

UMA ABORDAGEM BAYESIANA PARA MODELAR A ISOTERMA DE LANGMUIR

Dialilson Teixeira de CARVALHO¹
Luiz Alberto BEIJO¹
Joel Augusto MUNIZ²

- RESUMO: O objetivo deste trabalho foi utilizar o método Bayesiano no ajuste da isoterma de Langmuir considerando distribuições *a priori* informativas e não informativas. Realizou-se um estudo de simulação de dados considerando diferentes tamanhos amostrais para avaliar a precisão e acurácia das estimativas dos parâmetros de afinidade (k) e capacidade máxima de adsorção (M), obtidas com diferentes distribuições *a priori* informativas normais e uma não informativa uniforme, juntamente com as estimativas do parâmetro τ , para o qual foram propostas distribuições *a priori* gama informativa e não informativa. Amostras das distribuições marginais *a posteriori* dos parâmetros da isoterma foram obtidas pelo amostrador de Gibbs. As inferências foram realizadas e os resultados indicaram que o método Bayesiano é eficiente e as estimativas obtidas com uso das distribuições *a priori* informativas dos parâmetros apresentaram maiores precisão e acurácia mesmo em tamanhos amostrais inferiores. Realizou-se então um ajuste da isoterma sobre dados experimentais de adsorção, considerando as distribuições *a priori* propostas e os resultados obtidos corroboraram com o estudo de simulação, de modo que as estimativas obtidas com distribuições *a priori* informativas apresentaram maior precisão.
- PALAVRAS-CHAVE: Adsorção; informação *a priori*; regressão não linear; tamanho amostral.

1 Introdução

O alto desenvolvimento tecnológico e a expansão do agronegócio tem como provocado a geração cada vez maior de resíduos tóxicos das mais diversas

¹Universidade Federal de Alfenas - UNIFAL-MG, Instituto de Ciências Exatas, CEP: 37.130-000, Alfenas, MG, Brasil. E-mail: dialilson-carvalho@hotmail.com; luiz.beijo@unifal-mg.edu.br

²Universidade Federal de Lavras - UFLA, Departamento de Estatística, CEP: 37.200-000, Lavras, MG, Brasil. E-mail: joamuniz@des.ufla.br

composições, dentre elas os metais-traços. Tratam-se de metais como o chumbo, que mesmo em pequenas quantidades (da ordem de *ppm*) podem apresentar elevada toxicidade aos seres vivos. Para remoção destes metais são empregados vários processos, dos quais um dos mais comuns é a adsorção (BARBOSA *et al.*, 2014). Este processo pode ser entendido como a aderência ou acúmulo de substâncias sobre uma superfície mineral ou de partículas sólidas, que são a matriz adsorvente (ATKINS e JONES, 2011). Quanto aos estudos de adsorção de metais traço em laboratório, um problema enfrentado é a geração de materiais contaminados, os quais ocasionam uma elevação de custos, principalmente devido à necessidade de tratamento e descarte específicos dos resíduos tóxicos que são gerados. Desse modo há uma limitação prática na quantidade de amostra disponível.

A adsorção pode ser descrita por modelos estatísticos chamados isotermas de adsorção (PEREIRA e SILVA, 2009), dos quais o mais citado na literatura científica é a isoterma de Langmuir (RODRIGUES *et al.*, 2014). Esta expressão está no trabalho de Langmuir (1918) e pode ser descrita por:

$$q_i = \frac{M \times k \times c_i}{1 + k \times c_i} + e_i, \quad i = 1, \dots, n, \quad (1)$$

em que q_i refere-se à quantidade adsorvida de metal no equilíbrio (mgg^{-1}), c_i é a concentração de metal no equilíbrio (mgL^{-1}), M é a capacidade máxima de adsorção (mgg^{-1}), k é a "constante" de adsorção de Langmuir (Lmg^{-1}), e_i trata do erro aleatório, assumindo que o mesmo é independente e normalmente distribuído com média 0 e variância constante σ^2 , n é o tamanho amostral.

A estimação dos parâmetros deste modelo geralmente é feita empregando-se técnicas baseadas no método dos mínimos quadrados. Os estimadores fornecidos por este método apresentam boas propriedades, tais como robustez, alto poder e baixo erro quadrático médio (MOOD *et al.*, 1982) assintoticamente, entretanto, para pequenas quantidades de amostras, não é assegurado que tais propriedades sejam mantidas (MAZUCHELI e ACHCAR, 2002). Diante dessa limitação prática na quantidade de amostra disponível em estudos de adsorção, a qualidade do ajuste da isoterma pode ser comprometida, não havendo garantias de que as estimativas mantenham boas propriedades como alta precisão e acurácia.

Uma alternativa para contornar esta limitação é a aplicação da inferência Bayesiana, a qual não possui pressuposições fortes acerca do tamanho amostral. A inferência Bayesiana baseia-se em uma generalização do teorema de Bayes, em que o parâmetro é assumido como uma variável aleatória, podendo ser utilizada uma distribuição de probabilidade para quantificar a incerteza relacionada ao mesmo (DEGROOT e SCHERVISH, 2012). De acordo com Paulino *et al.* (2003), toda informação probabilística a respeito do parâmetro é fornecida por sua distribuição *a posteriori* $p(\theta | q)$, a qual é proporcional ao produto de uma distribuição *a priori* $p(\theta)$ pela verossimilhança $L(\theta | q)$:

$$p(\theta | q) \propto L(\theta | q) p(\theta) \quad (2)$$

A expressão (2) pode ser entendida como *posteriori* \propto *verossimilhança* \times

priori, em que q é a variável resposta quantidade de metal adsorvida e α representa proporcionalidade.

De acordo com Degroot e Schervich (2012) e Paulino *et al.* (2003), quando são disponíveis informações, passíveis de serem especificadas probabilisticamente, sobre uma quantidade de interesse θ , constitui-se uma distribuição a *priori*, trata-se de uma distribuição de probabilidade de um parâmetro θ antes da obtenção dos dados experimentais, sendo esse parâmetro abordado como uma variável aleatória.

Segundo Paulino *et al.* (2003), quando se dispõe de informações, contidas de elementos subjetivos, radicadas em fontes objetivas (dados históricos do problema ou de situações análogas, fatos, artigos científicos ou experiência de um especialista) que levam o especialista a formalizar probabilisticamente suas crenças sobre as incertezas relacionadas a θ , tem-se uma *priori* informativa.

Por outro lado, conforme Reis *et al.* (2008) e Resende (2000), se em determinado estudo o pesquisador tem pouca ou nenhuma informação para ser incorporada na modelagem, considera-se uma *priori* não informativa ou vaga. Existem algumas formas de implementar uma *priori* não informativa, tais como o método de Jeffreys (1961), ou com uma *priori* uniforme, entre outros. Para Box e Tiao (1992), quando a verossimilhança é dominante ou deseja-se representar o desconhecimento sobre θ , a ideia inicial é considerar todos os possíveis valores para θ como equiprováveis, da forma $p(\theta) \propto \nu$, sendo ν uma constante e θ limitado em um intervalo real.

Considerando θ um vetor de parâmetros e q a variável resposta quantidade de metal adsorvida, para se inferir sobre qualquer elemento θ_i , para $i > 1$, a distribuição a *posteriori* conjunta dos parâmetros, $p(\theta | q)$, deve ser integrada em relação a todos os elementos desse vetor (ROSA, 1998). Assim, se o interesse do pesquisador se concentra sobre determinado elemento de θ , tem-se a necessidade da obtenção da distribuição marginal $p(\theta_i | q)$:

$$p(\theta_i | q) = \int_{\theta \neq \theta_i} p(\theta | q) d\theta_{\theta \neq \theta_i} \quad (3)$$

Conforme Paulino *et al.* (2003), a integração da distribuição a *posteriori* conjunta para a obtenção das distribuições marginais pode não possuir expressão analítica, havendo a necessidade de utilizarem-se algoritmos iterativos de aproximação especializados, tais como o amostrador de Gibbs e o Algoritmo de Metropolis-Hastings (1970).

De acordo com Gamerman (1997), o amostrador de Gibbs é um processo Markoviano no qual a variável gerada só depende da variável gerada anteriormente. O amostrador fornece uma forma alternativa para gerar sucessivos valores de uma distribuição condicional completa, utilizando a simulação de Monte Carlo para a geração de dados ou mais especificamente, de valores para os parâmetros a partir de distribuições de probabilidades conhecidas. Mais detalhes podem ser encontrados em Gamerman (1997) e Nogueira (2004).

Os métodos MCMC (do inglês *Markov Chain Monte Carlo*) são simulações iterativas, por isso necessitam de métodos (critérios) para diagnóstico de

convergência. Gamerman (1997) define os métodos de convergência em formais e informais. Quanto aos métodos formais, Gamerman (1997) faz referência aos métodos propostos por Raftery e Lewis (1992), Geweke (1992), Heidelberger e Welch (1993). Os informais consistem na análise visual da trajetória da cadeia ao longo das iterações, de modo que conclui-se pela convergência quando o gráfico após um certo período apresenta exaustivamente o mesmo comportamento tanto quantitativamente quanto qualitativamente.

Diante da necessidade de se identificar uma metodologia alternativa para minimizar a questão da limitação prática com relação ao tamanho de amostra em estudos de adsorção, de modo que a precisão e a acurácia das estimativas dos parâmetros da isoterma de Langmuir não sejam afetadas, objetivou-se com este trabalho ajustar a isoterma de Langmuir sob enfoque Bayesiano considerando distribuições *a priori* não informativas e também informativas, avaliando, via simulação, a acurácia e a precisão das estimativas obtidas sob diferentes tamanhos amostrais. Pretende-se também avaliar as distribuições *a priori* propostas no ajuste da isoterma considerando-se os dados experimentais de adsorção obtidos por Souza *et al.* (2012).

