtém-se:

$$p(a|b, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i^* - a\}^2\right]. \quad (8)$$

Desenvolvendo-se o quadrado em (8), tem-se:

$$p(a|b, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i^{*2} - 2ay_i^* + a^2\}\right]. \quad (9)$$

Resolvendo-se o somatório em (9) e considerando que os termos que não dependem do parâmetro a são termos constantes para este parâmetro, a distribuição condicional completa para o parâmetro a do modelo pode ser reescrita como:

$$p(a|b, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau}{2} \left(na^2 - 2a \sum_{i=1}^n y_i^* \right)\right]. \quad (10)$$

Rearranjando-se os termos da expressão (10) de maneira adequada, ou seja, colocando o termo de acompanha a^2 em evidência e utilizando o completamento de quadrado, a expressão pode ser reescrita como:

$$p(a|b, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau n}{2} \left(a - \frac{\sum_{i=1}^n y_i^*}{n} \right)^2\right]. \quad (11)$$

Lembrando que, na expressão (11), $y_i^* = y_i - b(1 - \exp(-ct_i))$, obtém-se:

$$p(a|b, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau n}{2} \left(a - \frac{\sum_{i=1}^n (y_i - b(1 - \exp(-ct_i)))}{n} \right)^2\right]. \quad (12)$$

A expressão (12) representa o núcleo de uma distribuição normal com média e variância dadas, respectivamente, por:

$$\mu^* = \frac{\sum_{i=1}^n (y_i - b(1 - \exp(-ct_i)))}{n}, \quad \sigma^{2*} = \frac{\sigma^2}{n}.$$

Analogamente, a distribuição condicional completa para o parâmetro b , fração potencialmente degradável, do modelo de Orskov & McDonald (1979) segundo a distribuição *a posteriori* conjunta, pode ser escrita de forma proporcional, como se segue:

$$p(b|a, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-ct_i))\}^2\right]. \quad (13)$$

Na expressão (13), fazendo $y_i^* = y_i - a$, obtém-se:

$$p(b|a, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i^* - b(1 - \exp(-ct_i))\}^2\right]. \quad (14)$$

Desenvolvendo-se o quadrado em (14), tem-se:

$$\propto \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i^{*2} - 2y_i^*b(1 - \exp(-ct_i)) + b^2(1 - \exp(-ct_i))^2\}\right]. \quad (15)$$

Na expressão (15), resolvendo o somatório e considerando que os termos que não dependem do parâmetro b são constantes para este parâmetro, obtém-se a expressão (16):

$$\propto \exp\left[-\frac{\tau}{2} \left\{ -2b \sum_{i=1}^n y_i^* (1 - \exp(-ct_i)) + b^2 \sum_{i=1}^n (1 - \exp(-ct_i))^2 \right\}\right]. \quad (16)$$

Rearranjando-se os termos da expressão (16) de maneira adequada, ou seja, colocando o termo que acompanha b^2 em evidência e utilizando o completamento de quadrado, a expressão pode ser reescrita como:

$$\propto \exp \left[-\frac{\tau \left[\sum_{i=1}^n (1 - \exp(-ct_i))^2 \right]}{2} \right] \left\{ b - \frac{\sum_{i=1}^n y_i^* (1 - \exp(-ct_i))}{\sum_{i=1}^n (1 - \exp(-ct_i))^2} \right\}^2. \quad (17)$$

Lembrando que, na expressão (17), $y_i^* = y_i - a$, obtém-se:

$$\propto \exp \left[-\frac{\tau \left[\sum_{i=1}^n (1 - \exp(-ct_i))^2 \right]}{2} \right] \left\{ b - \frac{\sum_{i=1}^n (y_i - a)(1 - \exp(-ct_i))}{\sum_{i=1}^n (1 - \exp(-ct_i))^2} \right\}^2. \quad (18)$$

A expressão (18) representa o núcleo de uma distribuição normal com média e variância dadas, respectivamente, por:

$$\mu^* = \frac{\sum_{i=1}^n (y_i - a)(1 - \exp(-ct_i))}{\sum_{i=1}^n (1 - \exp(-ct_i))^2}; \quad \sigma^{2*} = \frac{\sigma^2}{\sum_{i=1}^n (1 - \exp(-ct_i))^2}.$$

A distribuição condicional completa para o parâmetro c , taxa de degradação da fração potencialmente degradável, do modelo de Orskov & McDonald (1979), segundo a distribuição *a posteriori* conjunta, pode ser escrita de forma proporcional, como se segue:

$$p(c|a, b, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-ct_i))\}^2 \right]. \quad (19)$$

A expressão acima não representa uma distribuição de forma conhecida e, para este caso, foi necessária a utilização do algoritmo Metropolis-Hastings. Esse algoritmo utiliza a idéia de que um valor é gerado de uma distribuição auxiliar e aceito com uma dada probabilidade que garante a convergência da cadeia para a distribuição de equilíbrio que, neste caso, é a distribuição *a posteriori* do parâmetro c . A distribuição auxiliar utilizada no algoritmo foi uma

distribuição gama com parâmetros alfa e beta iguais a 1.000 e 25.000, respectivamente. Considerando que a média da distribuição gama é dada pelo quociente entre os parâmetros alfa e beta, os valores dos parâmetros da distribuição auxiliar foram escolhidos de forma a garantir que a média dessa distribuição estivesse próxima do valor paramétrico utilizado no estudo de simulação de 0,04.

Para o parâmetro precisão, τ , segundo a distribuição *a posteriori* conjunta pode ser escrita de forma proporcional, como se segue:

$$p(\tau|a,b,c,y,\alpha,\beta) \propto \tau^{n/2} \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-ct_i))\}^2\right] \tau^{\alpha-1} \exp(-\beta\tau). \quad (20)$$

Rearranjando-se os termos da expressão (20) de maneira adequada, essa pode ser reescrita como:

$$\propto \tau^{\frac{n+2\alpha}{2}-1} \exp\left[-\tau \left\{ \beta + \frac{1}{2} \sum_{i=1}^n (y_i - a - b(1 - \exp(-ct_i)))^2 \right\}\right]. \quad (21)$$

A expressão (21) representa o núcleo de uma distribuição gama com parâmetros, dado por:

$$\alpha^* = \frac{n+2\alpha}{2} \text{ e } \beta^* = \beta + \frac{1}{2} \sum_{i=1}^n (y_i - a - b(1 - \exp(-ct_i)))^2,$$

em que: α e β são os hiperparâmetros da distribuição gama utilizada como *priori* para o parâmetro precisão e assumindo valores iguais a 250 e 500, respectivamente, tanto para o ajuste dos dados simulados como para o ajuste dos dados reais.

4.2 Modelo de McDonald (1981)

De posse da verossimilhança obtida para o modelo de McDonald (1981), das distribuições *a priori* consideradas e por intermédio do teorema de Bayes foi obtida a distribuição *a posteriori* conjunta dos parâmetros do modelo, dada por:

$$p(a, b, c, L, \tau | y, \alpha, \beta) \propto \tau^{n/2} \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-c(t_i - L)))\}^2 \right] I_{(L>0)} \tau^{\alpha-1} \exp(-\beta\tau).$$

A partir da distribuição *a posteriori* conjunta foram obtidas as distribuições condicionais completas para cada parâmetro, nas quais foi aplicada a amostragem Gibbs e algoritmo Metropolis-Hastings, no intuito de obter aproximações das distribuições marginais dos parâmetros.

4.2.1 Obtenção das distribuições condicionais completas para os parâmetros do modelo de McDonald (1981)

A distribuição condicional completa para o parâmetro a , fração solúvel do componente nutritivo analisado, do modelo de McDonald (1981), segundo a distribuição *a posteriori* conjunta dos parâmetros, pode ser escrita como:

$$p(a|b, c, L, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-c(t_i - L)))\}^2 \right]. \quad (22)$$

Na expressão (22), fazendo $y_i^* = y_i - b(1 - \exp(-c(t_i - L)))$, obtém-se:

$$p(a|b, c, L, \tau, y_i, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i^* - a\}^2 \right]. \quad (23)$$

Desenvolvendo-se o quadrado em (23), tem-se:

$$p(a|b, c, L, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i^{*2} - 2ay_i^* + a^2\} \right]. \quad (24)$$

Resolvendo o somatório em (24) e considerando que os termos que não dependem do parâmetro a são termos constantes para este parâmetro, tem-se:

$$p(a|b, c, L, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \left(na^2 - 2a \sum_{i=1}^n y_i^* \right) \right]. \quad (25)$$

Rearranjando-se os termos da expressão (25) de maneira adequada, ou seja, colocando o termo de acompanha a^2 em evidência e utilizando o completamento de quadrado, a expressão pode ser reescrita como:

$$p(a|b, c, L, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau n}{2} \left(a - \frac{\sum_{i=1}^n y_i^*}{n} \right)^2 \right]. \quad (26)$$

Lembrando que, na expressão (26) $y_i^* = y_i - b(1 - \exp(-c(t_i - L)))$, obtém-se:

$$p(a|b, c, L, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau n}{2} \left(a - \frac{\sum_{i=1}^n (y_i - b(1 - \exp(-c(t_i - L))))}{n} \right)^2 \right]. \quad (27)$$

A expressão (27) representa o núcleo de uma distribuição normal com média e variância dadas, respectivamente, por:

$$\mu^* = \frac{\sum_{i=1}^n (y_i - b(1 - \exp(-c(t_i - L))))}{n}, \quad \sigma^{2*} = \frac{\sigma^2}{n}.$$

A distribuição condicional completa para o parâmetro b , fração potencialmente degradável, do modelo de McDonald (1981), segundo a distribuição *a posteriori* conjunta, pode ser escrita de forma proporcional, como se segue:

$$p(b|a, c, L, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-c(t_i - L)))\}^2 \right]. \quad (28)$$

Na expressão (28), fazendo-se $y_i^* = y_i - a$, obtém-se:

$$p(b|a, c, L, \tau, y, \alpha, \beta) \propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \left\{ y_i^* - b(1 - \exp(-c(t_i - L))) \right\}^2 \right]. \quad (29)$$

Desenvolvendo-se o quadrado em (29), tem-se:

$$\propto \exp \left[-\frac{\tau}{2} \sum_{i=1}^n \left\{ y_i^{*2} - 2y_i^*b(1 - \exp(-c(t_i - L))) + b^2(1 - \exp(-c(t_i - L)))^2 \right\} \right]. \quad (30)$$

Na expressão (30), resolvendo o somatório e considerando que os termos que não dependem do parâmetro b são constantes para este parâmetro, então:

$$\propto \exp \left[-\frac{\tau}{2} \left\{ -2b \sum_{i=1}^n y_i^* (1 - \exp(-c(t_i - L))) + b^2 \sum_{i=1}^n (1 - \exp(-c(t_i - L)))^2 \right\} \right]. \quad (31)$$

Rearranjando-se os termos da expressão (31) de maneira adequada, ou seja, colocando o termo que acompanha b^2 em evidência e utilizando o completamento de quadrado, a expressão pode ser reescrita como:

$$\propto \exp \left[-\frac{\tau \left[\sum_{i=1}^n (1 - \exp(-c(t_i - L)))^2 \right]}{2} \left\{ b - \frac{\sum_{i=1}^n y_i^* (1 - \exp(-c(t_i - L)))}{\sum_{i=1}^n (1 - \exp(-c(t_i - L)))^2} \right\}^2 \right]. \quad (32)$$

Lembrando que, na expressão (32) $y_i^* = y_i - a$, obtém-se:

$$\propto \exp \left[-\frac{\tau \left[\sum_{i=1}^n (1 - \exp(-c(t_i - L)))^2 \right]}{2} \left\{ b - \frac{\sum_{i=1}^n (y_i - a)(1 - \exp(-c(t_i - L)))}{\sum_{i=1}^n (1 - \exp(-c(t_i - L)))^2} \right\}^2 \right]. \quad (33)$$

A expressão (33) representa o núcleo de uma distribuição normal com média e variância dadas, respectivamente, por:

$$\mu^* = \frac{\sum_{i=1}^n (y_i - a)(1 - \exp(-c(t_i - L)))}{\sum_{i=1}^n (1 - \exp(-c(t_i - L)))^2}; \quad \sigma^{*2} = \frac{\sigma^2}{\sum_{i=1}^n (1 - \exp(-c(t_i - L)))^2}.$$

Para o parâmetro c , taxa de degradação da fração potencialmente degradável, do modelo de McDonald (1981), segundo a distribuição *a posteriori* conjunta, pode ser escrita de forma proporcional, como se segue:

$$p(c|a, b, L, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-c(t_i - L)))\}^2\right]. \quad (34)$$

A expressão (34) não representa uma distribuição de forma conhecida e, para este caso, foi necessária a utilização do algoritmo Metropolis-Hastings. A distribuição auxiliar utilizada no algoritmo foi uma distribuição gama com parâmetros alfa e beta iguais a 1.000 e 25.000, respectivamente, no ajuste dos dados simulados e iguais a 2.000 e 50.000 no ajuste dos dados reais. Considerando que a média da distribuição gama é dada pelo quociente entre os parâmetros alfa e beta, os valores dos parâmetros da distribuição auxiliar foram escolhidos de forma que garantissem que a média dessa distribuição estivesse próxima do valor paramétrico utilizado no estudo de simulação de 0,04.

