



**ANÁLISE BAYESIANA DO MODELO
AUTO-REGRESSIVO PARA DADOS EM
PAINEL: APLICAÇÃO NA AVALIAÇÃO
GENÉTICA DE TOUROS DA RAÇA NELORE**

FABYANO FONSECA E SILVA

2006

FABYANO FONSECA E SILVA

**ANÁLISE BAYESIANA DO MODELO AUTO-REGRESSIVO
PARA DADOS EM PAINEL: APLICAÇÃO NA AVALIAÇÃO
GENÉTICA DE TOUROS DA RAÇA NELORE**

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

BIBLIOTECA CENTRAL Orientadora

UFLA

Nº CLAS

T519.542

Prof. Dra. Thelma Sáfydi

SIK

ana

Nº REGISTRO

217235

DATA

15 / 05 / 07

LAVRAS

MINAS GERAIS - BRASIL

2006

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

Silva, Fabyano Fonseca e

Análise Bayesiana do modelo auto-regressivo para dados em painel:
aplicação na avaliação genética de touros da raça Nelore / Fabyano Fonseca e
Silva. – Lavras: UFLA, 2006.

100 p. : il.

Orientadora: Thelma Sáfydi.

Tese (Doutorado) – UFLA.

Bibliografia.

1. Análise Bayesiana. 2. Dados em painel. 3. Modelo auto-regressivo.
4. Touros Nelore. I. Universidade Federal de Lavras. II. Título.

CDD-591.542

FABYANO FONSECA E SILVA

**ANÁLISE BAYESIANA DO MODELO AUTO-REGRESSIVO
PARA DADOS EM PAINEL: APLICAÇÃO NA AVALIAÇÃO
GENÉTICA DE TOUROS DA RAÇA NELORE**

Tese apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-graduação em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de "Doutor".

Aprovada em quinta-feira, 14 de dezembro de 2006.

Prof. Dr. Carlos Henrique Osório Silva

UFV

Prof. Dr. Gerson Barreto Mourão

ESALQ/USP

Prof. Dr. Joel Augusto Muniz

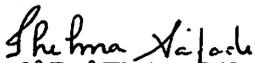
UFLA

Prof. Dr. Júlio Sílvio Souza Bueno Filho

UFLA

Prof. Dr. Renato Ribeiro de Lima

UFLA


Prof.^a Dr.^a Thelma Sáfyadi

DEX – UFLA

(Orientadora)

**LAVRAS
MINAS GERAIS – BRASIL**

2006

OFEREÇO

À minha filha Lara

À minha esposa Taciana

Aos meus pais, Gonçalves e Vera, e a minha irmã Tati

**“ ...coragem, coragem, se o que você quer é aquilo que pensa e faz,
coragem, coragem, eu sei que você pode mais... “**

Raul Seixas

AGRADECIMENTOS

A Deus, por ter concedido saúde, força e coragem para a realização deste trabalho.

À Universidade Federal de Lavras e ao Departamento de Ciências Exatas pela oportunidade e confiança no meu trabalho.

Ao CNPq pela concessão da bolsa de estudos.

À Prof.^a Thelma Sáfydi pela valiosa orientação e dedicação ao nosso trabalho.

Ao Prof. Luiz Henrique de Aquino, cujo apoio foi fundamental para a conclusão deste trabalho, por todos os ensinamentos e principalmente pela grande amizade.

Ao Prof. Joel Augusto Muniz pela co-orientação e ao grande incentivo em toda minha formação profissional.

Ao Prof. Gerson Barreto Mourão (ESALQ-USP) e ao GMA/FZEA (USP-Pirassununga), respectivamente, pela grande contribuição e pelo fornecimento dos dados.

À todos os professores do Departamento de Informática da Universidade Federal de Viçosa, em especial ao Prof. Paulo Cecon, Prof. Antônio Policarpo, Prof. Carlos Henrique, Prof. Adair Regazzi, Prof. Sebastião e Prof. Alexandre Peternelli, pela receptividade e pelo ótimo convívio em Viçosa.