2 Material e métodos

2.1 Dados

Os dados simulados referem-se à quantidade de chumbo (Pb^{2+}) adsorvido (mg) em cascas de laranja (g), considerando uma aproximação dos dados obtidos nos respectivos trabalhos de referência (Tabela 1). Estimou-se os parâmetros M e k e a variância residual do modelo, por meio da função *nls*, disponível no pacote *stats* do Sistema computacional R (R CORE TEAM, 2016).

Tabela 1 - Estimativas dos parâmetros, respectivos desvios-padrão, variância residual do modelo e referência

\hat{M}	DP	\hat{k}	DP	Var. res.	Referência
113,500	5,189	0,034	0,007	6,342	Feng et al., (2011)
89,764	7,348	0,030	0,007	5,075	Liang et al., (2009)
55,523	4,017	0,022	0,006	4,103	Massocatto et al., (2010)

DP: Desvio-padrão; Var. res.: variância residual.

A partir da Tabela 1, obtiveram-se os quartis das estatísticas descritivas dos parâmetros k e M , utilizando-se o 1º e o 3º quartis como valores aproximados nos cenários para a avaliação da precisão e acurácia das estimativas dos parâmetros para o primeiro cenário: $k = 0,025 Lmg^{-1}$ e $M = 72 mgg^{-1}$, para o segundo: $k = 0,032 Lmg^{-1}$ e $M = 100 mgg^{-1}$. Para cada cenário considerou-se os seguintes tamanhos amostrais (quantidade de concentrações) $n = 5, 6, 7, 9, 12, 15, 20, 25$ e 30 e os

valores destas concentrações c_i foram definidos no intervalo de 5 a 500 mgL^{-1} , com base nos trabalhos de Liang et al., (2009) e Feng e Guo (2012).

O processo de simulação foi realizado utilizando-se a isoterma de Langmuir (1), com o erro aleatório e_i seguindo uma distribuição normal com média 0 e variância constante $\sigma^2 = 4$, que foi obtida a partir da variância residual dos modelos, conforme apresentado na Tabela 1. Os valores da variável q_i foram obtidos a partir da expressão (1), utilizando os valores dos parâmetros M e k dos dois cenários, as concentrações c_i definidas e por último acrescidas do erro e_i . Foram realizadas 1.200 simulações para cada tamanho amostral em cada cenário.

Os dados experimentais foram obtidos a partir do trabalho de Souza *et al* (2012), no qual há a descrição de todo processo de condução do experimento e obtenção dos dados.

2.2 Distribuições *a priori*

Como distribuições *a priori* informativas, assumiu-se que os parâmetros k e M seguem uma distribuição normal, com hiperparâmetros média μ_0 e precisão τ_0 . Para especificação do hiperparâmetro média de ambas as distribuições *a priori* de M , utilizou-se a mediana das três estimativas de M apresentadas na Tabela 1. Para especificação do hiperparâmetro precisão (τ_{M0}), na *priori* normal I, utilizou-se o inverso do quadrado do maior valor do desvio-padrão dentre os três obtidos ($\tau_{M0} = 1/\sigma^2$), já na normal II, utilizou-se inverso do quadrado do produto do maior valor do desvio-padrão por 1,5 ($\tau_{M0} = 1/(\sigma \times 1,5)^2$).

Para a distribuição *a priori* normal I, os valores dos hiperparâmetros média μ_0 e desvio-padrão σ_0 da distribuição normal foram 90 e 7,35, ou seja, $M \sim N(90; 7,35)$. Para a distribuição *priori* normal II esses valores foram 90 e 11,02, respectivamente, ou seja, $M \sim N(90; 11,02)$.

Para especificação dos hiperparâmetros das distribuições *a priori* normais de k , utilizou-se o mesmo procedimento que para as de distribuições *a priori* de M . Para a de distribuição *a priori* normal I de k , os valores dos hiperparâmetros μ_0 e σ_0 foram 0,030 e 0,007, ou seja, $k \sim N(0,030; 0,007)$. Para a normal II esses valores foram 0,030 e 0,011 ($k \sim N(0,030; 0,011)$), respectivamente.

Para o parâmetro precisão τ , que é o inverso da variância residual, foi assumida uma distribuição *a priori* gama, pelo fato de que a variância segue uma distribuição gama inversa. Para especificação dos hiperparâmetros a_0 e b_0 , baseou-se no valor aproximado da média da raiz quadrada das variâncias residuais ($\sigma = 2$) dos modelos apresentados na Tabela 1. Tais valores foram $\alpha_0 = 4$ e $\beta_0 = 12$.

Como distribuições *a priori* não informativas, para k e M assumiu-se uma distribuição uniforme, com hiperparâmetros a e b com valores de 0,0001 e 1000 respectivamente, ou seja, $k \sim U(0,0001; 1000)$ e $M \sim U(0,0001; 1.000)$. Tais valores para os hiperparâmetros foram selecionados pelo fato de que k e M não assumem valores negativos. Para a precisão τ foi assumida uma distribuição gama, com a forma $\tau \sim G(0,01; 0,01)$, que é considerada vaga.

2.3 Obtenção das distribuições *a posteriori*

Com as respectivas distribuições *a priori* propostas e a verossimilhança da isoterma de Langmuir, obtiveram-se as distribuições *a posteriori* conjuntas. Considerando θ como sendo o vetor de parâmetros $\theta = (M, k, \tau)$ e as respectivas distribuições *a priori* não informativas, a distribuição *a posteriori* conjunta de θ é dada pela expressão:

$$P(\theta | q_i) \propto \frac{1}{b_{M0} - a_{M0}} \times \frac{1}{b_{k0} - a_{k0}} \times \frac{\beta_0^{\alpha_0} \tau^{\alpha_0 - 1} e^{-\beta_0 \tau}}{\Gamma(\alpha_0)} \times \frac{\tau^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}} \times \exp \left\{ -\frac{\tau}{2} \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 \right\} \quad (4)$$

Considerando-se as distribuições *a priori* informativas para os elementos de θ , a distribuição *a posteriori* conjunta é dada por:

$$P(\theta | q_i) \propto \exp \left\{ -\frac{\tau_{k0} \times (k - k_0)^2}{2} \right\} \times \exp \left\{ -\frac{\tau_{M0} \times (M - M_0)^2}{2} \right\} \times e^{-\beta_0 \tau} \times \tau^{\alpha_0 + \frac{n}{2} - 1} \times \exp \left\{ -\frac{\tau}{2} \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 \right\} \quad (5)$$

Como as expressões 4 e 5 não possuem solução analítica, foi necessário o emprego de um método iterativo para obtenção das marginais. O algoritmo escolhido foi o amostrador de Gibbs, que está implementado no programa *OpenBUGS*, utilizado em interface com o R por meio do pacote *R2OpenBUGS* (STURTZ *et al.*, 2005).

Cada cadeia de MCMC foi obtida utilizando-se 104.000 iterações, um *burn-in* de 4.000, e um pulo (*thin*) de 50. Foram descartadas as simulações em que houve a violação de pelo menos um dos critérios de convergência.

A convergência das cadeias foi monitorada utilizando-se o pacote CODA - *Convergence Diagnostics and Output Analysis Software for Gibbs Sampling Output* (PLUMMER *et al.*, 2006), disponível no sistema computacional R (R CORE TEAM, 2016). Os critérios para verificação da convergência das cadeias geradas por MCMC foram o critério de Raftery e Lewis (1992), sendo rejeitadas as cadeias em que o fator de dependência foi maior ou igual a 5, a estatística Z de Geweke (1992), sendo rejeitadas aquelas cadeias em que $|z| \geq 1,96$ e o critério de Heidelberger e Welch (1993), adotando o nível de significância de 5%, rejeitando-se as cadeias cujo *valor-p* foi menor que 5%, ($p < 0,05$).

2.3.1 Distribuições condicionais

A partir da distribuição *a posteriori* conjunta, obtida com distribuições *a priori* não informativas, dada pela expressão 4, a distribuição condicional para o parâmetro

M é dada por:

$$p(M | q_i, k, \tau) \propto \exp \left\{ -\frac{\tau}{2} \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 \right\} \quad (6)$$

Para o parâmetro k tem-se:

$$p(k | q_i, M, \tau) \propto \exp \left\{ -\frac{\tau}{2} \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 \right\} \quad (7)$$