A distribuição condicional completa para o parâmetro L , que contempla o tempo de colonização das partículas, pode ser escrita de forma proporcional, como se segue:

$$p(L|a, b, c, \tau, y, \alpha, \beta) \propto \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-c(t_i - L)))\}^2\right] I_{(L>0)} \quad (35)$$

A expressão (35) não representa uma distribuição de forma conhecida e, para este caso, foi necessária a utilização do algoritmo Metropolis-Hastings. A distribuição auxiliar utilizada no algoritmo foi uma distribuição uniforme entre os limites 0 e 8. Os valores dos parâmetros da distribuição auxiliar foram escolhidos de forma que o limite inferior garantisse que o valor do parâmetro fosse positivo e que o limite superior mantivesse uma interpretação coerente do parâmetro, uma vez que este é medido em horas e significa o tempo para que a degradação ruminal se inicie.

A distribuição condicional completa para o parâmetro precisão, τ , segundo a distribuição *a posteriori* conjunta, pode ser escrita de forma proporcional, como se segue:

$$p(\tau|a,b,c,L,y,\alpha,\beta) \propto \tau^{n/2} \exp\left[-\frac{\tau}{2} \sum_{i=1}^n \{y_i - a - b(1 - \exp(-c(t_i - L)))\}^2\right] \tau^{\alpha-1} \exp(-\beta\tau). \quad (36)$$

Rearranjando os termos da expressão (36) de maneira adequada, essa pode ser reescrita como:

$$\propto \tau^{\frac{n+2\alpha}{2}-1} \exp\left[-\tau \left\{ \beta + \frac{1}{2} \sum_{i=1}^n (y_i - a - b(1 - \exp(-c(t_i - L))))^2 \right\}\right]. \quad (37)$$

A expressão (37) representa o núcleo de uma distribuição gama com parâmetros, dado por:

$$\alpha^* = \frac{n+2\alpha}{2} \text{ e } \beta^* = \beta + \frac{1}{2} \sum_{i=1}^n (y_i - a - b(1 - \exp(-c(t_i - L))))^2,$$

em que α e β são os hiperparâmetros da distribuição gama utilizada como *priori* para o parâmetro precisão e assumindo valores iguais a 250 e 500, respectivamente, para ajuste dos dados simulados e valores iguais a 500 e 1.000, para o ajuste dos dados reais.

4.3 Análise dos dados simulados

4.3.1 Modelo de Orskov e McDonald (1979) – Modelo 1

Na Figura 2 podem-se observar os valores simulados da porcentagem de degradação da FDN em função dos tempos de incubação, em horas, segundo o modelo de Orskov & McDonald (1979). Foram simulados 120 valores dessa variável, no intuito de representar uma coleta por hora. Os valores assumidos

como verdadeiros para os parâmetros do modelo estão descritos no item 3.1 e apresentados na Tabela 3.

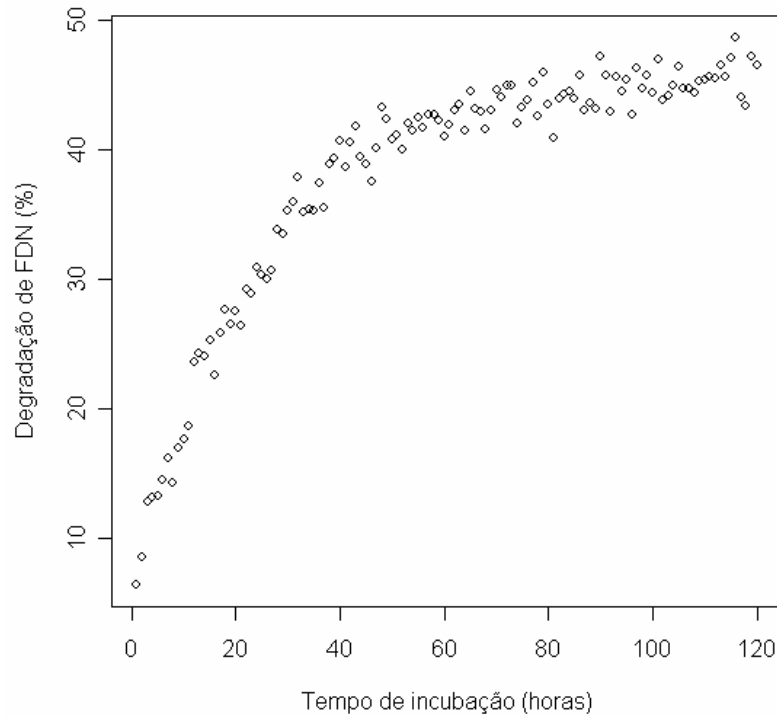


FIGURA 2. Valores simulados da porcentagem de degradação de fibra em detergente neutro (FDN), em função dos tempos de incubação, em horas.

A convergência das cadeias de cada um dos parâmetros do modelo de Orskov & McDonald (1979), bem como das funções dos parâmetros, dadas pela degradabilidade efetiva (DE) e potencial (DP), foi monitorada por meio da visualização gráfica do traço da cadeia gerada pelo método MCMC e pelos critérios propostos Geweke (1992), Gelman & Rubin (1992), Heidelberger & Welch (1993) e Raftery & Lewis (1992), além do Erro de Monte Carlo (EMC).

Visualizam-se, na Tabela 2, os resultados para os critérios de convergência das cadeias. Pelo critério de Geweke, que consiste em indicar a convergência da média *a posteriori*, o valor p obtido foi sempre maior que o nível de significância pré-fixado de 5% (0,05), não apresentando evidências contra a convergência para os parâmetros do modelo e para as degradabilidades, efetiva e potencial, que são funções dos parâmetros.

Segundo Gelman & Rubin (1992), quando a estimativa de R for grande, sugere-se que qualquer das estimativas das variâncias entre as cadeias geradas pode ser reduzida com a utilização de mais iterações. Entende-se que, se a estimativa de R aproxima-se de um, cada m conjunto de n iterações se aproxima da distribuição de interesse. Para a análise desse critério, foram consideradas duas cadeias com valores iniciais distintos em cada parâmetro. Os valores estimados de R podem ser observados na Tabela 2 e também não apresentam qualquer evidência contra a convergência das cadeias.

A regra de decisão do critério de Raftery & Lewis (1992) é baseada no fator de dependência (FD), que é responsável pelo acréscimo multiplicativo ao número de iterações necessárias para se alcançar a convergência. De acordo com os autores, se o fator de dependência for maior que cinco, pode-se concluir que a cadeia não atingiu convergência. Para a análise dos dados simulados, verifica-se que, por este critério, a cadeia amostrada de todos os parâmetros também atingiu a convergência (Tabela 2).

Outro indicativo da qualidade de convergência das cadeias, apresentado na Tabela 3, são os valores encontrados para o Erro de Monte Carlo (EMC), que foram de baixa magnitude em relação à estimativa a posteriori dos parâmetros do modelo, o que, segundo Blasco et al. (2003), também assegura a convergência.

Segundo o critério de Heidelberger & Welch (1993), que utiliza testes estatísticos para avaliar a hipótese nula de estacionariedade da amostra gerada,

não é necessário um número maior de iterações, além das computadas na análise, ou seja, a hipótese nula foi aceita para todos os parâmetros e o tamanho do *burn-in* sugerido foi igual a zero.

TABELA 2. Critérios de Geweke (valor p), Gelman e Rubin (\hat{R}) e Raftery Lewis (fator de dependência - FD) utilizados na detecção da convergência das cadeias amostradas.

Parâmetros	Geweke (valor p)	Gelman e Rubin (\hat{R})	Raftery e Lewis (FD)
a	0,7393	0,9995	1,045
b	0,8934	0,9995	1,026
c	0,4086	0,9997	0,970
τ	0,1795	0,9998	1,007
¹ DE	0,8924	0,9997	1,085
² DP	0,5121	0,9996	1,065

¹ Degradabilidade efetiva com taxa de passagem de 5%/h; ² Degradabilidade potencial.

Uma das vantagens da inferência bayesiana via amostras das distribuições *a posteriori* para o cálculo de estimativas é que quaisquer combinações lineares, ou não lineares, dos valores amostrais, uma vez calculada em cada passo de amostragem, geram diretamente um ponto amostral para a estatística de interesse, com base na qual a inferência é bastante simples (Rosa, 1998). Assim, foram geradas cadeias para as degradabilidades efetiva (DE) e potencial (DP), que são dadas por uma combinação dos parâmetros do modelo, as quais também tiveram suas convergências constatadas pelos critérios adotados (Tabela 2).

Observam-se, pelos dados da Tabela 3, os valores paramétricos utilizados no estudo de simulação, bem como as estimativas da média e moda *a posteriori*, o intervalo de máxima densidade *a posteriori* e o Erro de Monte Carlo (EMC). Os resultados mostram estimativas bastante próximas da média e da moda *a posteriori*, evidenciando distribuições do tipo simétrica, para todos os

parâmetros (Figuras 3 a 8). Nesses casos, um resumo pontual da informação contida nas distribuições *a posteriori* pode ser dado tanto pela média *a posteriori* como pela moda *a posteriori*. Ainda pela Tabela 3 é possível observar que a metodologia bayesiana mostrou-se efetiva para estimar os parâmetros do modelo de Orskov & McDonald (1979), bem como a precisão do erro (parâmetro τ), pois os valores paramétricos utilizados no estudo de simulação encontram-se todos dentro do intervalo de máxima densidade *a posteriori* estimado. Com este resultado foi possível validar a metodologia e a rotina computacional utilizada.

TABELA 3. Valor paramétrico, média e moda *a posteriori*, intervalo de máxima densidade *a posteriori* (HPD) e Erro de Monte Carlo (EMC) para os parâmetros do modelo de Orskov & McDonald (1989).

Parâmetros	Valor paramétrico	Média	Moda	Intervalo HPD 95%		EMC
				¹ LI	LS	
<i>a</i>	6,00	5,75	5,75	4,61	6,79	$8,33 \times 10^{-3}$
<i>b</i>	40,00	40,14	40,10	39,03	41,25	$8,24 \times 10^{-3}$
<i>c</i>	0,040	0,040	0,041	0,039	0,042	$1,50 \times 10^{-5}$
τ	0,50	0,51	0,50	0,45	0,56	$4,64 \times 10^{-4}$
² DE	23,78	23,77	23,74	23,34	24,21	$3,57 \times 10^{-3}$
³ DP	46,00	45,90	45,89	45,49	46,33	$3,51 \times 10^{-3}$

¹Limites inferior e superior do intervalo HDP; ²Degradabilidade efetiva para taxa de passagem de 5%/h; ³Degradabilidade potencial.

Os resultados apresentados na Tabela 2 podem ser confirmados por meio da visualização gráfica do traço da cadeia, nas Figuras de 3 a 8, após o descarte das 20.000 iterações iniciais (*burn in*) e o intervalo entre pontos amostrados (*thin*) de 40 observações.

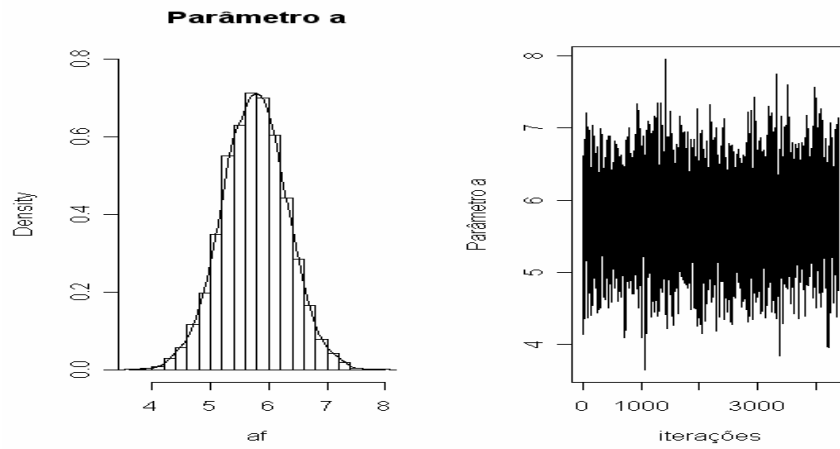


FIGURA 3. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro *a* do modelo de Orskov & McDonald (1979).