Às grandes amigas construídas no DEX-UFLA: Telde, Mário Vivanco, Eric (barba), Celminha, Edila, Washington, Denis, Marcelo Cirillo, Natascha, Ricardo, Patrícia e Roberta, pelos felizes momentos de descontração.

Às grandes amigas construídas em Viçosa: Willian, Giselda, Cassiano, Juliana, Débora, Anderson e Dayse, pela receptividade e incentivo.

SUMÁRIO

Lista de Tabelas.....	i
Lista de Quadros.....	ii
Lista de Figuras.....	iii
Resumo.....	iv
Abstract.....	v
1 – Introdução.....	01
2 – Referencial Teórico.....	04
2.1 – Dados em painel.....	04
2.2 – Modelos auto-regressivos, AR(p), para dados em painel.....	05
2.3 – Inferência Bayesiana.....	07
2.4 – Algoritmos MCMC.....	10
2.5 – Avaliação da convergência dos algoritmos MCMC.....	13
2.6 – Função de verossimilhança exata em modelos auto-regressivos.....	16
2.7 – Informação a priori em modelos auto-regressivos.....	17
2.8 – Comparação de modelos.....	20
2.9 – Distribuições preditivas.....	23
2.10 – Dados de avaliações genéticas de touros da raça Nelore.....	25
3 – Material e Métodos.....	30
3.1 Função de verossimilhança exata.....	30
3.2 Distribuições a priori.....	33
3.2.1 Priori conjugada Normal multivariada - Gama Inversa.....	33
3.2.2 Priori independente t-Student multivariada - Gama Inversa.....	34
3.2.3 Priori não informativa de Jeffreys.....	34
3.3 Distribuições conjuntas a posteriori.....	35

3.3.1 Distribuição conjunta a posteriori (Priori conjugada Normal multivariada - Gama Inversa).....	35
3.3.2 Distribuição conjunta a posteriori (Priori independente t-Student multivariada - Gama Inversa).....	36
3.3.2 Distribuição conjunta a posteriori (Priori não informativa de Jeffreys).....	37
3.4 Distribuições condicionais completas a posteriori.....	37
3.4.1 Distribuições condicionais completas a posteriori (Priori conjugada Normal multivariada - Gama Inversa).....	37
3.4.2 Distribuições condicionais completas a posteriori (Priori independente t-Student multivariada - Gama Inversa).....	38
3.4.3 Distribuições condicionais completas a posteriori (Priori não informativa de Jeffreys).....	39
3.5 Implementação dos algoritmos MCMC.....	41
3.6 Distribuições preditivas.....	42
3.7 Fator de Bayes e Pseudo-Fator de Bayes: comparação de prioris.....	44
3.8 Simulação de dados.....	46
3.9 Análise dos dados reais.....	50
4 Resultados e Discussão.....	52
4.1. Apresentação dos dados simulados.....	52
4.2 Resultados obtidos para os dados simulados.....	55
4.3 Resultados obtidos para os dados reais.....	66
5 Conclusões.....	77
6 Referências Bibliográficas.....	78
Anexos A.....	86

LISTA DE TABELAS

Tabela 1. Organização dos dados em estrutura de painel.....	7
Tabela 2. Esquema de comparação das prioris consideradas.....	49
Tabela 3. Dados simulados por meio do modelo 1.....	52
Tabela 4. Dados simulados por meio do modelo 2.....	52
Tabela 5. Valor do parâmetro ϕ_1 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke, e “burn-in” considerando o modelo 1.....	55
Tabela 6. Valor do parâmetro ϕ_1 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke, e “burn-in” considerando o modelo 2.....	56
Tabela 7. Valor do Parâmetro ϕ_2 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke, e “burn-in” considerando o modelo 1.....	56
Tabela 8. Valor do Parâmetro ϕ_2 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke, e “burn-in” considerando o modelo 2.....	57
Tabela 9. Variâncias residuais usadas na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke, e “burn-in” considerando os modelos 1 e 2.....	57
Tabela 10. Estimativas (médias a posteriori) para a última observação (y_{12}) e intervalos de credibilidade (95%) para ambos os modelos.....	62
Tabela 11. Modelos utilizados na simulação dos dados e respectivos critérios de comparação.....	64
Tabela 12. Estimativas de máxima verossimilhança e desvio-padrão (DP).....	68