As distribuições condicionais apresentadas nas expressões 6 e 7 não possuem expressões fechadas. Por outro lado, para o parâmetro τ tem-se a expressão:

$$p(\tau | q_i, M, k) \propto \tau^{\alpha_0 + \frac{n}{2} - 1} \times \exp \left\{ -\tau \times \left(\frac{1}{2} \times \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 + \beta_0 \right) \right\} \quad (8)$$

que é o núcleo de uma distribuição Gama cuja função de densidade de probabilidade é:

$$p(y | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \times y^{\alpha-1} \times \exp(-\beta \times y) \quad (9)$$

de parâmetros dados por:

$$\alpha = \alpha_0 + \frac{n}{2}$$

$$\beta = \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 + \beta_0$$

Considerando-se agora a distribuição *a posteriori* conjunta obtida com o uso de distribuições *a priori* informativas (expressão 5), obteve-se a distribuição condicional para o parâmetro M , dada por:

$$p(M | q_i, k, \tau) \propto \exp \left\{ -\frac{1}{2} \times \left[\tau_{M0} + \tau \times k^2 \times \sum_{i=1}^n \left[\frac{c_i}{(1 + k \times c_i)^2} \right] \right] \times \right. \\ \left. \times \left[M - \frac{M_0 \times \tau_{M0} + k \times \tau \times \sum_{i=1}^n \left[\frac{q_i \times c_i}{1 + k \times c_i} \right]}{\tau_{M0} + \tau \times k^2 \times \sum_{i=1}^n \left[\frac{c_i}{(1 + k \times c_i)^2} \right]} \right]^2 \right\} \quad (10)$$

que é o núcleo de uma distribuição normal com parâmetros média μ^* e variância σ^{2^*} dados respectivamente por:

$$\mu^* = \frac{M_0 \times \tau_{M0} + k \times \tau \times \sum_{i=1}^n \left[\frac{q_i \times c_i}{1+k \times c_i} \right]}{\tau_{M0} + \tau \times k^2 \times \sum_{i=1}^n \left[\frac{c_i}{(1+k \times c_i)^2} \right]}$$

$$\sigma^{2*} = \frac{\sigma_{M0}^2 \times \sigma^2}{\sigma^2 + \sigma_M^2 \times k^2 \times \sum_{i=1}^n \left[\frac{c_i}{(1+k \times c_i)^2} \right]}$$

Para o parâmetro k tem-se:

$$p(k | q_i, M, \tau) \propto \exp \left\{ -\frac{1}{2} \times \left[\tau_{k0} \times (k - k_0)^2 + \tau \times \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 \right] \right\} \quad (11)$$

a qual não possui uma forma analítica. Para o parâmetro σ tem-se:

$$p(\tau | q_i, M, k) \propto \tau^{\alpha_0 + \frac{n}{2} - 1} \exp \left\{ -\tau \times \left(\frac{1}{2} \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 + \beta_0 \right) \right\} \quad (12)$$

que, assim como a distribuição condicional para τ mostrada em 8, é o núcleo de uma distribuição Gama cuja função de densidade de probabilidade é apresentada na expressão 9, com parâmetros dados por:

$$\alpha = \alpha_0 + \frac{n}{2}$$

$$\beta = \sum_{i=1}^n \left[q_i - \frac{k \times M \times c_i}{1 + k \times c_i} \right]^2 + \beta_0$$

Obtiveram-se as mesmas distribuições condicionais para o parâmetro τ , conforme observado nas expressões 8 e 12. Isso se justifica no fato de que, para obtenção das distribuições *a posteriori* conjuntas, utilizaram-se como distribuições *a priori*, uma gama não informativa e uma gama informativa. Nesse caso essas distribuições diferem apenas nos valores dos hiperparâmetros, os quais, como mostrado na seção 2.2, foram $\alpha_0 = 0,01$ e $\beta_0 = 0,01$ para a distribuição *a priori* gama não informativa e $\alpha_0 = 4$ e $\beta_0 = 12$ para a distribuição *a priori* gama informativa.

2.4 Avaliação das estimativas

A avaliação das estimativas obtidas para os parâmetros da isoterma de Langmuir foi baseada em três medidas: para a precisão (não confundir com o parâmetro precisão τ) utilizou-se a amplitude do intervalo HPD ao nível de 95% de credibilidade e o valor do erro quadrático médio (EQM), para a acurácia baseou-se nos valores do viés médio relativo absoluto (VMR).

Conforme Paulino *et al.* (2003), o intervalo $a < \theta < b$ será chamado de intervalo de credibilidade $100(1 - \alpha)\%$ para θ se:

$$\int_{-\infty}^a \pi(\theta | x) d\theta = \int_b^{\infty} \pi(\theta | x) d\theta = \frac{\alpha}{2} \quad (13)$$

O intervalo de credibilidade engloba todos os valores de θ mais credíveis. Existe um único intervalo, conhecido como região de alta densidade a *posteriori*, abreviadamente região HPD (*Highest Posterior Density*), que apresenta a menor amplitude (maior precisão) dentre todos os possíveis intervalos.

O erro quadrático médio (EQM), segundo Casella e Berger (2010), pode ser obtido pela equação:

$$EQM = \frac{1}{ns} \sum_{i=1}^{ns} (\hat{\theta}_i - \theta)^2. \quad (14)$$

em que ns é o número final de simulações, $\hat{\theta}_i$ é o valor estimado na i -ésima simulação e θ é o valor real do parâmetro.

O viés médio relativo absoluto (VRM) utilizado neste trabalho pode ser obtido a partir da equação:

$$VMR = \frac{1}{ns} \sum_{i=1}^{ns} \left| \frac{\hat{\theta}_i - \theta}{\theta} \right| \times 100. \quad (15)$$

3 Resultados e discussões

Na Tabela 2 encontram-se os resultados da análise de convergência das cadeias a *posteriori* para ambos os cenários estudados. Nota-se que os valores de máximo absoluto das estatísticas Z para o critério de Geweke estão todos abaixo de 1,96. Os *valores-p* para o teste de Heidelberger e Welch não foram significativos ao nível de significância adotado (5%) e os fatores de dependência de Raftery e Lewis não foram maiores que 5, logo não há evidências de não convergência das cadeias a *posteriori* em ambos os cenários estudados. Assim os resultados apresentados nas Tabelas 3 a 8 e nas Figuras 1 e 2 são válidos.

Tabela 2 - Valores dos critérios de convergência para as cadeias *a posteriori* obtidas nos cenários I e II considerando as três distribuições *a priori*: Estatística Z de Geweke (valor máximo absoluto), fator de dependência de Raftery-Lewis (máximo) e *valor-p* mínimo do teste de Heidelbergger-Welch

		Dist. <i>a priori</i> uniforme			Dist. <i>a priori</i> normal I			Dist. <i>a priori</i> normal II		
		Cenário I								
P.	n	ZG	RL	pHW	ZG	RL	pHW	ZG	RL	pHW
<i>M</i>	5	1,9451	4,98	0,0503	1,9547	1,030	0,0523	1,953	1,04	0,0506
<i>k</i>	5	1,9599	4,21	0,0513	1,9590	1,040	0,0514	1,936	1,03	0,0503
σ	5	1,9338	4,98	0,0503	1,9442	1,030	0,0515	1,952	1,02	0,0530
<i>M</i>	6	1,9543	3,99	0,0501	1,9509	1,040	0,0524	1,935	1,04	0,0503
<i>k</i>	6	1,9313	3,14	0,0511	1,9347	1,040	0,0502	1,952	1,03	0,0529
σ	6	1,9493	3,99	0,0515	1,9588	1,030	0,0508	1,954	1,02	0,0506
<i>M</i>	7	1,9380	1,990	0,0515	1,9570	1,009	0,0514	1,953	1,04	0,0513
<i>k</i>	7	1,9151	1,160	0,0510	1,9550	1,040	0,0520	1,936	1,03	0,0503
σ	7	1,9598	2,980	0,0508	1,9083	1,002	0,0503	1,952	1,02	0,0512
<i>M</i>	9	1,9441	1,150	0,0502	1,9354	1,040	0,0543	1,953	1,04	0,0511
<i>k</i>	9	1,9526	1,090	0,0518	1,9363	1,040	0,0501	1,936	1,03	0,0514
σ	9	1,9303	1,030	0,0501	1,9547	1,020	0,0501	1,952	1,02	0,0519
<i>M</i>	12	1,9493	1,090	0,0516	1,9591	1,050	0,0513	1,956	1,06	0,0544
<i>k</i>	12	1,9556	1,080	0,0515	1,9573	1,060	0,0504	1,951	1,07	0,0520
σ	15	1,9506	1,020	0,0502	1,9552	1,030	0,0517	1,955	1,03	0,0513
<i>M</i>	15	1,9587	1,100	0,0502	1,9580	1,050	0,0508	1,959	1,05	0,0512
<i>k</i>	15	1,9335	1,060	0,0502	1,9583	1,060	0,0505	1,958	1,05	0,0508
σ	15	1,9584	1,020	0,0504	1,9494	1,030	0,0502	1,948	1,02	0,0503
<i>M</i>	20	1,9561	1,100	0,0504	1,9444	1,090	0,0515	1,957	1,07	0,0505
<i>k</i>	20	1,9578	1,080	0,0504	1,9508	1,080	0,0509	1,957	1,08	0,0511
σ	20	1,9592	1,020	0,0510	1,9255	1,030	0,0511	1,957	1,03	0,0507
<i>M</i>	25	1,9580	1,150	0,0509	1,9444	1,100	0,0505	1,938	1,11	0,0515
<i>k</i>	25	1,9356	1,140	0,0504	1,9206	1,110	0,0514	1,952	1,12	0,0522
σ	25	1,9596	1,030	0,0503	1,9532	1,030	0,0505	1,939	1,03	0,0505
<i>M</i>	30	1,9539	1,140	0,0503	1,9530	1,100	0,0508	1,958	1,11	0,0515
<i>k</i>	30	1,9498	1,110	0,0509	1,9430	1,110	0,0527	1,934	1,12	0,0515
σ	30	1,9536	1,030	0,0502	1,9586	1,020	0,0505	1,958	1,03	0,0510
		Cenário II								
<i>M</i>	5	1,9451	4,98	0,0503	1,9547	1,030	0,0523	1,953	1,04	0,0506
<i>k</i>	5	1,9599	4,21	0,0513	1,9590	1,040	0,0514	1,936	1,03	0,0503
σ	5	1,9338	4,98	0,0503	1,9442	1,030	0,0515	1,952	1,02	0,0530
<i>M</i>	6	1,9543	3,99	0,0501	1,9509	1,040	0,0524	1,935	1,04	0,0503
<i>k</i>	6	1,9313	3,14	0,0511	1,9347	1,040	0,0502	1,952	1,03	0,0529
σ	6	1,9493	3,99	0,0515	1,9588	1,030	0,0508	1,954	1,02	0,0506
<i>M</i>	7	1,9380	1,990	0,0515	1,9570	1,009	0,0514	1,953	1,04	0,0513
<i>k</i>	7	1,9151	1,160	0,0510	1,9550	1,040	0,0520	1,936	1,03	0,0503
σ	7	1,9598	2,980	0,0508	1,9083	1,002	0,0503	1,952	1,02	0,0512
<i>M</i>	9	1,9441	1,150	0,0502	1,9354	1,040	0,0543	1,953	1,04	0,0511
<i>k</i>	9	1,9526	1,090	0,0518	1,9363	1,040	0,0501	1,936	1,03	0,0514
σ	9	1,9303	1,030	0,0501	1,9547	1,020	0,0501	1,952	1,02	0,0519
<i>M</i>	12	1,9493	1,090	0,0516	1,9591	1,050	0,0513	1,956	1,06	0,0544
<i>k</i>	12	1,9556	1,080	0,0515	1,9573	1,060	0,0504	1,951	1,07	0,0520
σ	15	1,9506	1,020	0,0502	1,9552	1,030	0,0517	1,955	1,03	0,0513
<i>M</i>	15	1,9587	1,100	0,0502	1,9580	1,050	0,0508	1,959	1,05	0,0512
<i>k</i>	15	1,9335	1,060	0,0502	1,9583	1,060	0,0505	1,958	1,05	0,0508
σ	15	1,9584	1,020	0,0504	1,9494	1,030	0,0502	1,948	1,02	0,0503
<i>M</i>	20	1,9561	1,100	0,0504	1,9444	1,090	0,0515	1,957	1,07	0,0505
<i>k</i>	20	1,9578	1,080	0,0504	1,9508	1,080	0,0509	1,957	1,08	0,0511
σ	20	1,9592	1,020	0,0510	1,9255	1,030	0,0511	1,957	1,03	0,0507
<i>M</i>	25	1,9580	1,150	0,0509	1,9444	1,100	0,0505	1,938	1,11	0,0515
<i>k</i>	25	1,9356	1,140	0,0504	1,9206	1,110	0,0514	1,952	1,12	0,0522
σ	25	1,9596	1,030	0,0503	1,9532	1,030	0,0505	1,939	1,03	0,0505
<i>M</i>	30	1,9539	1,140	0,0503	1,9530	1,100	0,0508	1,958	1,11	0,0515
<i>k</i>	30	1,9498	1,110	0,0509	1,9430	1,110	0,0527	1,934	1,12	0,0515
σ	30	1,9536	1,030	0,0502	1,9586	1,020	0,0505	1,958	1,03	0,0510