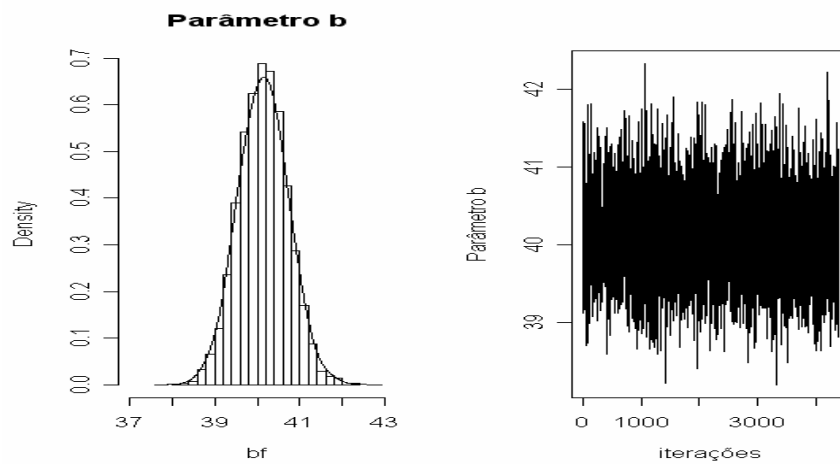


FIGURA 4. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro *b* do modelo de Orskov & McDonald (1979).

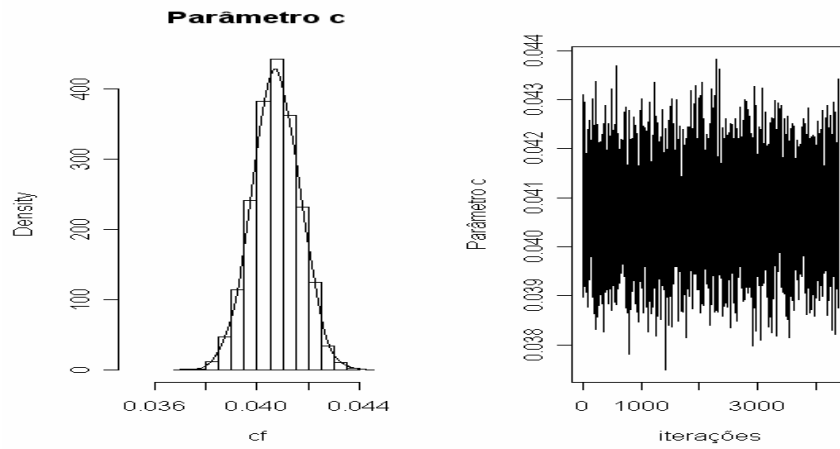


FIGURA 5. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro c do modelo de Orskov & McDonald (1979).

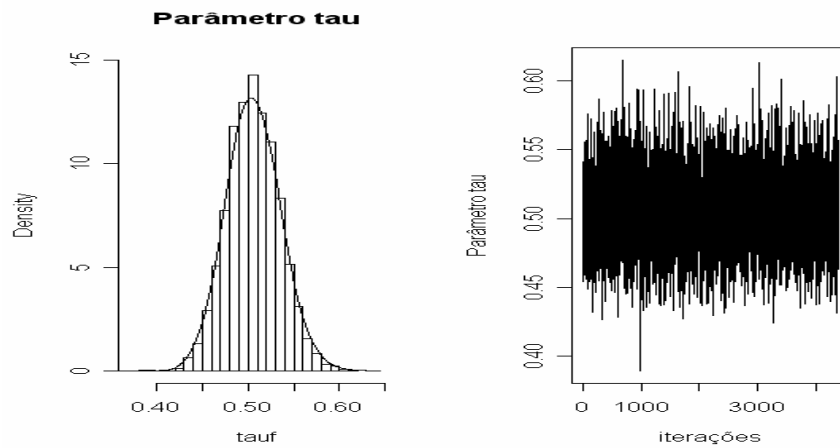


FIGURA 6. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a precisão.

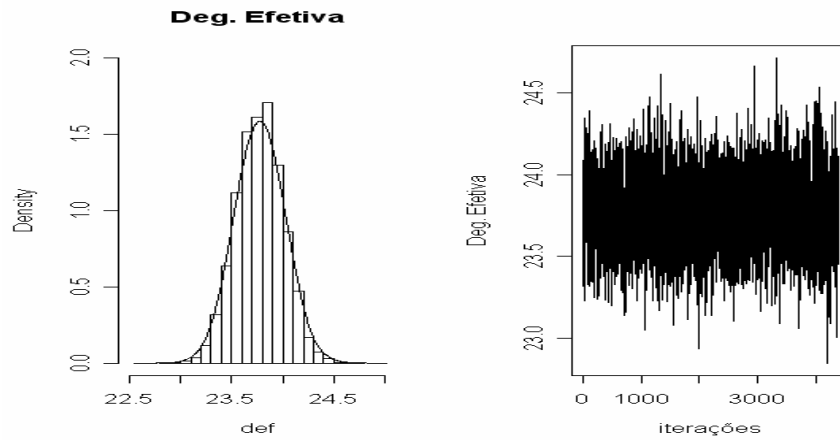


FIGURA 7. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade efetiva, em função dos parâmetros do modelo de Orskov & McDonald (1979).

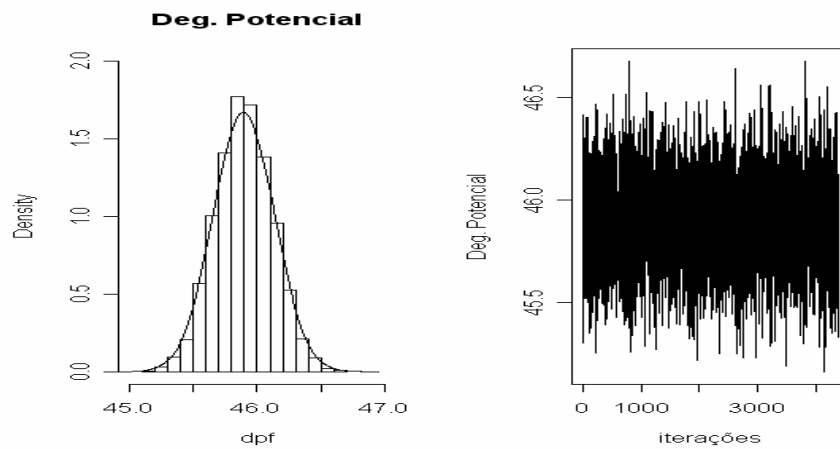


FIGURA 8. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade potencial, em função dos parâmetros do modelo de Orskov & McDonald (1979).

Pelo gráfico da Figura 3, observa-se que a cadeia gerada para o parâmetro a do modelo oscilou em torno da média estimada, não apresentando

picos discrepantes ao longo da cadeia. O mesmo comportamento pode ser observado para a cadeia dos demais parâmetros (Figuras 4 a 6), bem como para as cadeias geradas para as degradabilidades efetiva e potencial (Figuras 7 e 8).

Com base na simetria das densidades e dos histogramas, apresentados nas Figuras de 3 a 8, podem-se propor novas distribuições *a priori* para estudos futuros. Para os histogramas e densidades simétricas pode-se sugerir a distribuição normal ou a distribuição t de Student.

4.3.2 Modelo de McDonald (1981) – Modelo 2

Na Figura 9 observam-se os valores simulados da porcentagem de degradação da FDN, em função dos tempos de incubação, em horas, segundo o modelo de McDonald (1981). Foram simulados também 120 valores dessa variável, no intuito de representar uma coleta por hora. Os valores assumidos como verdadeiros para os parâmetros do modelo estão descritos no item 3.1. e são apresentados na Tabela 5.

A convergência das cadeias de cada um dos parâmetros do modelo de McDonald (1981), bem como das funções dos parâmetros dadas pela degradabilidade efetiva (DE) e potencial (DP), foi monitorada pelos mesmos critérios adotados para o modelo anterior, ou seja, visualização gráfica do traço da cadeia gerada pelo método MCMC, pelos critérios propostos por Geweke (1992), Gelman & Rubin (1992), Heidelberger & Welch (1993) e Raftery & Lewis (1992), além do Erro de Monte Carlo (EMC).

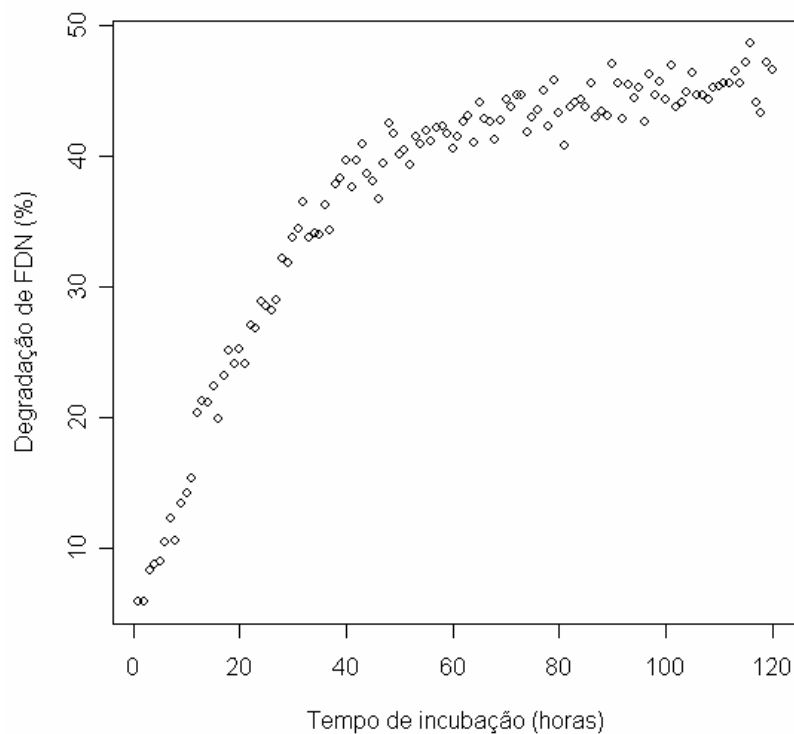


FIGURA 9. Valores simulados da porcentagem de degradação de fibra em detergente neutro (FDN), em função dos tempos de incubação, em horas.

Pelo critério de Geweke, o valor p obtido para esse modelo também foi sempre maior que o nível de significância pré-fixado de 5%, não apresentando evidências contra a convergência. Para a análise do critério de Gelman e Rubin, foram consideradas duas cadeias com valores iniciais distintos em cada parâmetro e os valores estimados de R , que podem ser observados na Tabela 4, também não apresentam qualquer evidência contra a convergência das cadeias.

A regra de decisão do critério de Raftery & Lewis (1992), baseado no fator de dependência (FD), mostra que, para a análise dos dados simulados a

partir no modelo de McDonald (1981), a cadeia amostrada de todos os parâmetros também atingiu a convergência (Tabela 4).

Os valores encontrados para o Erro de Monte Carlo (EMC) foram de baixa magnitude (Tabela 5) em relação à estimativa *a posteriori* dos parâmetros, assegurando a convergência das cadeias. Segundo o critério de Heidelberger & Welch (1993), a aceitação da hipótese nula de estacionariedade da amostra gerada indica que não é necessário um número maior de iterações além das computadas na análise.

As cadeias geradas para as degradabilidades efetiva (DE) e potencial (DP), dadas pela combinação dos parâmetros do modelo, tiveram sua convergência constatada (Tabela 4).

TABELA 4. Critérios de Geweke (valor p), Gelman e Rubin (\hat{R}) e Raftery e Lewis (fator de dependência - FD), utilizados na detecção da convergência das cadeias amostradas.