Tabela 13. Modelos considerados na análise dos dados reais referentes à Dep's calculadas no decorrer dos anos de 2000 a 2004 e respectivos critérios de comparação.....	70
Tabela 14. Resultados referentes à porcentagem de significância dos parâmetros do modelo AR(2) ajustado às séries individuais de Dep's.....	71
Tabela 15. Estimativas (médias a posteriori) para a última observação (y_6) e intervalos de credibilidade (95%) considerando a estrutura de acurácia baixa...	73
Tabela 16. Estimativas (médias a posteriori) para a última observação (y_6) e intervalos de credibilidade (95%) considerando a estrutura de acurácia média...	74
Tabela 17. Estimativas (médias a posteriori) para a última observação (y_6) e intervalos de credibilidade (95%) considerando a estrutura de acurácia alta.....	75
Tabela 18. Estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, e “burn-in” para cada estrutura de dados.....	75

LISTA DE QUADROS

Quadro 1. Descrição do algoritmo Metropolis-Hastings.....	12
Quadro 2. Descrição do algoritmo Gibbs Sampler.....	13
Quadro 3. Distribuições de probabilidade utilizadas no processo de simulação dos dados por meio do modelo 1.....	47
Quadro 4. Distribuições de probabilidade utilizadas no processo de simulação dos dados por meio do modelo 2.....	47

LISTA DE FIGURAS

Figura 1. Série original estudada por Powell et al. (2004).....	28
Figura 2. Série original estudada por MacNeil (2005).....	29
Figura 3. Distribuição Normal multivariada (a) e Distribuição t-Student multivariada (b) utilizadas na simulação dos dados.....	48
Figura 4. Distribuição Gama Inversa utilizada na simulação dos dados.....	48
Figura 5. Séries obtidas para os indivíduos 1, 5 e 10 , modelo 1.....	53
Figura 6. Séries obtidas para os indivíduos 1, 5 e 10 , modelo 2.....	54
Figura 7. Box-plot para os parâmetros de auto-regressão de todos os indivíduos	57
Figura 8. Cadeias geradas pelo método MCMC e densidade a posteriori para os parâmetros ϕ_1 e ϕ_2 referentes ao indivíduo 1 (modelo 1).....	60
Figura 9. Cadeias geradas pelo método MCMC e densidade a posteriori para os parâmetros ϕ_1 e ϕ_2 referentes ao indivíduo 1 (modelo 2).....	60
Figura 10. Cadeias geradas pelo método MCMC e densidade a posteriori para a variância do erro.....	61
Figura 11. Distribuições preditivas de y_{12} para o indivíduo 1 (modelos 1 e 2)	61
Figura 12. Amplitudes do intervalo de credibilidade de 95% para y_{12} referentes a todos os indivíduos.....	62
Figura 13. Box-plot para valores do conjunto representado pela acurácia alta...66	66
Figura 14. Box-plot para valores do conjunto representado pela acurácia média	66
Figura 15. Box-plot para valores do conjunto representado pela acurácia alta...67	67
Figura 16. Box-plot para valores do conjunto representado pela ausência da divisão por classes de acurácia.....	67
Figura 17. Série obtida com os dados médios para conjunto classificado de acordo com a acurácia: baixa (a), média (b), alta (c) e geral (d).....	69
Figura 18. Cadeias geradas pelo método MCMC e densidade a posteriori para a variância do erro em cada estrutura de dados considerada.....	76