P.: Parâmetro; ZG: Estatística Z de Geweke; RL: Fator de dependência de Raftery-Lewis; pHW: *valor-p* do teste de

Para o cenário I, todas as estimativas dos parâmetros M , k e σ foram significativas, uma vez que os respectivos intervalos HPD ao nível de 95% de credibilidade não continham o valor zero em sua extensão (Tabelas 3 a 5).

Na Tabela 3 são apresentadas as estatísticas descritivas associadas ao parâmetro M obtidas a partir das diferentes distribuições *a priori* propostas sobre os diferentes tamanhos amostrais.

Tabela 3 - Número de cadeias convergentes para cada tamanho amostral, médias *a posteriori* do parâmetro M e respectivos desvios-padrão (DP) e intervalo HPD ao nível de 95% de credibilidade obtidos sob diferentes distribuições *a priori* e tamanhos amostrais (n) no cenário I

Dist. <i>a priori</i>	n	N ^o cadeias	\hat{M}	DP	LI	LS	Amp.
Uniforme	5	282	100,07	1,865	96,65	103,42	6,772
Uniforme	6	665	99,93	20,438	95,91	103,85	7,938
Uniforme	7	814	100,04	19,128	96,22	103,78	7,560
Uniforme	9	975	99,96	1,699	96,56	103,3	6,741
Uniforme	12	1036	99,98	1,539	96,91	102,99	6,085
Uniforme	15	1048	100,01	14,975	97,03	102,94	5,902
Uniforme	20	1058	99,95	1,323	97,33	102,52	5,195
Uniforme	25	953	100,02	13,088	97,43	102,56	5,125
Uniforme	30	1062	99,96	0,961	98,07	101,81	3,746
Normal I	5	1007	99,83	1,718	96,4	103,18	6,772
Normal I	6	1052	99,74	1,666	96,42	102,98	6,562
Normal I	7	1033	99,85	1,534	96,8	102,84	6,016
Normal I	9	1032	99,83	1,376	97,09	102,5	5,404
Normal I	12	1035	99,88	1,323	97,26	102,44	5,223
Normal I	15	1044	99,79	1,333	97,15	102,37	5,171
Normal I	20	1030	99,86	1,224	97,44	102,23	4,781
Normal I	25	1073	99,88	1,213	97,49	102,23	4,733
Normal I	30	1048	99,92	0,905	98,13	101,65	3,517
Normal II	5	999	99,947	1,795	96,373	103,461	7,088
Normal II	6	1011	100,071	1,737	96,612	103,465	6,854
Normal II	7	1029	99,978	1,592	96,818	103,087	6,27
Normal II	9	1035	99,95	1,406	97,156	102,683	5,527
Normal II	12	1052	99,984	1,361	97,293	102,632	5,339
Normal II	15	1060	99,948	1,358	97,259	102,579	5,32
Normal II	20	1057	99,941	1,246	97,48	102,353	4,873
Normal II	25	1069	99,873	1,227	97,455	102,243	4,788
Normal II	30	1031	99,984	0,915	98,182	101,74	3,558

Dist.: Distribuição; DP: Desvio-padrão; LI = Limite inferior; LS = limite superior; Amp.: Amplitude do intervalo HPD ao nível de 95% de credibilidade.

Na Tabela 4 são apresentadas as estatísticas descritivas associadas ao parâmetro k e na Tabela 5 as estatísticas associadas ao parâmetro σ .

Tabela 4 - Número de cadeias convergentes para cada tamanho amostral, médias a *posteriori* do parâmetro k e respectivos desvios-padrão (DP) e intervalo HPD ao nível de 95% de credibilidade obtidos sob diferentes distribuições a *priori* e tamanhos amostrais (n) no cenário I

Dist. a <i>priori</i>	n	Nº cadeias	\hat{k}	DP	LI	LS	Amp.
Uniforme	5	282	0,032	0,005	0,028	0,037	0,009
Uniforme	6	665	0,032	0,003	0,027	0,038	0,01
Uniforme	7	814	0,032	0,003	0,027	0,038	0,011
Uniforme	9	975	0,032	0,002	0,028	0,037	0,009
Uniforme	12	1036	0,032	0,002	0,028	0,036	0,007
Uniforme	15	1048	0,032	0,001	0,029	0,035	0,006
Uniforme	20	1058	0,032	0,001	0,029	0,035	0,005
Uniforme	25	953	0,032	0,001	0,03	0,034	0,005
Uniforme	30	1062	0,032	0,001	0,03	0,034	0,005
Normal I	5	1007	0,032	0,002	0,028	0,037	0,009
Normal I	6	1052	0,032	0,002	0,028	0,036	0,008
Normal I	7	1033	0,032	0,002	0,028	0,036	0,008
Normal I	9	1032	0,032	0,002	0,029	0,036	0,007
Normal I	12	1035	0,032	0,002	0,029	0,035	0,006
Normal I	15	1044	0,032	0,001	0,029	0,035	0,005
Normal I	20	1030	0,032	0,001	0,03	0,035	0,005
Normal I	25	1073	0,032	0,001	0,03	0,034	0,004
Normal I	30	1048	0,032	0,001	0,03	0,034	0,004
Normal II	5	999	0,032	0,002	0,027	0,037	0,01
Normal II	6	1011	0,032	0,002	0,028	0,036	0,009
Normal II	7	1029	0,032	0,002	0,028	0,037	0,009
Normal II	9	1035	0,032	0,002	0,029	0,036	0,007
Normal II	12	1052	0,032	0,002	0,029	0,035	0,006
Normal II	15	1060	0,032	0,001	0,029	0,035	0,006
Normal II	20	1057	0,032	0,001	0,03	0,035	0,005
Normal II	25	1069	0,032	0,001	0,03	0,034	0,005
Normal II	30	1031	0,032	0,001	0,03	0,034	0,004

Dist.: Distribuição; DP: Desvio-padrão; LI = Limite inferior; LS = limite superior; Amp.: Amplitude do intervalo HPD ao nível de 95% de credibilidade.

Conforme se observa nas Tabelas 3, 4 já apresentadas e na Tabela 5 a seguir, houve um aumento no número de cadeias que atingiram a convergência, conforme aumento do tamanho amostral e devido à utilização de distribuições a *priori* informativas, o que é evidenciado quando se comparam as distribuições a

posteriori obtidas com a distribuição *a priori* uniforme ou com as distribuições *a priori* normais, considerando o mesmo tamanho amostral.