Parâmetros	Geweke (valor p)	Gelman e Rubin (\hat{R})	Raftery e Lewis (FD)	EMC
a	0,1161	0,9997	1,077	$2,53 \times 10^{-2}$
b	0,1195	0,9999	2,759	$2,58 \times 10^{-2}$
c	0,5992	1,0023	0,981	$1,54 \times 10^{-5}$
L	0,3133	0,9997	1,923	$2,22 \times 10^{-2}$
τ	0,0931	1,0001	0,968	$4,53 \times 10^{-4}$
¹ DE	0,1380	0,9999	1,106	$1,52 \times 10^{-2}$
² DP	0,7851	1,0013	1,065	$3,14 \times 10^{-3}$

¹ Degradabilidade efetiva com taxa de passagem de 5%/h; ² Degradabilidade potencial.

Observam-se, pelos dados da Tabela 5, os valores paramétricos utilizados no estudo de simulação, bem como as estimativas da média e moda *a posteriori* e o intervalo de máxima densidade *a posteriori*. Os resultados mostram as estimativas da média e da moda *a posteriori* bastante próximas apenas para o parâmetro c do modelo de McDonald (1981) e para o parâmetro precisão (τ), evidenciando distribuições do tipo simétricas para estes

parâmetros (Figuras 12 e 14). Novamente, nesses casos, um resumo pontual da informação contida nas distribuições *a posteriori* pode ser dado tanto pela média *a posteriori* como pela moda *a posteriori*. Para os parâmetros *a* e *L*, as distribuições mostram-se assimétricas à direita (Figuras 10 e 13) e, neste caso, a estimativa que melhor resume a distribuição *a posteriori* é a moda *a posteriori* e não a média. O mesmo se pode inferir para o parâmetro *b* do modelo em estudo, esse, porém, apresenta uma distribuição assimétrica à esquerda (Figura 11).

Ainda na Tabela 5 é possível observar que a metodologia bayesiana mostrou-se efetiva para estimar os parâmetros do modelo de McDonald (1981), bem como a precisão do erro (parâmetro τ), pois os valores paramétricos utilizados no estudo de simulação encontram-se todos dentro do intervalo de máxima densidade *a posteriori* estimado com 95% de credibilidade. Com este resultado foi possível validar a metodologia e a rotina computacional utilizadas.

TABELA 5. Valor paramétrico, média e moda *a posteriori*, intervalo de máxima densidade *a posteriori* (HPD) e Erro de Monte Carlo (EMC), para os parâmetros do modelo de McDonald (1981).

Parâmetros	Valor paramétrico	Média	Moda	Intervalo HPD 95%		EMC
				¹ LI	LS	
<i>a</i>	6,00	3,20	2,67	0,89	6,10	$2,53 \times 10^{-2}$
<i>b</i>	40,00	42,65	43,14	39,88	44,96	$2,58 \times 10^{-2}$
<i>c</i>	0,040	0,041	0,041	0,039	0,043	$1,54 \times 10^{-5}$
<i>L</i>	3,00	1,83	1,36	0,11	3,92	$2,22 \times 10^{-2}$
τ	0,50	0,50	0,50	0,45	0,57	$4,53 \times 10^{-4}$
² DE	23,78	22,48	22,16	21,12	24,08	$1,52 \times 10^{-2}$
³ DP	46,00	45,86	45,85	45,40	46,32	$3,14 \times 10^{-3}$

¹Limites inferior e superior do intervalo HDP; ²Degradabilidade efetiva com taxa de passagem de 5%/h; ³Degradabilidade potencial.

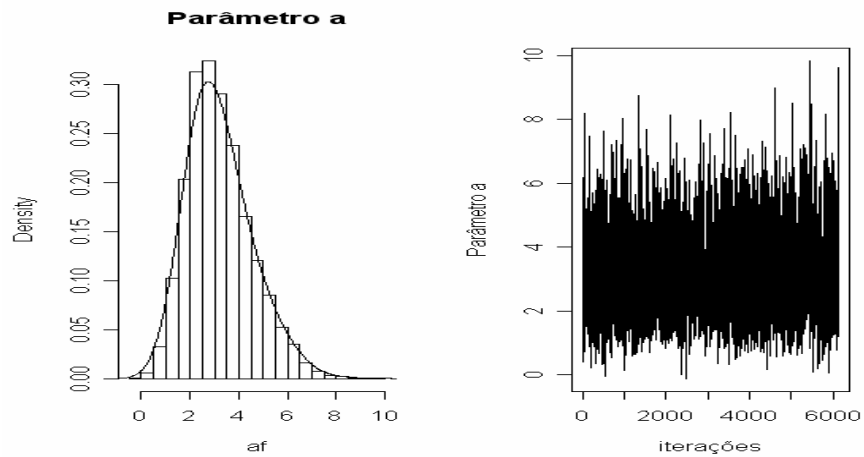


FIGURA 10. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro a do modelo de McDonald (1981).

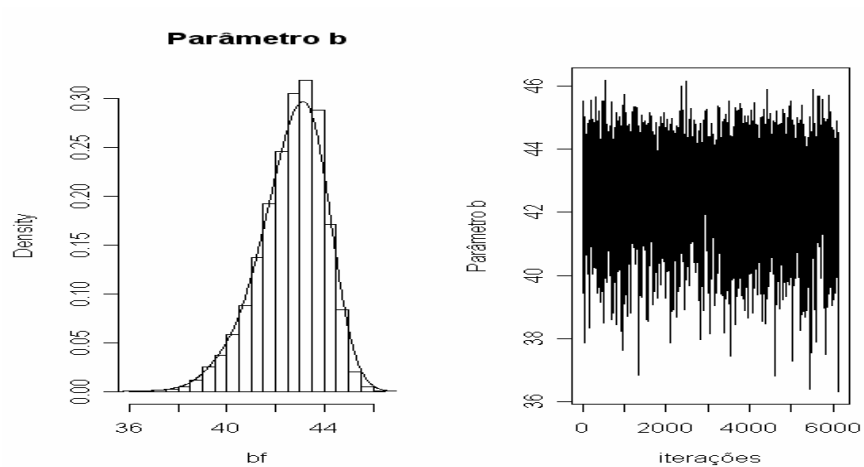


FIGURA 11. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro b do modelo de McDonald (1981).

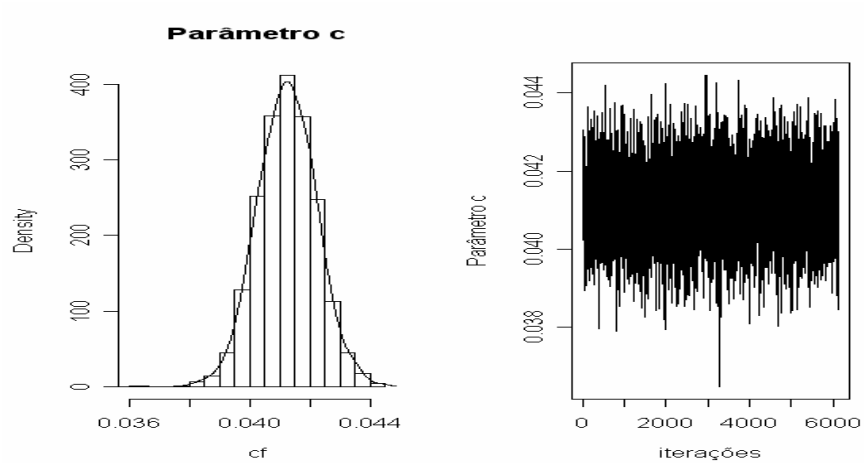


FIGURA 12. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro c do modelo de McDonald (1981).

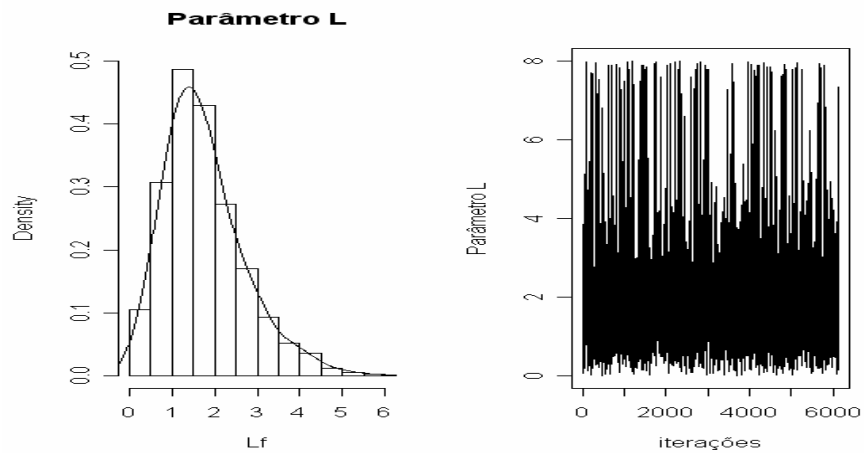


FIGURA 13. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro L do modelo de McDonald (1981).

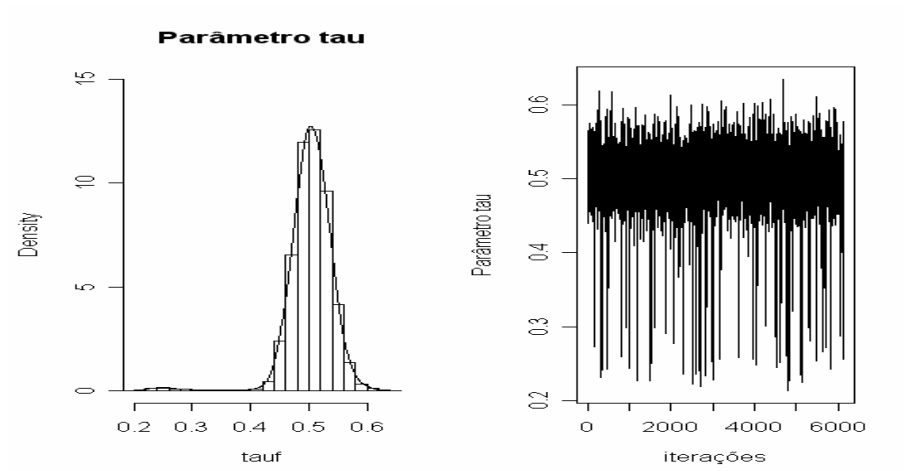


FIGURA 14. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a precisão.

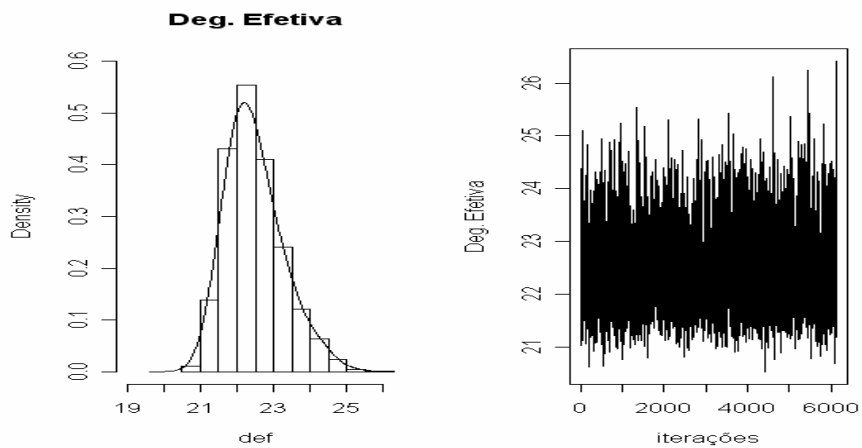


FIGURA 15. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade efetiva, em função dos parâmetros do modelo de McDonald (1981).

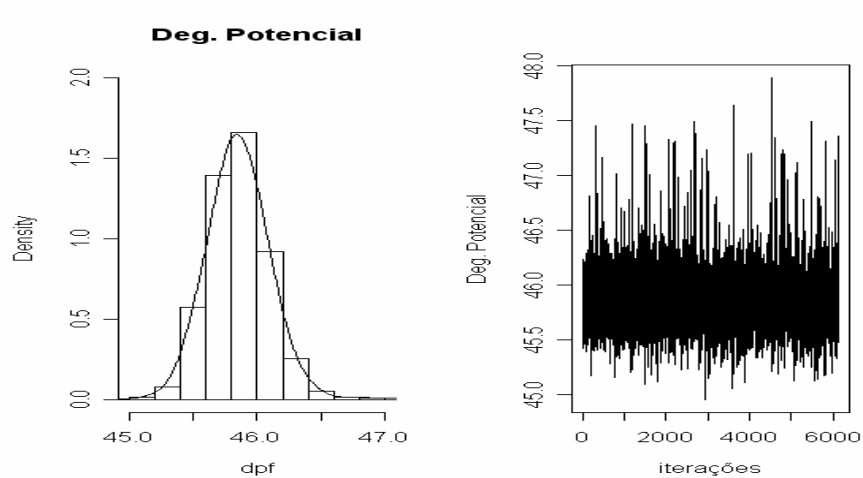


FIGURA 16. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade potencial, em função dos parâmetros do modelo de McDonald (1981).