ABSTRACT

SILVA, Fabyano Fonseca e. **Bayesian analysis of autoregressive panel data model: application in genetic evaluation of Nelore sires.** 2006. 100 p. Thesis (Phd Program in Statistics and Agricultural Experimentation) – Federal University of Lavras, Lavras, Minas Gerais, Brazil*

We considered a Bayesian analysis of an autoregressive, $AR(p)$, panel data model, using exact likelihood function, comparative analysis of priors distributions and predictive distributions of future observations. The methodology was tested by a simulation study using three priors: hierarchical Multivariate Normal-Inverse Gamma (model 1), independent Multivariate t-Student – Inverse Gamma (model 2) and Jeffreys (model 3). These comparisons were realized by Bayes Factor and Pseudo-Bayes Factor, and the results showed the model 2 superiority. We also applied the proposed methodology application to real data from Nelore sires Expected Progenie Difference (EPD), observed during a five year period (2000-2005). Results indicated the importance of sires grouping by accuracy values, and also showed forecast efficiency around 80%.

*Guidance Committe: Prof^a Thelma Sáfaci, Prof. Joel Augusto Muniz, Prof. Luiz Henrique de Aquino.

RESUMO

SILVA, Fabyano Fonseca e. **Análise Bayesiana do modelo auto-regressivo para dados em painel: aplicação na avaliação genética de touros da raça Nelore.** 2006. 100 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG. *

Considerou-se uma análise Bayesiana de modelos auto-regressivos de ordem p , $AR(p)$, para dados em painel, de forma a utilizar a função de verossimilhança exata, a análise de comparação de distribuições a priori e a obtenção de distribuições preditivas de dados futuros. A metodologia utilizada foi testada mediante um estudo de simulação usando a priori hierárquica Normal multivariada-Gama inversa (modelo 1) a priori independente t-Student Gama inversa (modelo 2) e a priori de Jeffreys (modelo 3). As comparações entre os modelos, realizadas por meio do Fator de Bayes e do Pseudo-Fator de Bayes, indicaram uma superioridade do modelo 2 em relação aos demais. Realizou-se uma aplicação em dados reais referentes às DEPs de touros da raça Nelore publicadas durante os anos de 2000 a 2005. Os resultados apontaram para a importância de se dividir os animais em grupos homogêneos de acordo com a acurácia. Constatou-se também, que em média, a eficiência de previsão dos valores de DEPs para um ano futuro foi próxima de 80%.

*Comitê Orientador: Prof^a Thelma Sáfyadi, Prof. Joel Augusto Muniz, Prof. Luiz Henrique de Aquino.

1 INTRODUÇÃO

Os dados em estrutura de painel estão essencialmente orientados para estudar a heterogeneidade relativa aos diferentes indivíduos que apresentam uma trajetória longitudinal. Este fato permite identificar determinados aspectos que são de difícil quantificação usando apenas modelos de regressão individuais. Modelos de regressão adaptados para estudar dados em painel são de grande aplicabilidade em diversas áreas da ciência, como em Economia, Sociologia e Biologia.

Segundo Marques (2000) a análise estatística de dados em painel é essencial, mesma que seja complexa, pois a grande quantidade de informação disponível neste tipo de conjunto de dados aumenta a eficiência da estimação, ou seja, os dados em painel permitem identificar e medir efeitos que não serão pura e simplesmente detectáveis em estudos exclusivamente individuais.

Uma classe de modelos de regressão bastante retratada em estudos de dados em painel é a de Séries Temporais. Um modelo de Série Temporal é utilizado quando se dispõe de uma seqüência de dados obtidos em intervalos regulares de tempo durante um período específico. Dentre estes modelos, destaca-se o auto-regressivo, pois ele se aplica às diversas situações práticas e geralmente apresenta boa qualidade de ajuste quando comparado com modelos mais complexos. De forma geral, ao estudar estes modelos, primeiramente deseja-se compreender o fenômeno focado para, a partir daí, descrever o comportamento da série, obter as estimativas para os coeficientes dos modelos e, por último, realizar previsões fundamentadas nas estimativas obtidas anteriormente.