Tabela 5 - Número de cadeias convergentes para cada tamanho amostral, médias *a posteriori* do parâmetro σ e respectivos desvios-padrão (DP) e intervalo HPD ao nível de 95% de credibilidade, obtidas a partir de uma distribuição *a priori* gama não informativa (G NI) e uma informativa (G I), sob os diferentes tamanhos amostrais (n) no cenário I

Dist. <i>a priori</i>	n	Nº cadeias	$\hat{\sigma}$	DP	LI	LS	Amp.
G NI	5	282	1,112	1,230	0,689	37,308	37,753
G NI	6	665	1,575	1,872	1,020	11,703	10,679
G NI	7	814	1,770	2,225	1,183	7,352	6,169
G NI	9	975	1,969	2,695	1,375	5,243	3,868
G NI	12	1.036	1,962	2,935	1,429	3,991	2,562
G NI	15	1.048	1,980	3,163	1,485	3,523	2,038
G NI	20	1.058	1,963	3,409	1,524	3,086	1,562
G NI	25	953	1,998	3,679	1,586	2,936	1,350
G NI	30	1.062	1,995	3,860	1,612	2,800	1,187
G I	5	1.007	1,816	2,909	1,363	3,211	2,091
G I	6	1.052	1,804	2,833	1,343	3,305	1,962
G I	7	1.033	1,816	2,909	1,363	3,211	1,841
G I	9	1.032	1,839	3,049	1,401	3,077	1,676
G I	12	1.035	1,853	3,213	1,438	2,916	1,489
G I	15	1.044	1,893	3,411	1,491	2,848	1,345
G I	20	1.030	1,905	3,619	1,529	2,716	1,186
G I	25	1.073	1,928	3,827	1,571	2,645	1,074
G I	30	1.048	1,933	3,982	1,595	2,579	0,984

Dist.: Distribuição; DP: Desvio-padrão; LI = Limite inferior; LS = limite superior; Amp.: Amplitude do intervalo HPD ao nível de 95% de credibilidade.

A partir da Tabela 3, nota-se que as estimativas de M apresentaram valores bem próximos, diferindo nos valores do desvio-padrão associado, os quais oscilaram muito conforme o tamanho amostral para as estimativas obtidas com uso da distribuição *a priori* uniforme, de modo que não foi possível afirmar que exista uma relação. Por outro lado, para as estimativas obtidas com as normais I e II, nota-se um decréscimo gradativo do desvio-padrão conforme o tamanho amostral. Da Tabela 4, observa-se que o parâmetro k apresentou as mesmas estimativas, independente do tamanho amostral, apresentando uma pequena diferença quanto aos valores do desvio-padrão das mesmas, de modo que estes valores foram menores na estimação com uso das distribuições *a priori* informativas.

O parâmetro σ (Tabela 5), apresentou maior sensibilidade quanto à estimação com uso de diferentes distribuições *a priori* gama e em relação ao aumento do

tamanho amostral. Percebe-se que com a gama não informativa os valores das estimativas e do desvio-padrão de σ aumentaram conforme o tamanho amostral, entretanto, como será discutido adiante (Figura 1(c)), há um alto VMR associado às estimativas sob tamanhos amostrais menores.

Na Figura 1 são apresentados os gráficos relativos ao viés médio relativo absoluto (VMR) e ao erro quadrático médio (EQM) associados aos parâmetros M , k e σ obtidos sob diferentes tamanhos de amostra no cenário I.

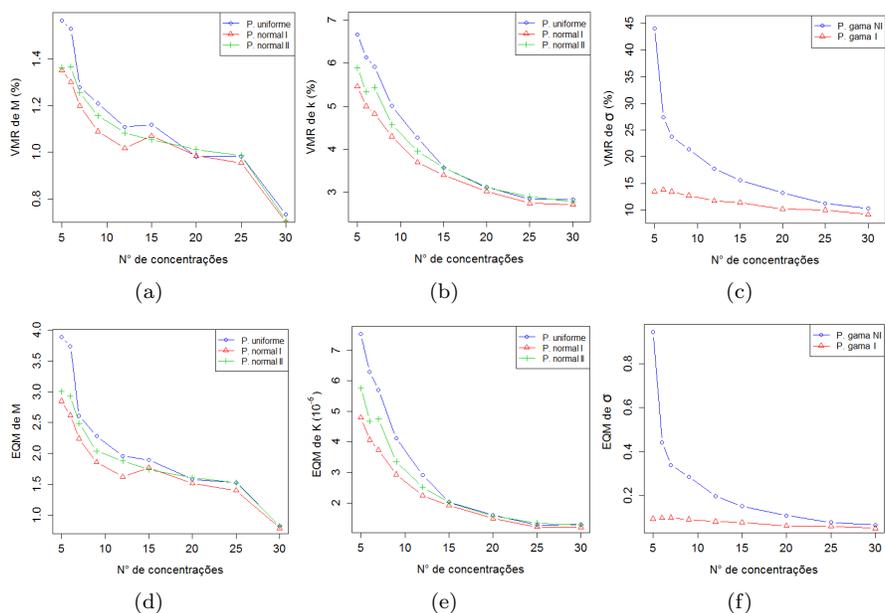


Figura 1 - Viés médio relativo absoluto e erro quadrático médio das estimativas dos parâmetros M , k e σ para diferentes tamanhos amostrais no Cenário I, considerando distribuições *a priori* informativas e também não informativas.

A partir dos resultados apresentados nas Tabelas 3 e 4 e na Figura 1(d), nota-se um aumento na precisão das estimativas de M , evidenciado pela redução tanto da amplitude do intervalo HPD ao nível de 95% de credibilidade, como dos valores de EQM, com o uso das distribuições *a priori* normais e conforme o tamanho amostral. O aumento da acurácia de M ficou evidenciado pela redução do VMR (Figura 1(a)) conforme o tamanho amostral, apresentando valores de 1,53% para $n = 5$ e 0,73% para $n = 30$ com a distribuição *a priori* uniforme e de 1,35% a 0,70% com a normal I nesses mesmos tamanhos amostrais.

Esse aumento de precisão e acurácia com o uso da distribuição *a priori* informativa normal (I ou II), do ponto de vista prático, representa uma importante vantagem, por exemplo, a partir da Figura 1(a) pode-se observar que para o cenário

estudado, ao utilizar-se a *priori* normal I, obteve-se o mesmo VMR de estimativa (1,20%) no tamanho amostral $n = 7$ comparado a $n = 9$ uso da *priori* uniforme, situação parecida observa-se para o VMR de M em $n = 9$ na informativa (1,09%), que é menor que o VMR de $n = 15$ na uniforme (1,12%). As estimativas obtidas com uso da *priori* normal I apresentaram precisão e acurácia um pouco maiores que aquelas obtidas com a normal II (1). Resultados semelhantes foram observados para as estimativas do parâmetro k , com o aumento da precisão e da acurácia com uso da *priori* informativa e conforme aumento do tamanho amostral.

Para as estimativas do parâmetro σ , notam-se elevados valores da amplitude do intervalo HPD 95% de credibilidade (Tabela 5) e do EQM (Figura 1(f)), o que se traduz em uma baixa precisão. Situação semelhante se observa pelos altos valores do VMR (Figura 1(c)), mostrando a baixa acurácia das estimativas. Por outro lado quando na estimação foi considerado a *priori* gama informativa e o aumento do tamanho amostral, observa-se um aumento nítido na precisão e na acurácia das estimativas obtidas. É importante Ressaltar que com a gama informativa, o tamanho amostral pouco influenciou na precisão e acurácia das estimativas de σ , visto que os valores de amplitude do intervalo HPD 95%, EQM e VMR pouco variaram.

Embora o tipo de dados a que estes modelos são aplicados não possuem nenhuma relação com dados de adsorção, o uso da metodologia Bayesiana aplicados a alguns tipos de testes, bem como outros modelos de regressão não linear apresentam resultados semelhantes aos obtidos com este estudo. Quanto à amplitude do intervalo HPD 95% de credibilidade, Neves *et al.* (2015) reduziram consideravelmente a amplitude do intervalo HPD da estimativa do parâmetro com inclusão da *priori* informativa, em um estudo de análise Bayesiana para avaliar testes de detecção de *Otodectes cynotis* em cães.

Blasco *et al.* (2003) em estudos sobre o efeito da seleção sobre a taxa de crescimento de coelhos, utilizando para isso um modelo de Gompertz sob enfoque Bayesiano, com o uso de uma distribuição *a priori* normal, obtiveram estimativas mais precisas, uma redução do número de estimativas viesadas e uma menor amplitude do intervalo de credibilidade, mesmo utilizando uma quantidade menor de informações.

A seguir, nas Tabelas 6, 7 e 8, são apresentadas as estatísticas descritivas relacionadas aos parâmetros M , k e σ obtidas no cenário II ($k = 0,025 \text{ Lmg}^{-1}$ e $M = 72 \text{ mgg}^{-1}$).

Conforme apresentado nas Tabelas 6, 7 e 8, todas as estimativas dos parâmetros M , k e σ foram significativas, uma vez que os respectivos intervalos HPD ao nível de 95% de credibilidade não continham o valor zero em sua extensão.

Percebe-se pelas referidas tabelas que houve um aumento no número de cadeias que atingiram a convergência, conforme aumento do tamanho amostral e devido à utilização de distribuições *a priori* informativas, o que é evidenciado quando se comparam as distribuições *a posteriori* obtidas com a distribuição *a priori* uniforme e com as distribuições *a priori* normais para os parâmetros M e k , bem como com as distribuições *a priori* gama informativa e não informativa para o parâmetro σ ,

considerando o mesmo tamanho amostral.