Com base na simetria das densidades e dos histogramas, apresentados nas Figuras 12 e 14, podem-se propor novas distribuições *a priori* para estudos futuros como a distribuição normal ou a distribuição t de Student. Para os histogramas e as densidades assimétricas (Figuras 10, 11 e 13), poder-se-ia propor a distribuição gama.

4.3.3 Comparação dos modelos

O resultado da comparação dos modelos de Orskov e McDonald (modelo 1) e de McDonald (modelo 2) por meio do Fator de Bayes encontram-se na Tabela 6. Se o Fator de Bayes for maior que 1, o modelo 1, contemplado no numerador, é mais plausível ou verossímil do que o modelo 2, então, opta-se por escolher o modelo 1 para prever o fenômeno em questão. Segundo Jeffreys (1961), as magnitudes atingidas pelo Fator de Bayes demonstram os graus de

plausibilidade do modelo. Os resultados na Tabela 6 demonstram que o modelo de Orskov & McDonald (1979) é mais plausível que o modelo proposto por McDonald (1981) e que a evidência a favor deste modelo é considerada muito forte.

TABELA 6. Fator de Bayes para a comparação dos modelos de Orskov & McDonald (1979) e McDonald (1981).

Modelos	FB_{12}
Orskov e McDonald (1979)	538,49
McDonald (1981)	

4.4 Análise dos dados reais

4.4.1 Modelo de Orskov e McDonald (1979) – Modelo 1

A convergência das cadeias de cada um dos parâmetros do modelo de Orskov & McDonald (1979), bem como das funções dos parâmetros, dadas pela degradabilidade efetiva (DE) e potencial (DP), foi monitorada por meio da visualização gráfica do traço da cadeia gerada pelo método MCMC e pelos critérios propostos Geweke (1992), Gelman & Rubin (1992), Heidelberger & Welch (1993) e Raftery & Lewis (1992), além do Erro de Monte Carlo (EMC).

Visualizam-se, pelos dados da Tabela 7, os resultados para os critérios de convergência das cadeias. Pelo critério de Geweke (1992), o valor p obtido foi sempre maior que o nível de significância pré-fixado de 5% (0,05), não apresentando evidências contra a convergência para os parâmetros do modelo e para as degradabilidades, efetiva e potencial, que são funções dos parâmetros. Segundo o critério de Gelman & Rubin (1992), os valores estimados de R sugerem que não é necessário um maior número de iterações uma vez que estão próximos de 1. Para a análise desse critério foram consideradas duas cadeias

com valores iniciais distintos em cada parâmetro. A regra de decisão do critério de Raftery & Lewis (1992), baseado no fator de dependência (FD), demonstra que, para todos os parâmetros, o fator de dependência foi inferior a cinco, podendo-se concluir que a cadeia atingiu convergência. Outro indicativo da qualidade de convergência das cadeias, apresentado na Tabela 8, são os valores encontrados para o Erro de Monte Carlo (EMC), que foram de baixa magnitude em relação às estimativas dos parâmetros, o que também assegura a convergência.

TABELA 7. Critérios de Geweke (valor p), Gelman e Rubin (\hat{R}) e Raftery Lewis (fator de dependência - FD), utilizados na detecção da convergência das cadeias amostradas.

Parâmetros	Geweke (valor p)	Gelman e Rubin (\hat{R})	Raftery e Lewis (FD)
a	0,8657	0,9996	1,046
b	0,6414	0,9996	0,971
c	0,7492	0,9995	1,007
τ	0,1644	0,9998	0,989
¹ DE	0,8753	0,9995	1,007
² DP	0,5275	0,9996	0,971

¹ Degradabilidade efetiva com taxa de passagem de 5%/h; ² Degradabilidade potencial.

O critério de Heidelberger & Welch (1993), cujos resultados não foram apresentados, demonstrou a aceitação da hipótese nula de estacionariedade da amostra gerada, não sendo necessário um número maior de iterações além das computadas na análise.

Foram geradas cadeias para as degradabilidades efetiva (DE) e potencial (DP), que são dadas por uma combinação dos parâmetros do modelo, as quais também tiveram suas convergências constatadas pelos critérios adotados (Tabela 7). Os resultados apresentados na Tabela 7 podem ser confirmados por meio da

visualização gráfica do traço da cadeia, nas Figuras de 17 a 22, após o descarte das 20.000 iterações iniciais (*burn in*) e o intervalo entre pontos amostrados (*thin*) de 40 observações.

Observam-se, por meio dos dados da Tabela 8, as estimativas da média e moda *a posteriori*, o intervalo de máxima densidade *a posteriori* e Erro de Monte Carlo (EMC). Os resultados mostram estimativas da média e da moda *a posteriori* bastante próximas, evidenciando distribuições do tipo simétricas para todos os parâmetros (figuras 17 a 22). Nesses casos, um resumo pontual da informação contida nas distribuições *a posteriori* pode ser dado tanto pela média *a posteriori* como pela moda *a posteriori*.

TABELA 8. Valores estimados da média e moda *a posteriori*, e intervalo de máxima densidade *a posteriori* (HPD) para os parâmetros do modelo de Orskov e McDonald no ajuste aos dados reais.

Parâmetros	Média	Moda	Intervalo HPD 95%		EMC
			¹ LI	LS	
<i>a</i>	13,50	13,48	11,78	15,21	$1,34 \times 10^{-2}$
<i>b</i>	31,98	31,97	29,46	34,52	$1,86 \times 10^{-2}$
<i>c</i>	0,039	0,040	0,037	0,042	$1,66 \times 10^{-5}$
τ	0,49	0,49	0,43	0,55	$4,08 \times 10^{-4}$
² DE	27,66	27,66	26,71	28,67	$7,93 \times 10^{-3}$
³ DP	45,49	45,48	43,98	46,96	$1,07 \times 10^{-2}$

¹Limites inferior e superior do intervalo HDP; ²Degradabilidade efetiva com taxa de passagem de 5%/h; ³Degradabilidade potencial.

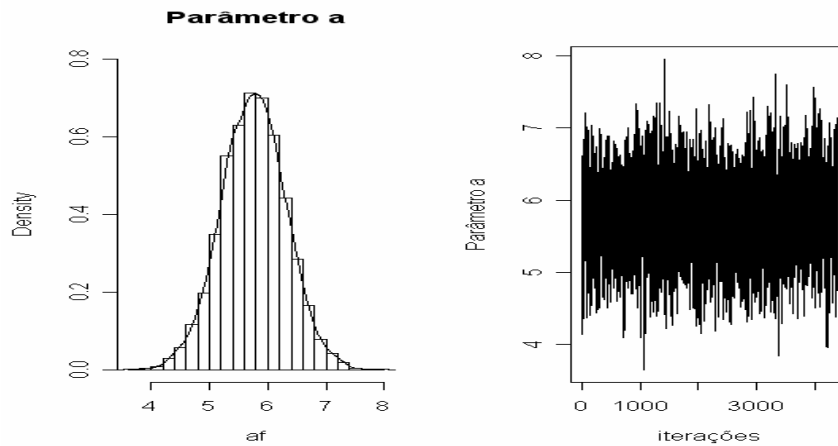


FIGURA 17. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro a do modelo de Orskov & McDonald (1979).

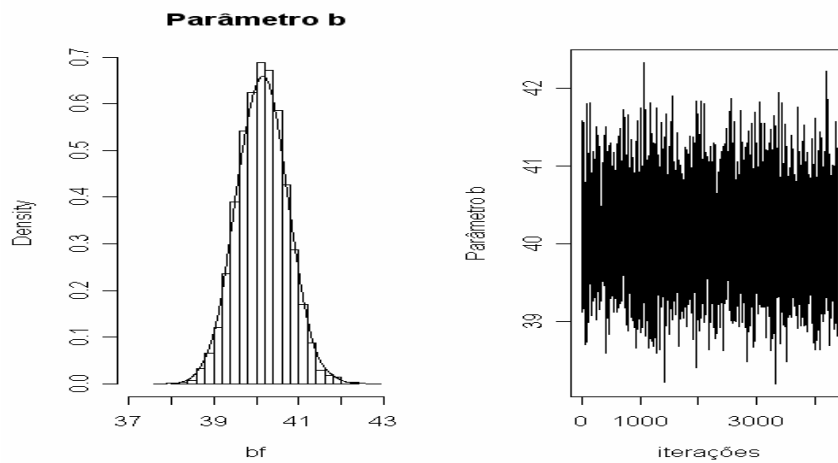


FIGURA 18. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro b do modelo de Orskov & McDonald (1979).

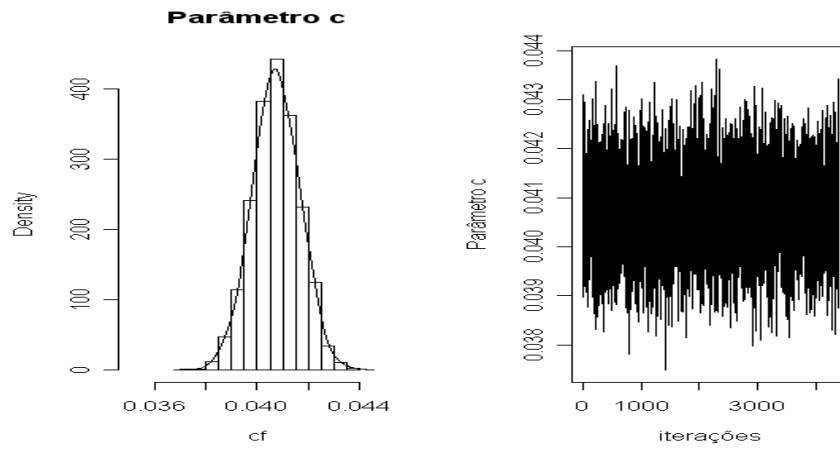


FIGURA 19. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro c do modelo de Orskov & McDonald (1979).

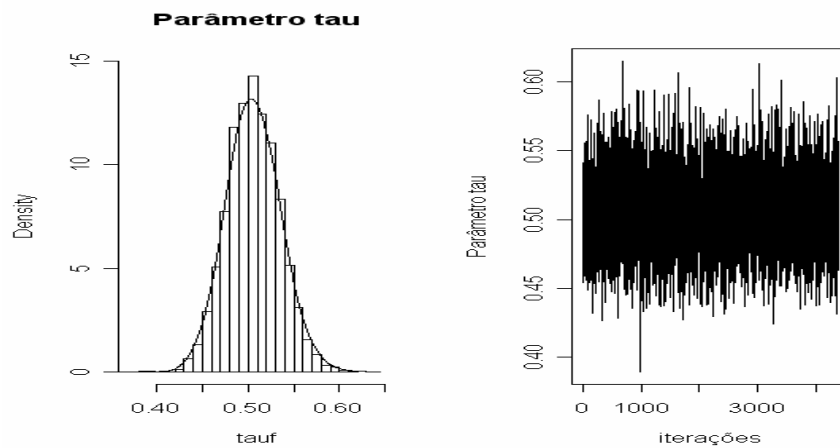


FIGURA 20. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a precisão.

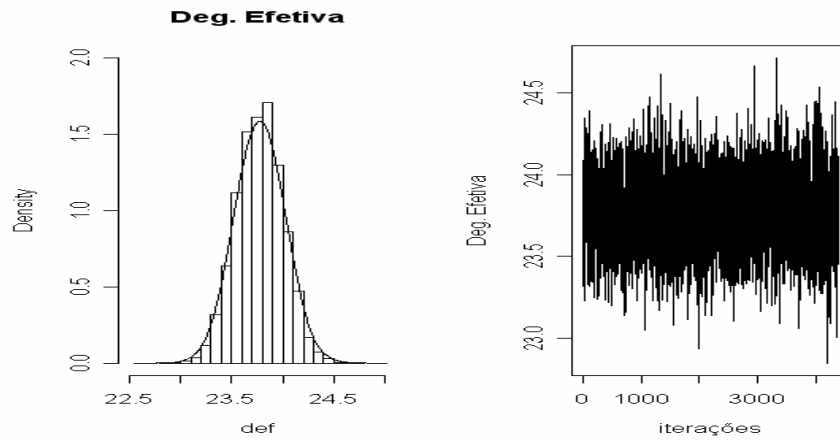


FIGURA 21. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade efetiva, em função dos parâmetros do modelo de Orskov & McDonald (1979).