Quando o objetivo é ajustar um modelo de série temporal a dados em estrutura de painel deve-se considerar que cada indivíduo, ou componente, apresenta o seu conjunto particular de observações ou sua série, e a combinação

de informações de todos estes indivíduos é utilizada para estimar os coeficientes individuais de cada série, de forma a considerá-los provenientes de uma mesma população. Dessa maneira, também é importante estimar os parâmetros da distribuição representativa da população de coeficientes, o que implica na utilização de modelos hierárquicos, os quais muitas vezes tornam-se inviáveis de serem considerados pela inferência Frequentista (Liu & Tiao, 1980).

Sob o ponto de vista da inferência Bayesiana, esta hierarquia pode ser facilmente incorporada, devido ao fato de se considerar informações a priori a respeito dos parâmetros a serem estimados. Portanto, assume-se uma distribuição a priori para os coeficientes dos modelos e assume-se também distribuições a priori para os parâmetros da distribuição adotada para estes coeficientes, obtendo-se assim a inserção da hierarquia requerida.

Sendo a teoria Bayesiana fundamentada em uma distribuição conjunta dos dados amostrais, denominada função de verossimilhança, e nas distribuições a priori a respeito dos parâmetros, torna-se necessário ampliar o leque de conhecimento a respeito destes dois fatores.

É importante realçar que ao se utilizar modelos auto-regressivos de ordem p , os quais apresentam termos de defasagem temporal, verifica-se uma redução na quantidade de observações que compõem a função de verossimilhança quando esta se apresenta condicionada as p primeiras observações, uma vez que essas observações não são consideradas para a análise. Ao se tratar de dados em painel, este problema relacionado com a utilização da função de verossimilhança condicional é ampliado, pois nesta situação a condicionalidade é inerente a todos os indivíduos, o que acarreta uma perda considerável de informações. Uma maneira de solucionar este problema é a designação de uma função de verossimilhança exata, ou incondicional, pois apesar de proporcionar uma maior complexidade ao processo de análise, permite examinar todas as observações.

Em relação à escolha de distribuições a priori para os parâmetros dos modelos auto-regressivos, também é importante mencionar que esta deve ser realizada de acordo com uma metodologia apropriada, que seja capaz de indicar quais distribuições são mais adequadas. Este estudo de comparação de prioris que, atualmente, vem sendo bastante usado na inferência Bayesiana, é denominado de análise de robustez ou análise de sensibilidade de prioris.

Segundo Box et al. (1994) a complexidade de modelos auto-regressivos e de médias móveis, com especificações sazonais e/ou com estimação de tendências, utilizados em estudos de Séries Temporais pode ser justificada por meio de aplicações em situações reais. Uma possível utilização prática de modelos auto-regressivos para dados em painel está relacionada com modelagem e previsão de valores genéticos, também denominados de Diferenças Esperadas nas Progênieis (DEP). Esta aplicação é possível porque as DEP's são calculadas e publicadas anualmente em sumários, e, ao verificar que um mesmo touro consta em diferentes sumários, ou seja, tem seus valores genéticos calculados ao longo de vários anos, tem-se então uma série temporal. Ao se considerar vários touros, nascidos e criados sob as mesmas condições, torna-se possível construir um arquivo de dados estruturado em forma de painel.

O desenvolvimento deste trabalho teve como objetivo a apresentação de um estudo relacionado com a análise Bayesiana de modelos auto-regressivos de ordem p , $AR(p)$, para dados em painel, de forma a considerar a função de verossimilhança exata, a análise de comparação de distribuições a priori e a obtenção de distribuições preditivas de dados futuros. A metodologia utilizada foi testada mediante um estudo de simulação considerando a priori hierárquica Normal multivariada-Gama inversa (modelo 1) a priori independente t-Student-Gama inversa (modelo 2) e a priori de Jeffreys (modelo 3). As comparações entre os modelos foram testadas por meio do Fator de Bayes e do Pseudo-Fator

de Bayes. Realizou-se uma aplicação em dados reais referentes às DEPs de touros da raça Nelore calculadas durante os anos de 2000 a 2005.