Tabela 6 - Número de cadeias convergentes para cada tamanho amostral, médias *a posteriori* do parâmetro M e respectivos desvios-padrão (DP) e intervalo HPD ao nível de 95% de credibilidade obtidos sob diferentes distribuições *a priori* e tamanhos amostrais (n) no cenário II

Dist. <i>a priori</i>	n	N cadeias	\hat{M}	DP	LI	LS	Amp.
Uniforme	5	182	72,07	1,660	69,01	75,07	6,054
Uniforme	6	555	72,00	2,059	67,96	75,95	7,986
Uniforme	7	614	72,00	1,892	68,22	75,73	7,512
Uniforme	9	1017	71,9	1,815	68,26	75,5	7,238
Uniforme	12	1062	71,88	1,674	68,52	75,19	6,663
Uniforme	15	1062	71,98	1,638	68,72	75,22	6,502
Uniforme	20	1087	71,98	1,478	69,04	74,88	5,845
Uniforme	25	1036	71,94	1,436	69,08	74,75	5,670
Uniforme	30	1091	71,97	1,057	69,87	74,04	4,166
Normal I	5	1035	73,24	2,079	69,26	77,4	8,143
Normal I	6	1022	73,15	1,966	69,36	77,08	7,725
Normal I	7	1038	73,01	1,814	69,51	76,63	7,125
Normal I	9	1033	72,77	1,589	69,67	75,93	6,256
Normal I	12	1041	72,75	1,535	69,74	75,79	6,043
Normal I	15	1041	72,78	1,522	69,8	75,79	5,992
Normal I	20	1067	72,62	1,406	69,86	75,39	5,532
Normal I	25	1061	72,61	1,366	69,93	75,3	5,373
Normal I	30	1072	72,31	1,018	70,3	74,3	4,001
Normal II	5	1013	72,44	1,937	68,63	76,29	7,664
Normal II	6	1001	72,44	1,919	68,66	76,25	7,594
Normal II	7	1017	72,31	1,618	69,11	75,52	6,405
Normal II	9	1031	72,28	1,582	69,16	75,41	6,255
Normal II	12	1037	72,28	1,524	69,27	75,29	6,017
Normal II	15	1059	72,20	1,515	69,20	75,18	5,976
Normal II	20	1061	72,17	1,409	69,39	74,94	5,553
Normal II	25	1081	72,22	1,369	69,52	74,91	5,388
Normal II	30	1056	72,11	1,015	70,10	74,09	3,994

Dist.: Distribuição; DP: Desvio-padrão; LI = Limite inferior; LS = limite superior; Amp.: Amplitude do intervalo HPD ao nível de 95% de credibilidade.

Nota-se que as estimativas de M apresentaram valores próximos, diferindo pouco nas medidas do desvio-padrão associado. Quanto à precisão, o uso das distribuições *a priori* uniforme e normal I levaram a obtenção amplitudes de intervalo HPD de 95% de credibilidade bem próximas, ao passo que essa amplitude com a normal II apresentou um menor valor. Observa-se um inesperado valor

menor de amplitude (6,054) HPD 95% de credibilidade em $n = 5$ (Tabela 6), acredita-se que isso ocorreu em decorrência do baixo número de cadeias *a posteriori* que não violaram os critérios de convergência (Tabela 2), de modo que este valor de amplitude não condiz com a realidade.

As informações referentes às estimativas do parâmetro k são apresentadas a seguir na Tabela 7.

Tabela 7 - Número de cadeias convergentes para cada tamanho amostral, médias *a posteriori* do parâmetro k e respectivos desvios-padrão e intervalo HPD ao nível de 95% de credibilidade obtidos sob diferentes distribuições *a priori* e tamanhos amostrais (n) no cenário II

Dist. <i>a priori</i>	n	N cadeias	\hat{k}	DP	LI	LS	Amp.
Uniforme	5	182	0,025	0,005	0,021	0,030	0,009
Uniforme	6	555	0,026	0,005	0,02	0,031	0,010
Uniforme	7	614	0,025	0,003	0,02	0,031	0,010
Uniforme	9	1017	0,025	0,002	0,021	0,030	0,010
Uniforme	12	1062	0,025	0,002	0,021	0,029	0,008
Uniforme	15	1062	0,025	0,002	0,022	0,029	0,007
Uniforme	20	1087	0,025	0,001	0,022	0,028	0,006
Uniforme	25	1036	0,025	0,001	0,022	0,028	0,005
Uniforme	30	1091	0,025	0,001	0,023	0,028	0,005
Normal I	5	1035	0,024	0,003	0,019	0,029	0,010
Normal I	6	1022	0,024	0,002	0,020	0,029	0,009
Normal I	7	1038	0,024	0,002	0,020	0,029	0,009
Normal I	9	1033	0,024	0,002	0,021	0,028	0,008
Normal I	12	1041	0,024	0,002	0,021	0,028	0,007
Normal I	15	1041	0,024	0,001	0,021	0,028	0,006
Normal I	20	1067	0,025	0,001	0,022	0,027	0,005
Normal I	25	1061	0,025	0,001	0,022	0,027	0,005
Normal I	30	1072	0,025	0,001	0,022	0,027	0,005
Normal II	5	1013	0,025	0,003	0,020	0,030	0,011
Normal II	6	1001	0,025	0,002	0,020	0,030	0,009
Normal II	7	1017	0,025	0,002	0,020	0,030	0,009
Normal II	9	1031	0,025	0,002	0,021	0,029	0,008
Normal II	12	1037	0,025	0,002	0,021	0,028	0,007
Normal II	15	1059	0,025	0,002	0,022	0,028	0,006
Normal II	20	1061	0,025	0,001	0,022	0,028	0,006
Normal II	25	1081	0,025	0,001	0,022	0,027	0,005
Normal II	30	1056	0,025	0,001	0,023	0,027	0,005

Dist.: Distribuição; DP: Desvio-padrão; LI = Limite inferior; LS = limite superior; Amp.: Amplitude do intervalo HPD ao nível de 95% de credibilidade.

Da Tabela 7, observa-se que o parâmetro k apresentou as mesmas estimativas, independente do tamanho amostral com a distribuição *a priori* normal II. Ficou evidente que, quando se comparam os valores da amplitude do intervalo HPD ao nível de 95% de credibilidade, ambas as distribuições *a priori* normais e uniforme apresentaram praticamente os mesmos resultados em termos de precisão.

As informações associadas às estimativas do parâmetro σ são apresentadas a seguir na Tabela 8.

Tabela 8 - Número de cadeias convergentes para cada tamanho amostral, médias *a posteriori* do parâmetro σ e respectivos desvios-padrão e intervalo HPD ao nível de 95% de credibilidade, obtidas a partir de uma distribuição *a priori* gama não informativa (G NI) e uma informativa (G I) sob os diferentes tamanhos amostrais (n) no cenário II

<i>Priori</i>	n	N cadeias	$\hat{\sigma}$	Dp	LI	LS	Amp.
G NI	5	182	0,902	0,998	0,559	31,662	31,103
G NI	6	555	1,438	1,708	0,931	11,859	10,928
G NI	7	614	1,583	1,989	1,058	6,642	5,584
G NI	9	1017	1,909	2,610	1,332	5,078	3,746
G NI	12	1062	1,930	2,886	1,406	3,929	2,524
G NI	15	1062	1,961	3,132	1,470	3,490	2,020
G NI	20	1087	1,960	3,403	1,521	3,079	1,558
G NI	25	1036	1,978	3,643	1,570	2,907	1,337
G NI	30	1091	1,977	3,825	1,597	2,773	1,176
G I	5	1035	1,820	2,765	1,335	3,586	2,250
G I	6	1022	1,837	2,856	1,362	3,448	2,087
G I	7	1038	1,837	2,918	1,374	3,306	1,932
G I	9	1033	1,856	3,058	1,410	3,141	1,730
G I	12	1041	1,882	3,248	1,457	2,983	1,525
G I	15	1041	1,902	3,413	1,495	2,876	1,380
G I	20	1067	1,916	3,629	1,536	2,739	1,203
G I	25	1061	1,922	3,806	1,565	2,643	1,078
G I	30	1072	1,936	3,985	1,598	2,587	0,989

Dist.: Distribuição; DP: Desvio-padrão; LI = Limite inferior; LS = limite superior; Amp.: Amplitude do intervalo HPD ao nível de 95% de credibilidade.

As discussões referentes aos resultados obtidos para o parâmetro σ apresentados na Tabela 8 são análogos aos já discutidos no cenário I.

Na Figura 2 são apresentados os gráficos relativos ao viés médio relativo absoluto (VMR) e ao erro quadrático médio (EQM) associados aos parâmetros M , k e σ obtidos sob diferentes tamanhos de amostra no cenário II.