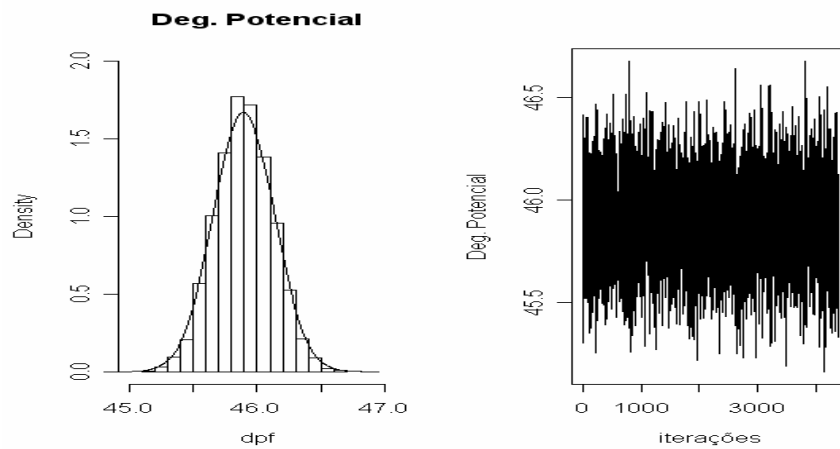


FIGURA 22. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade potencial, em função dos parâmetros do modelo de Orskov & McDonald (1979).

Pelo gráfico da Figura 17 observa-se que a cadeia gerada para o parâmetro a do modelo oscilou em torno da média estimada, não apresentando

picos discrepantes ao longo da cadeia. O mesmo comportamento pode ser observado para a cadeia dos demais parâmetros (Figuras 18 a 20), bem como para as cadeias geradas para as degradabilidades efetiva e potencial (Figuras 21 e 22).

Com base na simetria das densidades e dos histogramas, apresentados nas Figuras de 17 a 22, podem-se propor distribuições *a priori* do tipo simétrica, como a distribuição normal ou a distribuição t de Student.

Quanto às estimativas dos parâmetros, destaca-se o valor encontrado para a fração potencialmente degradável (*b*) que foi de, aproximadamente, 32%. No entanto, a degradabilidade efetiva (DE) da FDN foi próxima de 28%, mostrando que o alimento pode permanecer mais tempo no rúmen para atingir seu máximo potencial de degradação. Oliveira et al. (1999) encontraram valores para degradabilidade efetiva da FDN do feno de capim-coastcross em torno de 39,75% para uma taxa de passagem de 2,66%/hora. Malafaia et al. (1998) determinaram as frações que constituem os carboidratos totais e a cinética de degradação ruminal da FDN de alguns alimentos, entre eles o capim Tifton 85 e o feno do capim-coastcross. Estes volumosos apresentaram 29,4% e 21,1% de degradabilidade efetiva da FDN, respectivamente, para taxa de passagem de 3%/h.

Assis et al. (1999), avaliando a degradabilidade efetiva da FDN do capim Tifton85, cortado aos 35 dias, encontraram, para este parâmetro, um valor de 34,26%. Mendonça (1983) afirma que a composição química das forrageiras é afetada, principalmente, pela idade da planta, a fertilidade do solo e o clima, podendo, ainda, variar entre espécies e cultivares. O avanço da idade interfere no valor nutritivo de gramíneas forrageiras por meio de mudanças na composição química, devido ao aumento do teor de fibra e ao decréscimo do teor de proteína bruta.

Gonçalves (2001) observou, para Tifton85, Tifton44 e coastcross, valores médios de 31,0% de DE da FDN, para uma taxa de passagem de 2%/h. Malafaia et al. (1998) também relataram DE da FDN de 29,0% para a gramínea Tifton85, semelhante ao obtido neste experimento.

Para a fração solúvel (parâmetro a), foi encontrado, neste estudo, um valor de 14% para a gramínea coastcross cortada aos 120 dias. Resultado semelhante foi encontrado por Soares et al. (2008), ao compararem equações de predição de consumo com valores obtidos por métodos diretos utilizando capim-elefante fornecido picado para vacas mestiças lactantes e canuladas no rúmen, que encontraram, para fração solúvel (parâmetro a), um valor próximo a 19% para o capim-elefante, porém este cortado aos 60 dias. Savian et al. (2007), estudando a degradabilidade do capim-coastcross, por meio do ajuste do modelo de Orskov & McDonald (1979), encontraram um valor, para a fração solúvel de FDN, de 14,2%, resultado bastante semelhante ao encontrado neste estudo.

4.4.2 Modelo de McDonald (1981) – Modelo 2

A convergência das cadeias de cada um dos parâmetros do modelo de McDonald (1981), bem como da degradabilidade efetiva (DE) e potencial (DP), foi monitorada pelos mesmos critérios adotados para o modelo anterior, ou seja, visualização gráfica do traço da cadeia gerada pelo método MCMC, pelos critérios propostos Geweke (1992), Gelman & Rubin (1992), Heidelberger & Welch (1993) e Raftery & Lewis (1992), além do Erro de Monte Carlo (EMC).

Pelo critério de Geweke, o valor p obtido para esse modelo também foi sempre maior que o nível de significância pré-fixado de 5%, não apresentando evidências contra a convergência. Para a análise do critério de Gelman e Rubin, foram consideradas duas cadeias com valores iniciais distintos em cada

parâmetro e os valores estimados de R , que podem ser observados na Tabela 9, também não apresentam qualquer evidência contra a convergência das cadeias.

A regra de decisão do critério de Raftery & Lewis (1992), baseado no fator de dependência (FD), mostra que, para a análise dos dados reais a partir no modelo de McDonald (1981), a cadeia amostrada de todos os parâmetros também atingiu a convergência (Tabela 9).

TABELA 9. Critérios de Geweke (valor p), Gelman e Rubin (\hat{R}) e Raftery e Lewis (fator de dependência - FD), utilizados na detecção da convergência das cadeias amostradas.

Parâmetros	Geweke (valor p)	Gelman e Rubin (\hat{R})	Raftery e Lewis (FD)
a	0,5716	0,9998	0,989
b	0,4790	1,0001	0,989
c	0,3406	1,0014	0,964
L	0,4244	1,0037	4,145
τ	0,8818	0,9998	0,977
¹ DE	0,7642	0,9996	1,015
² DP	0,5258	1,0013	1,001

¹ Degradabilidade Efetiva com taxa de passagem de 5%/h; ² Degradabilidade Potencial.

Os valores encontrados para o Erro de Monte Carlo (EMC) foram de baixa magnitude (Tabela 10) em relação às estimativas *a posteriori* dos parâmetros, assegurando a convergência das cadeias. Segundo o critério de Heidelberg & Welch (1993), a aceitação da hipótese nula de estacionariedade indica que não é necessário um número maior de iterações além das computadas na análise. As cadeias geradas para as degradabilidades efetiva (DE) e potencial (DP) também tiveram sua convergência constatada (Tabela 9).

Observam-se, por meio da Tabela 10, os valores estimados da média e moda *a posteriori*, o intervalo de máxima densidade *a posteriori* e Erro de Monte Carlo (EMC). Os resultados mostram que as estimativas da média e da

moda *a posteriori* estão bastante próximas em todos os parâmetros, exceto para o tempo de colonização, L , que apresenta distribuição assimétrica à direita e, em estudos futuros, pode ser considerada, para este parâmetro, uma distribuição *a priori* do tipo gama. As distribuições do tipo simétrica, para os demais parâmetros do modelo, demonstram que um resumo pontual da informação contida nas distribuições *a posteriori* pode ser dado tanto pela média *a posteriori* como pela moda *a posteriori*. Para o parâmetro L do modelo de McDonald (1981), a estimativa que melhor resume a distribuição *a posteriori* é a moda *a posteriori* e não a média.

TABELA 10. Valores estimados da média e moda *a posteriori*, e intervalo de máxima densidade *a posteriori* (HPD) para os parâmetros do modelo de McDonald (1981).

Parâmetros	Média	Moda	Intervalo HPD 95%		EMC
			¹ LI	LS	
a	3,10	3,15	1,80	4,52	$9,04 \times 10^{-3}$
b	44,42	44,41	42,24	46,74	$1,61 \times 10^{-2}$
c	0,041	0,041	0,040	0,043	$1,12 \times 10^{-5}$
L	0,045	0,013	$5,8 \times 10^{-5}$	0,14	$8,32 \times 10^{-4}$
τ	0,44	0,43	0,39	0,47	$2,43 \times 10^{-4}$
² DE	23,22	23,24	22,33	24,18	$5,20 \times 10^{-3}$
³ DP	47,53	47,51	46,06	49,09	$1,02 \times 10^{-2}$

¹Limites inferior e superior do intervalo HDP; ²Degradabilidade Efetiva com taxa de passagem de 5%/h; ³ Degradabilidade Potencial.

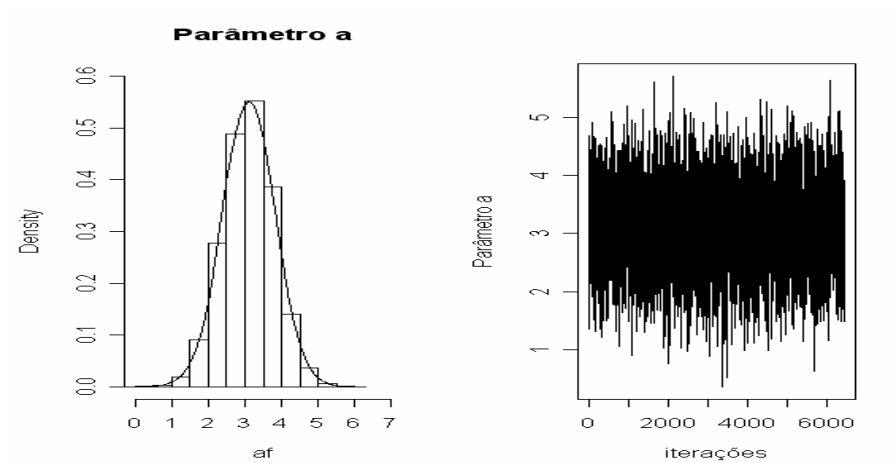


FIGURA 23. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro *a* do modelo de McDonald (1981).

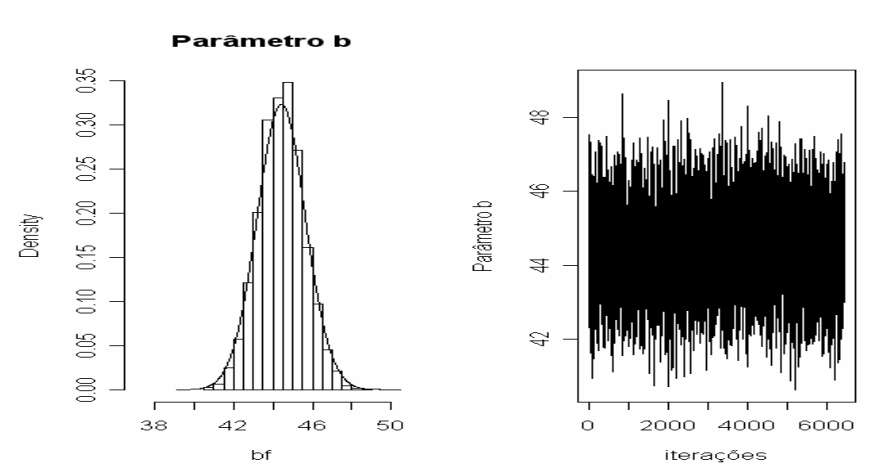


FIGURA 24. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro *b* do modelo de McDonald (1981).

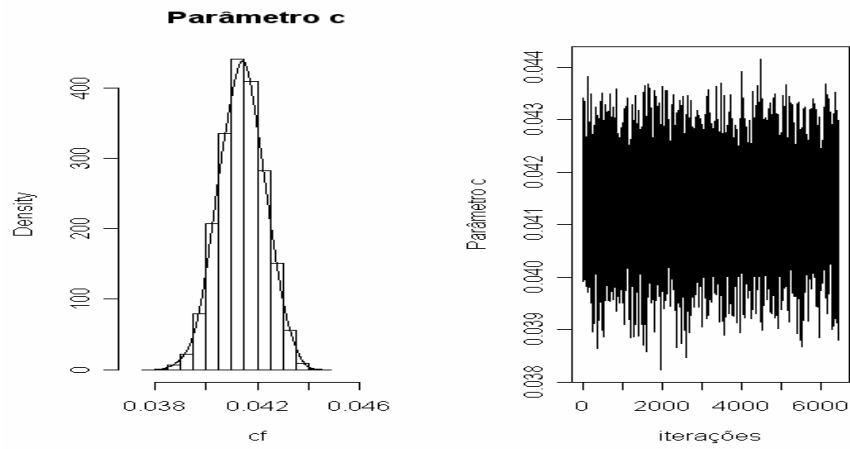


FIGURA 25. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro c do modelo de McDonald (1981).