2 REFERENCIAL TEÓRICO

2.1 Dados em painel

Os dados em painel sugerem a existência de características diferenciadoras dos indivíduos, entendidos como “unidade estatística de base”. Essas características podem não ser constantes ao longo do tempo, de tal forma que estudos temporais podem ser requeridos conjuntamente com a análise do painel (Hsiao, 2003).

De acordo com Hirano (2002) esta estrutura permite que sejam realizadas simultaneamente uma análise “cross-section” e uma análise temporal. Este autor explica esta característica por meio de um exemplo que consiste em dados de Pesquisas Industriais Anuais (PIA). É possível fazer a análise para um determinado ano usando várias firmas, isto é, usando dados em “cross-section”, e também para uma firma em vários anos, cuja estrutura caracteriza-se como dados temporais. Uma forma mais complexa, e mais informativa, é realizar a análise para várias firmas em vários anos conjuntamente, caracterizando assim a estrutura em painel.

Segundo Frees (2004) os dados em painel proporcionam uma maior quantidade de informação, maior número de graus de liberdade e maior eficiência na estimação. Adicionalmente, os estudos com amostras longitudinais possibilitam uma análise mais eficiente do ajustamento, pois os estudos seccionais, ao não contemplarem a variabilidade longitudinal, transmitem uma falsa idéia de estabilidade. Assim, a utilização de dados em painel permite

conjugar a diversidade de comportamentos individuais, com a existência de um padrão, ou seja, permite tipificar as respostas de diferentes indivíduos.

A maior quantidade de informação disponível aumenta a eficiência da estimação, isto é, os dados em painel permitem identificar e medir efeitos que não serão pura e simplesmente detectáveis em estudos exclusivamente seccionais ou temporais, bem como construir e testar modelos comportamentais complexos (Baltagi, 2000).

Em relação a algumas desvantagens de se utilizar uma análise de dados em painel, Hsiao & Sun (2000) relatam que se considerarmos uma população como um conjunto de decisões que se refletem em diferentes histórias individuais, estas terão que ser representadas como variáveis aleatórias específicas a cada indivíduo e que, certamente, estarão correlacionadas, não apenas com a variável dependente, mas também com as características que envolvem o indivíduo, o que causa diversos problemas relacionados com a identificação de modelos e estimação de seus parâmetros. Para evitar este problema os autores recomendam que é necessário optar por um conjunto de indivíduos homogêneos, que possam realmente ser identificados como elementos de uma mesma população.

2.2 Modelos auto-regressivos, AR(p), para dados em painel

A grande área da estatística denominada de Séries Temporais engloba uma gama enorme de modelos, cada um apresentando características próprias, que os diferenciam em relação às suas complexidades de análises. Dentre estes, um dos mais estudados e utilizados é o modelo auto-regressivo, denotado por AR(p), em que p é o número de parâmetros considerados, o qual define a ordem do modelo. Este modelo é apresentado pela Fórmula (1):

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t = \sum_{j=1}^p \phi_j Y_{t-j} + e_t \quad (1)$$

em que:

Y_t é o valor atual de um processo estocástico, cujos valores já assumidos no passado são dados por $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$;

$\phi_1, \phi_2, \dots, \phi_p$ são os parâmetros do modelo, denominados de parâmetros de auto-regressão;

e_t é o resíduo associado ao modelo, também denominado de ruído branco,

$$e_t \stackrel{iid}{\sim} N(0, \sigma_e^2).$$

De acordo com Morettin & Tolo (2004) o ajuste de um modelo AR(p) a um determinado conjunto de dados, que são descritos como uma série temporal, deve seguir os estágios: i) estimação dos parâmetros $\phi_1, \phi_2, \dots, \phi_p$ e de σ_e^2 ; ii) seleção da ordem do modelo, ou seja, do valor de p , de acordo com um critério estatístico; iii) utilização das estimativas obtidas para realizar predição de valores futuros.