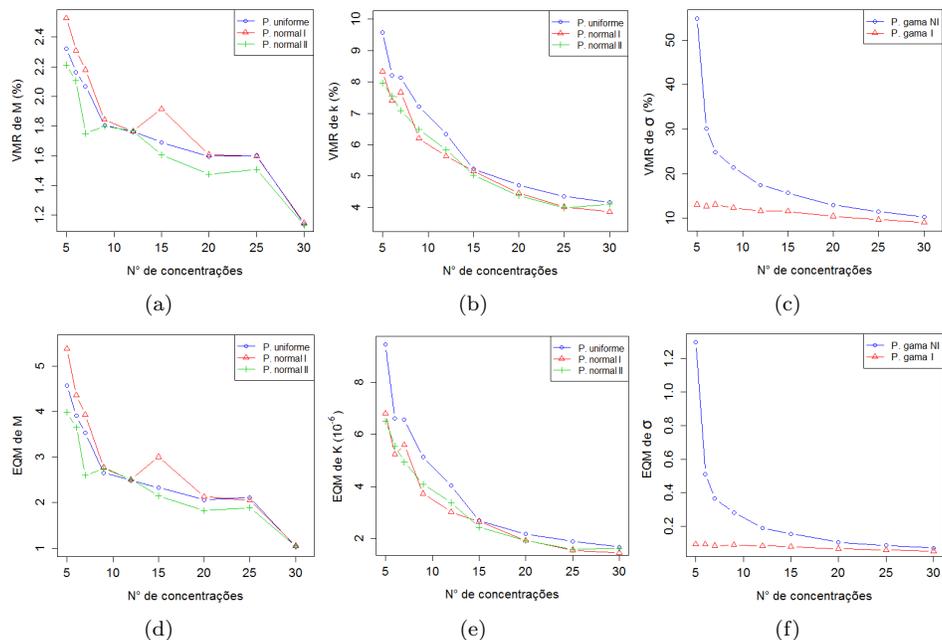


Figura 2 - Viés médio relativo absoluto (VMR) e erro quadrático médio (EQM) das estimativas dos parâmetros M , k e σ para diferentes tamanhos amostrais (nº de concentrações) no cenário II, considerando distribuições *a priori* informativas e também não informativas.

O uso da distribuição *a priori* normal I ocasionou um aumento do VRM (Figura 2(a)), do EQM (Figura 2(d)) e da amplitude HDP (Tabelas 6 das estimativas de M). A explicação mais plausível para esse aumento do VMR e do EQM, é de que, conforme observado nas curvas de densidade das distribuições *a priori* normais I e II (Figura 3), o valor $M = 72 \text{ mgg}^{-1}$, está distante da região mais densa para a distribuição normal I, desse modo possui menor probabilidade de ocorrência.

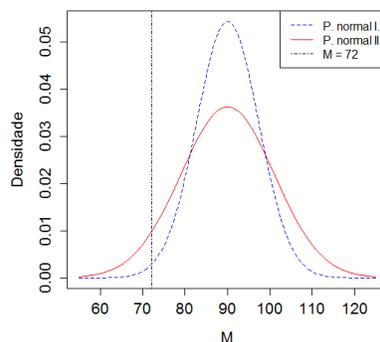


Figura 3 - Distribuições *a priori* para o parâmetro M , com representação do valor $M = 72 \text{ mgg}^{-1}$.

As discussões referentes às vantagens pela utilização da distribuição *a priori* normal II como forma de obterem-se estimativas com menores valores de VMR, EQM e amplitude do intervalo HPD 95% sob menores tamanhos amostrais, bem como os resultados referentes aos parâmetros k e σ , apresentados na Figura 2, são análogas às realizadas no cenário I.

A distribuição *a priori* normal II, a qual confere maior variabilidade aos parâmetros, mostrou-se mais aplicável à modelagem Bayesiana da isoterma de Langmuir que a uniforme, uma vez que possibilitou aumentos de acurácia e de precisão conforme o tamanho amostral em ambos os cenários estudados. De modo geral, os resultados obtidos para ambos os cenários corroboram com Gianola e Fernando (1986) no sentido de que a inferência Bayesiana proporcionou uma solução precisa para o conjunto de dados independente do seu tamanho, e com Paulino, Turkman e Murteira (2003) em que o uso de informação *a priori* possibilitou a obtenção de intervalos de credibilidade mais estreitos para os parâmetros estimados.

Em um estudo de melhoramento genético de gado nelore, utilizando uma abordagem Bayesiana de modelos auto-regressivos e considerando diferentes distribuições *a priori* informativas e não informativas no ajuste, Silva *et al.* (2008), ao utilizarem distribuições *a priori* informativas, obtiveram estimativas com precisão e acurácia expressivamente superiores em comparação às obtidas com distribuições *a priori* não informativas.

Este trabalho soma-se à um grande número de trabalhos, além dos supracitados, como Martins Filho *et al.* (2008), Silva *et al.* (2005), Nogueira *et al.* (2003) e Silva *et al.* (2007) e Andrade Filho *et al.* (2010) que mostram que a metodologia Bayesiana é eficiente na estimação de parâmetros, em que diferentes abordagens de distribuições *a priori* podem ser adotadas, variando conforme o tipo de modelo e tipo de informação que se tem disponível a respeito da variabilidade dos parâmetros dos respectivos modelos.

3.1 Estudo de caso

Os resultados da análise de convergência das cadeias *a posteriori* do ajuste Bayesiano da isoterma de Langmuir com os dados experimentais são apresentados na Tabela 9.

Tabela 9 - Critérios de convergência para as cadeias *a posteriori* obtidas com *prioris* uniforme e Normal II: Estatística Z de Geweke, fator de dependência de Raftery - Lewis (R-L) e *valor-p* do teste de Heidelberg - Welch (H-W)

Parâmetro	Dist. <i>a priori</i> normal II			Dist. <i>a priori</i> uniforme		
	Geweke	R-L	H-W	Geweke	R-L	H-W
M	1,129	1,00	0,5656	0,63	3,31	0,7814
k	-0,684	1,03	0,9573	0,79	2,15	0,3246
σ	-0,564	1,01	0,7367	0,04	3,29	0,5848

Dist.: Distribuição; Geweke: Estatística Z de Geweke; R-L: Fator de dependência de Raftery - Lewis; H-W: *valor-p* do teste de Heidelberg - Welch.

A partir da Tabela 9 conclui-se que não há evidências de não convergência das cadeias *a posteriori*, logo os resultados apresentados na Tabela 10 são válidos.

Tabela 10 - Estimativas dos parâmetros da isoterma de Langmuir com respectivos desvios-padrão, intervalos de credibilidade ao nível de 95% e sua amplitude obtidas a partir das distribuições *a priori* normal II e uniforme

Distribuição <i>a priori</i>	Par.: Parâmetros	Estimativa	DP	LI	LS	Amp.
Normal II	M	56,66	2,782	51,43	62,2	10,77
	k	0,0177	0,0031	0,0117	0,0239	0,0123
	σ	2,177	3,456	1,627	3,881	2,254
Uniforme	M	54,73	5,16	44,70	65,25	20,55
	k	0,0220	0,0165	0,0082	0,0391	0,0309
	σ	3,38	4,190	2,256	20,307	18,051

DP: Desvio padrão; LI: Limite inferior; LS: limite superior; Amp.: Amplitude do intervalo HPD ao nível de 95% de credibilidade

A partir da Tabela 10, observa-se que as estimativas de M , k e σ foram significativas, pois seus respectivos intervalos HPD ao nível de 95% de credibilidade não continham o valor zero em sua extensão. A estimação com uso da distribuição *a priori* normal que considera $M_0 \sim N(90; 11, 02)$ e $k_0 \sim N(0, 025; 0, 011)$ possibilitou a obtenção de estimativas mais precisas para M e k do que aquelas obtidas com a

uniforme, uma vez que apresentaram menores valores de desvio-padrão e amplitudes do intervalo HPD 95%. O alto valor da amplitude do intervalo HPD 95%, com uso da uniforme foi decorrente da dificuldade de obterem-se cadeias *a posteriori* que não violassem os critérios de convergência, mediante ao pequeno tamanho amostral do experimento ($n = 7$), prova disso são os valores relativamente altos dos fatores de dependência de Raftery - Lewis conforme Tabela 9.

Baseado na estimativa de M obtida com a distribuição *a priori* normal, a capacidade máxima de adsorção de chumbo (Pb^{2+}) por cascas de laranja é em média $56,66 \text{ mgg}^{-1}$, acredita-se com 95% de credibilidade, que este valor esteja entre 51,43 e $62,2 \text{ mgg}^{-1}$. Baseado na estimativa do parâmetro k , a energia de adsorção do chumbo (Pb^{2+}) nas cascas de laranja é em média $0,018 \text{ Lmg}^{-1}$, acredita-se com 95% de credibilidade, que este valor esteja entre 0,012 e $0,024 \text{ Lmg}^{-1}$.

As estimativas aqui obtidas para os parâmetros M e k são próximas das obtidas por Souza et al., (2012), que por meio da linearização da isoterma de Langmuir estimaram $M = 55,52 \text{ mgg}^{-1}$ e $k = 0,018 \text{ Lmg}^{-1}$.

Conclusões

As distribuições *a posteriori* foram obtidas e apresentaram resultados satisfatórios na estimação dos parâmetros da isoterma de Langmuir. Com isso a metodologia Bayesiana apresentou um elevado potencial para modelagem do processo de adsorção. O uso da distribuição *a priori* informativa normal com maior variabilidade proporcionou a obtenção de estimativas mais precisas e acuradas que aquelas obtidas com a distribuição *a priori* uniforme. Em algumas situações, com o uso da distribuição *a priori* informativa normal para estimação dos parâmetros foi possível obterem-se estimativas com maiores precisão e acurácia em tamanhos amostrais inferiores que as estimativas obtidas com a uniforme sob tamanhos amostrais superiores.

Ficou evidente que o uso da inferência Bayesiana, com distribuição *a priori* informativa na estimação dos parâmetros da isoterma de Langmuir, possibilitou a obtenção de estimativas com maior precisão e acurácia, mostrando-se uma alternativa viável e de grande potencial para contornar a limitação prática acerca do tamanho de amostra, de modo que seja possível obterem-se estimativas precisas utilizando-se a menor quantidade de amostra possível.