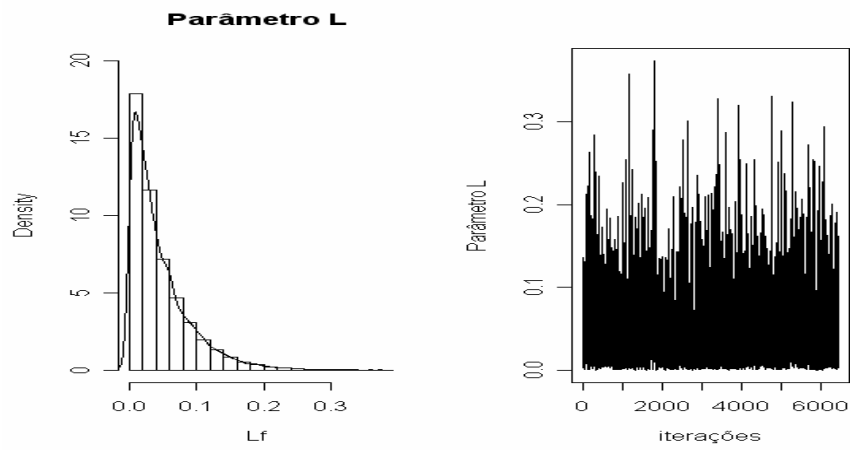


FIGURA 26. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para o parâmetro L do modelo de McDonald (1981).

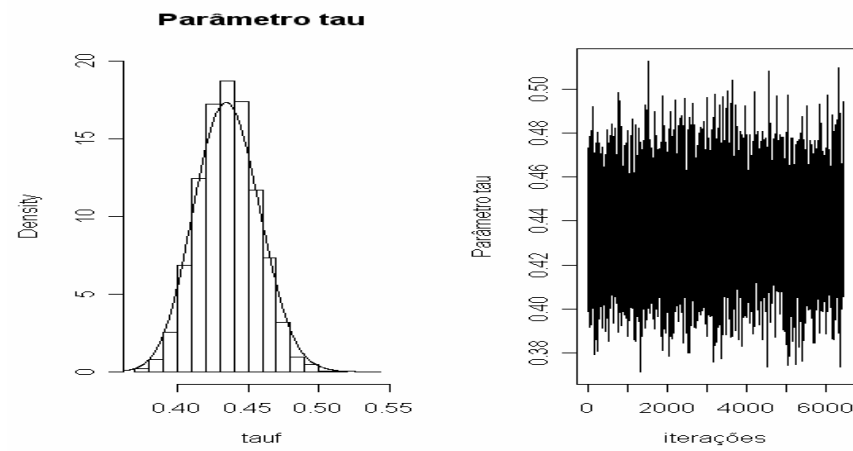


FIGURA 27. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a precisão.

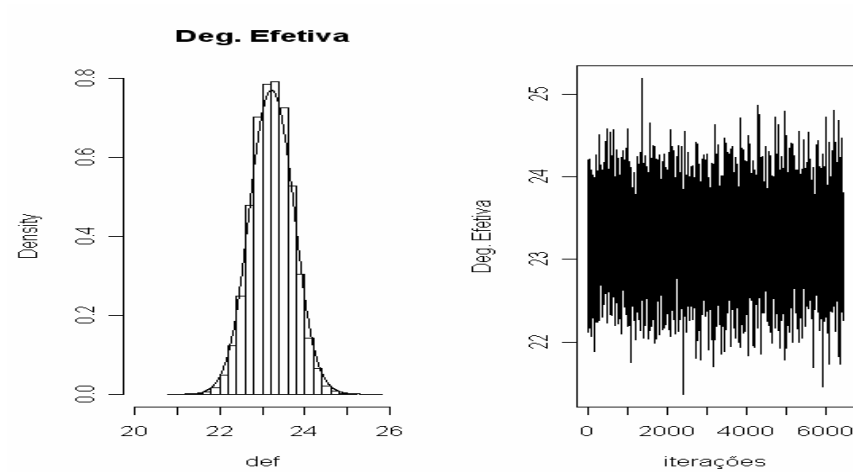


FIGURA 28. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade efetiva, em função dos parâmetros do modelo de McDonald (1981).

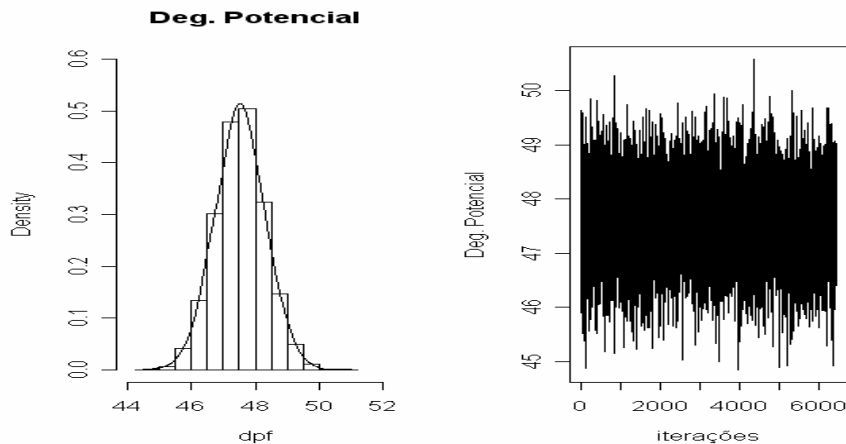


FIGURA 29. Gráfico da distribuição marginal *a posteriori* e traço da cadeia gerada pelo método MCMC, para a degradabilidade potencial, em função dos parâmetros do modelo de McDonald (1981).

A taxa de degradação da FDN, em torno de 4,10%/h (Tabela 10), está dentro da faixa sugerida pelo CNCPS, entre 3,00% e 8,10%/h, para plantas tropicais. As taxas de degradação da FDN obtidas por Ribeiro et al. (2001), para plantas de Tifton, entre 4,00% e 4,60%/h, são semelhantes às taxas encontradas no presente trabalho, ao contrário das obtidas por Vieira et al. (2000), em pastagens nativas, para os períodos de secas e águas, que foram, respectivamente de 9% e 6,2%/h.

Schmidt et al. (2007), estudando parâmetros ruminais e degradabilidade de frações fibrosas de cana-de-açúcar, encontraram teores médios estimados para as frações *a* e *b* de 8,45% e 34,5%, para degradação de FDN. No presente estudo encontrou-se, para a fração solúvel (parâmetro *a*), o teor de 3,10% e, para a fração potencialmente degradável (parâmetro *b*), o teor de 44,42%. Teoricamente, a proporção da fração “*a*” na FDN deveria ser próxima a zero, por ser uma fração prontamente solúvel no rúmen, no tempo zero de incubação. O valor encontrado para fração *a* deve, possivelmente, ser em virtude da perda

de partículas diminutas durante a lavagem dos sacos não incubados. O mesmo foi observado por outros autores em silagens de gramíneas tropicais (Balsalobre, 2002; Paziani, 2004).

Avaliando a degradabilidade da FDN de silagens de capim-tanzânia, Paziani (2004) verificou valores médios de *b* e *c* e tempo de colonização de 58,8%; 4,5%/hora e 3,0 horas, respectivamente. Pires et al. (2004) avaliaram a degradabilidade ruminal do bagaço de cana-de-açúcar e verificaram teores de fração *b* de 44,0% para a FDN. No presente estudo, a degradabilidade potencial da FDN apresentou valor médio de 47,5% e a degradabilidade efetiva da FDN estimada foi de 23,2% para a taxa de passagem de 5%/hora. Esses valores são bastante semelhantes aos encontrados por Schmidt et al. (2007).

Tonani et al. (2001) estudaram a degradabilidade ruminal *in situ* da FDN em silagens de híbridos de sorgo colhidos em diferentes épocas e encontraram, em média, valores para o tempo de colonização de 4,5 horas, os quais estão bem acima dos encontrados nesse estudo (Tabela 10). Veloso et al. (2000), estudando a degradabilidade ruminal da fibra de cinco forrageiras tropicais, encontraram que o tempo de colonização variou de duas a sete horas, sendo o mais rápido para a soja e o mais demorado para a leucena. Os valores encontrados pelos autores também estão bem acima dos encontrados neste estudo (Tabela 10).

4.4.3 Comparação dos modelos

O resultado da comparação dos modelos de Orskov e McDonald (modelo 1) e de McDonald (modelo 2), ajustados aos dados reais, por meio do Fator de Bayes, encontra-se na Tabela 11. Se o Fator de Bayes for maior que 1, o modelo 1, contemplado no numerador, é mais plausível ou verossímil do que o modelo 2. Então, opta-se por escolher o modelo 1 para predizer o fenômeno em

questão. Segundo Jeffreys (1961), as magnitudes atingidas pelo Fator de Bayes demonstram os graus de plausibilidade do modelo.

TABELA 11. Fator de Bayes para a comparação dos modelos de Orskov & McDonald (1979) e McDonald (1981), no ajuste aos dados reais.

Modelos	FB_{12}
Orskov e McDonald (1979)	651,16
McDonald (1981)	

Os resultados na Tabela 11 demonstram que o modelo de Orskov & McDonald (1979) é mais plausível que o modelo proposto por McDonald (1981) e que a evidência a favor deste modelo é considerada muito forte.

5 CONCLUSÕES

Diante dos resultados obtidos, pode-se concluir que:

- i.** a metodologia bayesiana mostrou-se bastante eficiente no estudo de modelos de degradabilidade ruminal, sendo esta avaliada e comprovada pelo processo de simulação, apresentando estimativas bem próximas ao valor verdadeiro;
- ii.** as estimativas obtidas para os parâmetros do modelo de degradabilidade ruminal de Orskov & McDonald (1979) por meio da abordagem bayesiana mostraram-se bastante coerente com os valores relatados na literatura;
- iii.** as estimativas obtidas para os parâmetros do modelo de degradabilidade ruminal de McDonald (1981) por meio da abordagem bayesiana mostraram-se bastante coerente com os valores relatados na literatura;
- iv.** o modelo de Orskov & McDonald (1979) na descrição dos dados de porcentagem de degradação da fibra em detergente neutro do capim-coastcross, com idade de corte de 120 dias, foi mais plausível que o modelo de McDonald (1981).

6 REFERÊNCIAS BIBLIOGRÁFICAS

AGRICULTURAL AND FOOD RESEARCH COUNCIL – AFRC. Technical Committee on responses to nutrients: Nutritive requirements of ruminant animals: protein. **Nutrition Abstracts and Reviews**: series B, London, v. 62, n. 9, p. 65-71, Sept. 1992.

ASSIS, M. A.; SANTOS, G. T.; CECATO, U.; DAMASCENO, J. C.; PETIT, H. V.; BETT, W.; GOMES, L. H.; DANIEL, M. Degradabilidade *in situ* de gramíneas do gênero *Cynodon* submetidas ou não a adubação nitrogenada. **Acta Scientiarum**, Maringá, v. 21, n. 3, p. 657-663, 1999.

BALDWIN, R. L.; THORNLEY, J. H. M.; BEAVER, D. E. Metabolism of the lactating cow: II. digestive elements of a mechanistic model. **Journal of Dairy Research**, Cambridge, v. 54, n. 1, p. 107-131, Feb. 1987.

BALSALOBRE, M. A. A. **Valor alimentar do capim Tanzânia irrigado**. 2002. 113 p. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, SP.

BLASCO, A.; PILES, M.; VARONA, L. A. Bayesian analysis of the effect of selection for growth rate curves in rabbits. **Genetics Selection Evolution**, Valencia, v. 35, p.21-41, 2003.

BOX, G. E. P.; TIAO, G. C. **Bayesian inference in statistical analysis**. New York: J. Wiley, 1992. 588 p.

BROEMELING, L. D. **Bayesian analysis of linear models**. New York: J. Wiley, 1989. 412 p.

CHIACCHIO, E. J. **Regressão não-linear**: desenvolvimento de um sistema computacional e aplicações. 1993. 149 p. Dissertação (Mestrado em Estatística e Experimentação Agronômica) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, SP.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 3. ed. New York: J. Wiley, 1998. 706 p.

EHLERS, R. S. **Introdução à inferência bayesiana**, 2007. 68p. Disponível em: <<http://leg.ufpr.br/~ehlers/bayes>>. Acesso em: 10 mar. 2007.

- FRANCE, J.; THORNLEY, H. M.; BEEVER, D. E. A mathematical model of the rumen. **Journal of Agricultural Science**, Cambridge, v. 99, n. 2, p. 343-353, Oct. 1982.
- GELFAND, A. E.; SMITH, A. F. M. Sampling based approaches for calculating marginal densities. **Journal of the American Statistical Association**, London, v. 85, n. 410, p. 398-409, June 1990.
- GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian data analysis**. London: Chapman Hall, 1997. 526 p.
- GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical Science**, Hayward, v. 7, n. 4, p. 457-511, 1992.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. **IEEE Transactions on pattern Analysis and Machine Intelligence**, Los Alamitos, v. 6, n. 6, p. 721-741, 1984.
- GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Ed.). **Bayesian Statistics 4**. New York: Oxford University, 1992. p.625-631.
- GONÇALVES, G. D. **Avaliação nutricional de gramíneas do gênero *Cynodon***. 2001. 91 p. Dissertação (Mestrado em Zootecnia) – Universidade Estadual de Maringá, Maringá.
- HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, London, v. 57, p. 97-109, 1970.
- HEIDELBERG, P.; WELCH, P. Simulation run length control in the presence of an initial transient. **Operations Research**, Landing, v. 31, n. 6, p. 1109-1144, Nov./Dec. 1993.
- JEFFREYS, H. **Theory of probability**. Oxford: Claredon, 1961. 325 p.
- KASS, R. E.; RAFTERY, A. E. Bayes factor. **Journal of the American Statistical Association**, Alexandria, v. 90, n. 430, p. 773-795, June 1995.
- KHURI, A. I.; CORNELL, J. A. **Response surface: designs and analyses**. New York: M. Dekker, 1987.

MALAFAIA, P. A. M.; VALADARES FILHO, S. C.; VIEIRA, R. A. M.; SILVA, J. F. C.; PEREIRA, J. C. Determinação das frações que constituem os carboidratos totais e da cinética ruminal da fibra em detergente neutro de alguns alimentos para ruminantes. **Revista Brasileira de Zootecnia**, Viçosa, v. 27, n. 4, p. 790-796, 1998.

MAZUCHELI, J.; ACHCAR, J. A. Algumas considerações em regressão não linear. **Acta Scientiarum**, Maringá, v. 24, n. 6, p. 1761-1770, 2002.

McDONALD, I. A revised model for the estimation of protein degradability in the rumen. **Journal of Agricultural Science**, Cambridge, v. 96, p. 251-252, 1981.

MEHREZ, A. Z.; ORSKOV, E. R. A study of the artificial fibre bag technique for determining the digestibility of feeds in the rumen. **Journal of Agricultural Science**, Cambridge, v. 88, n. 3, p. 645-650, 1977.

MENDONÇA, J. F. B. **Rendimento e valor nutritivo do campim elefante (*Pennisetum purpureum* Schum) cv. Cameroon**. 1983. 110 p. Dissertação (Mestrado em Zootecnia) – Escola Superior de Agricultura de Lavras, Lavras, MG.

MERTENS, D. R. Dietary fiber components: relationship to the rate and extent of ruminal digestion. **Federation Proceedings**, New York, v. 36, n. 2, p. 187-192, 1977.

MERTENS, D. R.; ELY, L. O. A dynamic model of fiber digestion and passage in the ruminant for evaluating forage quality. **Journal of Animal Science**, Champaign, v. 49, n. 4, p. 1085-1095, Apr. 1979.

MERTENS, D. R.; LOFTEN, J. R. The effects of starch on forage fiber digestion kinetics *in vitro*. **Journal of Dairy Science**, Champaign, v. 63, p. 1437-46, 1980.

MERTENS, D. R. Rate and extent of digestion. In: FORBES, J. M.; FRANCE, J. (Ed.). **Qualitative aspects of ruminant digestion and metabolism**. Wallingford, UK: Cambridge University, 1993. cap. 2, p. 13-51.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equations of state calculations by fast computing machines. **Journal of Chemical Physics**, Woodbury, v. 21, n. 6, p. 1087-1092, 1953.

NOCEK, J. E. In situ and other methods to estimate ruminal protein and energy digestibility: a review. **Journal of Dairy Science**, Champaign, v. 71, n. 8, p. 2051-2069, Aug. 1988.

NOGUEIRA, D. A.; SÁFADI, T.; FERREIRA, D. F. Avaliação de critérios de convergência para o método de Monte Carlo via Cadeias de Markov. **Revista Brasileira de Estatística**, Rio de Janeiro, v. 65, n. 224, p. 59-88, 2004.

OLIVEIRA, R. L.; PEREIRA, J. C.; SILVA, P. R. C.; VIEIRA, R. A. M. Degradabilidade ruminal da cama de frango e do feno de capim Coastcross e avaliação de modelos matemáticos para estimativa da taxa de passagem de partículas. **Revista Brasileira de Zootecnia**, Viçosa, v. 28, n. 4, p. 839-849, 1999.

ORSKOV, E. R.; McDONALD, I. The estimation of protein degradability in the rumen from incubation measurements weighted according to rate of passage. **Journal of Agricultural Science**, Cambridge, v. 92, n. 2, p. 499-503, 1979.

ORSKOV, E. R. **Protein nutrition in ruminants**. London: Academic, 1982. 160 p.

PAULINO, C. D.; TURKMAN, A. A.; MURTEIRA, B. **Estatística Bayesiana**, Lisboa: Fundação Calouste Gulbenkian, 2003. 446 p.

PAZIANI, S. F. **Controle de perdas na ensilagem, desempenho e digestão de nutrientes em bovinos de corte alimentados com rações contendo silagens de capim tanzânia**. 2004. 208 p. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, SP.

PEREIRA, J. C. **Degradación ruminal de diversos subproductos agroindustriales**. 1992. 223 p. Tese (Doutorado em Produção Animal) – Universidad Politécnica de Madrid, Madrid.

PIRES, A. J. V.; GARCIA, R.; VALADARES FILHO, S. C.; PEREIRA, G. O.; CECON, P. R.; SILVA, F. F. da; SILVA, P. A.; ÍTAVO, L. C. V. Degradabilidade do bagaço de cana-de-açúcar tratado com amônia anidra e, ou, sulfeto de sódio. **Revista Brasileira de Zootecnia**, Viçosa, v. 33, n. 4, p. 1071-1077, 2004.

QUINN, J. L.; WATH, J. G. van der; MYBURGH, S. Studies on the alimentary tract of merino sheep in southern Africa: IV. description of experiment technique. **Journal of Veterinary Science and Animal Industry**, Onderstepoort, v. 11, n.

2, p. 341-360, 1939. In: GILKS, W. R.; SPIEGELHALTER, D. J.; RICHARDSON, S. (Ed). **Practical Markov Chain Monte Carlo**. Oxford, USA: University Press, 1995. p.163-189.

RAFTERY, A. L.; LEWIS, S. Comment: one long run with diagnostics: implementation strategies for Markov chain Monte Carlo. **Statistical Science**, Hayward, v. 7, n. 4, p. 493-497. 1992.

R DEVELOPMENT CORE TEAM. **The R Manuals**: current version: 2. 6. 0. 2007. Disponível em: <<http://www.r-project.org>>. Acesso em: 10 mar. 2007.

REIS, R. A.; RODRIGUES, L. R. A. **Valor nutritivo de plantas forrageiras**. Jaboticabal: UNESP/FUNEP, 1993. 26 p.

REIS, S. T. dos. **Valor nutricional de gramíneas tropicais em diferentes idades de corte**. 2000. 99 p. Dissertação (Mestrado em Zootecnia) – Universidade Federal de Lavras, Lavras, MG.

RIBEIRO, K. G.; GARCIA, R.; PEREIRA, O. G.; VALADARES FILHO, S. de C.; CECON, P. R. Consumo e digestibilidades aparentes total e parcial de nutrientes, em bovinos recebendo rações contendo feno de capim-Tifton 85 de diferentes idades de rebrota. **Revista Brasileira de Zootecnia**, Viçosa, v. 30, n. 2, p. 573-580, 2001.

ROSA, G. J. M. **Análise Bayesiana de modelos lineares mistos robustos via Amostrador de Gibbs**. 1998. 57 p. Tese (Doutorado em Estatística e Experimentação Agronômica) – Universidade de São Paulo, Piracicaba, SP.

SAVIAN, T. V.; MUNIZ, J. A. A study of *in situ* degradability: heterogeneity of variances and correlated errors. **Scientia Agrícola**, Piracicaba, v. 64, n. 5, p. 548-554, 2007.

SAVIAN, T. V.; MUNIZ, J. A.; AQUINO, L. H. de; BANYS, V. L.; FERREIRA, D. F. Estudo da degradabilidade da fibra em detergente neutro do Coast Cross (*Cynodon dactylon x Cynodon nlemfunensis*). **Ciência e Agrotecnologia**, Lavras, v. 31, n. 4, p. 1184-1190, 2007.

SCHMIDT, P.; NUSSIO, L. G.; ZOPOLLATTO, M.; RIBEIRO, J. L.; SANTOS, V. P. dos; PIRES, A. V. Chemical and biological additives in sugar cane silages: 2. ruminal parameters and DM and fiber degradabilities. **Revista Brasileira de Zootecnia**, Viçosa, v. 36, suppl, p. 1676-1684, 2007.

SEARLE, S. R. **Linear models**. New York: J. Willey, 1971.

SILVA, F. F.; MUNIZ, J. A.; AQUINO, L. H. de.; SÁFADI, T. Abordagem bayesiana da curva de lactação de cabras Saanen de primeira e segunda ordens de parto. **Pesquisa Agropecuária Brasileira**, Brasília, v. 40, n. 1, p. 27-33, jan. 2005.

SILVA, N. A.; SILVA, F. F.; MUNIZ, J. A.; AQUINO, L. H. de. Estudo de parâmetros de crescimento de bezerros Nelore por meio de um modelo de regressão linear: uma abordagem bayesiana. **Ciência Animal Brasileira**, Goiânia, v. 7, n. 1, p. 57-65, jan./mar. 2006.

SOARES, J. P. G.; BERCHIELLI, T. T.; SALMAN, A. K. D.; DERESZ, F.; OLIVEIRA, A. D.; VERNEQUE, R. S. Comparações entre equações de predição de consumo de capim-elefante cortado para vacas lactantes. **Archivos de Zootecnia**, Córdoba, v. 57, p. 235-246, 2008.

SOUZA, G. da S. **Introdução aos modelos de regressão linear e não-linear**. Brasília: Embrapa-SPI/Embrapa-SEA, 1998. 489 p.

STENSIG, T.; WEISBJERG, M. R.; HVELPLUND, T. Estimation of ruminal digestibility of NDF from in sacco degradation and rumen fractional outflow rate. **Acta Agriculture Scandinavica**, Stockholm, v. 44, n. 1, p. 96-109, 1994.

TEIXEIRA, J. C.; HUBER, J. T.; WANDERLEY, R. C. A mobile nylon bag technique for estimating post-ruminal digestibility in dairy cows. **Journal of Dairy Science**, Champaign, v. 71, n. 1, p. 180, June 1988. (Abstr.)

TONANI, F. L.; RUGGIERI, A. C.; QUEIROZ, A. C.; ANDRADE, P. Ruminal in situ degradability of dry matter and neutral detergent fiber of sorghum (*Sorghum bicolor* L.) silages with different harvesting times. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 53, n. 1, p. 100-104, 2001.

VELOSO, C. M.; RODRIGUEZ, N. M.; SAMPAIO, I. B. M. et al.; GONÇALVES, L. C.; MOURÃO, G. B. Ruminal pH and ammonia, tropical forage leaves: stems ratio and fiber ruminal degradability. **Revista Brasileira de Zootecnia**, Viçosa, v. 29, n. 3, p. 871-879, 2000.

VIEIRA, R. A. M.; PEREIRA, J. C.; MALAFAIA, P. A. M. Fracionamento e cinética de degradação *in vitro* dos compostos nitrogenados da extrusa de bovinos a pasto. **Revista Brasileira de Zootecnia**, Viçosa, v. 29, n. 3, p. 880-888, 2000.