O ajuste Bayesiano da isoterma de Langmuir sobre dados experimentais de adsorção de chumbo em cascas de laranja evidenciaram o potencial da metodologia Bayesiana na modelagem do processo de adsorção, de modo que o uso da distribuição *a priori* normal informativa possibilitou um aumento expressivo da precisão das estimativas dos parâmetros.

Para prosseguimento deste trabalho sugere-se o emprego e comparação de outros tipos de distribuições *a priori* não informativas, tais como a de Jeffreys (1991) com e sem a suposição de independência entre os parâmetros M e k , densidades *a priori* normais com variância grande ou ainda uma distribuição *a priori flat* (uniforme imprópria). Outra sugestão é a abordagem Bayesiana, semelhante ao

que foi feito neste trabalho, de outras isotermas de adsorção, tais como a isoterma de adsorção de Freundlich.

4 Agradecimentos

À FAPEMIG pelo suporte financeiro (Processo N^o: APQ-02225-13), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de pesquisa concedida. Os autores gostariam de agradecer aos dois revisores pelos comentários que melhoraram o texto em vários aspectos.

CARVALHO, D. T.; BELJO, L. A.; MUNIZ, J. A. A Bayesian approach for modeling the Langmuir isotherm. *Rev. Bras. Biom.*, Lavras, v.35, n.2, p.376-401, 2017.

■ **ABSTRACT:** *The aim of this study was to utilize the Bayesian method for modeling the Langmuir isotherm considering informative and non-informative prior distributions. It was conducted a data simulation study considering different sample sizes to evaluate the precision and accuracy of the estimates of affinity parameter (k) and maximum adsorption capacity (M), where they were obtained with different normal informative priors distribution and not informative uniform distribution, together with the estimates of the parameter τ for which were proposed a Gama informative and uninformative prior distributions. The samples of the marginal posterior distributions of isotherm's parameters were obtained by Gibbs sampler. The inferences were made and the results indicated that the Bayesian method is efficient and the estimates obtained with use of informative prior distributions of the parameters had higher precision and accuracy in the same lower sample sizes. The Langmuir isotherm was modeled with experimental adsorption data considering prior distributions proposals and the results corroborate the simulation study so that the estimates obtained with the informative priors showed higher precision.*

■ **KEYWORDS:** *Adsorption; information a priori; nonlinear regression; sample size.*

Referências

ANDRADE FILHO, M. G. *et al.* Uma abordagem Bayesiana para curvas de crescimento com diferentes densidades a priori. *Revista Brasileira de Biometria*, São Paulo, v.28, n.2, p.161-181, 2010.

ATKINS, P.; JONES, L. *Princípios de Química: Questionando a vida moderna e o meio ambiente*. 3ed. Porto Alegre: Brookman, 2011. 1048p.

BARBOSA, C. S. *et al.* Remoção de compostos fenólicos de soluções aquosas utilizando carvão ativado preparado a partir do aguapé (*Eichhornia crassipes*): estudo cinético e de equilíbrio termodinâmico. *Química Nova*, v.37, n.3, p.447-453, 2014.

- BLASCO, A. *et al.* Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. *Genetics Selection Evolution*, v.35, p.21-41, 2003.
- BOX, G. E. P.; TIAO, G. *Bayesian inference in statistical analysis*. New York: Wiley, 1992. 608p.
- CASELLA, G.; BERGER, R. L. *Inferência estatística*. 2ed. São Paulo: Cengage Learning, 2010, 588p.
- DEGROOT, M. H.; SCHERVISH, M. J. *Probability and statistics*. Boston: Pearson Education. 4ed. 2012. 911p.
- FENG, N. *et al.* Biosorption of heavy metals from aqueous solutions by chemically modified orange peel. *Journal of Hazardous Materials*, v.185, p.49-54, 2011.
- FENG, N.; GUO, X. Characterization of adsorptive capacity and mechanisms on adsorption of copper, lead and zinc by modified orange peel. *Transactions of Nonferrous Metals Society of China*. v.22, p.1224-1231, 2012.
- GAMERMAN, D. *Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference*. Texts in Statistical Sciences. London: Chapman and Hall/CRC, 1997, 291p.
- GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (with discussion) In *Bayesian Statistics*. 4ed. Oxford: Oxford University Press, p.169-193, 1992.
- GIANOLA, D.; FERNANDO, R. L. Bayesian methods in animal breeding theories. *Journal Animal Science*, v.63, p.217-244, 1986.
- HASTINGS, W. K. Monte Carlo sampling methods using Markov Chain and their applications. *Biometrika*, v.57, p.97-109, 1970.
- HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. *Operations Research*, v.31, p.97-109, 1983.
- JEFFREYS, H. *Theory of probability*. Oxford: Clarendon, 1961.
- LANGMUIR I. The adsorption of gases on plane surfaces of glass, mica and platinum. *Journal of American Chemistry Society*, v.40, p.1361-1403, 1918.
- LIANG, S. *et al.* Application of orange peel xanthate for the adsorption of Pb^{2+} from aqueous solutions. *Journal of Hazardous Materials*, v.170, p.425-429, 2009.
- MARTINS FILHO, S. *et al.* Abordagem Bayesiana das curvas de crescimento de duas cultivares de feijoeiro. *Ciência Rural*, Santa Maria, v.38, n.6, p.1516-1521, 2008.
- MASSOCATTO, C. L. *et al.* Utilização da casca de laranja como adsorvente de chumbo. In: 33º REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE QUÍMICA. *Anais...* Águas de Lindóia, 2010.
- MAZUCHELI, J.; ACHCAR, J. A. Algumas considerações em regressão não linear. *Acta Scientiarum*. Maringá, v.24, n.6, p.1761-1770, 2002.

- MOOD, A. M.; GRAYBILL, F. A.; BOES, C. D. *Introduction to the theory of Statistics*. 3ed. New York: Joh Wiley & Sons, 1982. 564p.
- NEVES, R. C. S. M. *et al.* Avaliação dos métodos de otoscopia e exame do swab parasitológico no diagnóstico da otocariose canina: uma abordagem Bayesiana. *Pesquisa Veterinária Brasileira*. v.35, n.7, p.659-663, 2015.
- NOGUEIRA, D. A. *et al.* Análises clássica e Bayesiana de um modelo misto aplicado ao melhoramento animal: uma ilustração. *Ciência e Agrotecnologia*. Lavras, Edição especial, p.1614-1624, 2003.
- NOGUEIRA, D.A. *Proposta e avaliação de critérios de convergência para o método de Monte Carlo via Cadeias de Markov: casos uni e multivariados*. 2004. 135p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária). Universidade Federal de Lavras, Lavras - MG, 2004.
- PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. *Estatística Bayesiana*. Lisboa: Fundação Calouste Gulbenkian, 2003. 466p.
- PEREIRA, H. F. P.; SILVA, M. L. C. P. Estudo da adsorção de surfactante catiônico na matriz inorgânica fosfato de nióbio hidratado. *Química Nova*, v.32, n.1, p.7-11, 2009.
- PLUMMER, M. *et al.* coda: Convergence diagnosis and output analysis for MCMC. *R News*, v.6, n.1, p.7-11, 2006.
- R CORE TEAM: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Disponível em <http://www.r-project.org>. Acesso: 09 novembro de 2016.
- RAFTERY, A. E.; LEWIS, S. How many iterations in the Gibbs sampler, In *Bayesian Statistics*. 4ed. Oxford: Oxford University Press, p.763-773, 1992.
- REIS, R. L. *et al.* Inferência Bayesiana na análise genética de populações diplóides: estimação do coeficiente de endogamia e da taxa de fecundação cruzada. *Ciência Rural*, Santa Maria, v.38, n.5, p.1258-1265, 2008.
- RESENDE, M. D. V. Inferência Bayesiana e simulação estocástica (amostragem de Gibbs) na estimação de componentes de variância e de valores genéticos em plantas perenes. *Embrapa Florestas*. Documentos 46. Colombo, PR. 2000.
- RODRIGUES, G. M. A.; NEVES, R. F.; ANGÉLICA, R. S. Beneficiamento de uma argila tipo paligorskita da bacia de S. Luis-Grajaú, região de Alcântara, MA, e sua utilização como adsorvente de fósforo. *Cerâmica*, n.60, p.117-126, 2014.
- ROSA, G. J. M. *Análise Bayesiana de modelos mistos robustos via amostrador de Gibbs*. 1998. 57p. Tese (Doutorado em Estatística e Experimentação Agrônômica) - Universidade de São Paulo, Piracicaba, SP, 1998.
- SILVA, F. F. *et al.* Abordagem Bayesiana da curva de lactação de cabras Saanen de primeira e segunda ordem de parto. *Pesquisa agropecuária brasileira*, v.40, n.1, p.27-33, 2005.

- SILVA, N. A. M. *et al.* Aplicação do método Bayesiano na estimação de curva de crescimento em animais da raça nelore. *Revista Ceres*, v.54, n.312, p.192-199, 2007.
- SILVA, F. F. *et al.* Comparação Bayesiana de modelos de previsão de diferenças esperadas nas progênes no melhoramento genético de gado Nelore. *Pesquisa agropecuária brasileira*, Brasília, v.43, n.1, p.37-45, 2008.
- SOUZA, J. V. T. M. *et al.* Removal of Pb(II) from aqueous solution with Orange sub-products chemically modified as biosorbent. *BioResources*, v.7, n.2, p.2300-2318, 2012.
- STURTZ, S.; LIGGES, U.; GELMAN, A. R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software*, v.12, n.3, p.1-16, 2005.

Recebido em 28.04.2016.

Aprovado após revisão em 14.09.2016.