Na presença de dados em estrutura de painel, os quais podem ser visualizados na Tabela 1, o modelo AR(p) indicado na Fórmula (1) continua apresentando as mesmas características, porém a ele deve ser acrescentado um índice i referente ao indivíduo (Liu & Tiao, 1980), como na Fórmula (2):

$$Y_{it} = \phi_{i1} Y_{i(t-1)} + \phi_{i2} Y_{i(t-2)} + \dots + \phi_{ip} Y_{i(t-p)} + e_{it}$$

$$Y_{it} = \sum_{j=1}^p \phi_{ij} Y_{i(t-j)} + e_{it} \quad (2)$$

em que: $i=1,2,\dots,m$; $j=1,2,\dots,p$ e $t=1,2,\dots,n_i$.

De acordo com esta notação, tem-se m indivíduos, com n_i observações longitudinais cada um, indicando que cada indivíduo i pode apresentar um

número diferente de observações. Nota-se também que o modelo contempla p parâmetros por indivíduo.

No modelo apresentado em (2) assume-se que os resíduos associados às observações de cada série são amostras aleatórias provenientes de uma mesma distribuição, portanto $e_i \sim N(0, \sigma_e^2)$

Tabela 1. Organização dos dados em estrutura de painel.

Indivíduo (i)	Tempo (t)					n_i
	1	2	...	t	...	
1	Y_{11}	Y_{12}	...	Y_{1t}	...	Y_{1n_1}
2	Y_{21}	Y_{22}	...	Y_{2t}	...	Y_{2n_2}
.
.
.
m	Y_{m1}	Y_{m2}	...	Y_{mt}	...	Y_{mn_m}

2.3 Inferência Bayesiana

Métodos Bayesianos são atualmente um dos principais assuntos da comunidade científica envolvida com o desenvolvimento e aplicação de procedimentos estatísticos, pois estes têm apresentado grande versatilidade na resolução de problemas nunca antes solucionados, principalmente em relação à análise de modelos hierárquicos.. 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 60 ela ressurgiu em alguns trabalhos teóricos como o de Jeffreys (1961), mas somente em 1990 com o trabalho de Gelfan &

Smith (1990), os quais utilizaram algoritmos da classe MCMC (Markov Chain – Monte Carlo) é que o problema da resolução de integrais foi solucionado de uma maneira alternativa, uma vez que os resultados foram excelentes, o que atraiu a atenção de muitos outros pesquisadores.

Em recentes estudos envolvendo análise Séries Temporais e de modelos de regressão não linear (Barreto & Andrade, 2004; Silva et al., 2005 e Silva et al., 2006) a metodologia Bayesiana foi utilizada com sucesso, pois sua característica de considerar todos os parâmetros como variáveis aleatórias, segundo esses autores, reduziu substancialmente o número de estimativas viesadas. Além disso, relatam também que a utilização desta metodologia requer um número menor 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 considerado.

Segundo O'Hagan (1994) outro fato que merece destaque é que na metodologia Bayesiana a obtenção de intervalos de credibilidade é imediata, levando em consideração a incerteza existente sobre todos os parâmetros simultaneamente, sendo, portanto, a estimação por intervalo geralmente mais precisa em relação àquela apresentada pela metodologia frequentista, que em geral utiliza variâncias assintóticas e aproximadas.

A inferência Bayesiana consiste de uma informação a priori, dos dados amostrais e do cálculo da densidade a posteriori dos parâmetros. A informação a priori é dada pela densidade de probabilidade $P(\theta)$, a qual expressa o conhecimento do pesquisador sobre os parâmetros a serem estimados. Quando em determinado estudo o pesquisador tem pouca ou nenhuma informação para incorporar à priori considera-se uma não-informativa, por exemplo a priori de Jeffreys (Jeffreys, 1961). Os dados $Y = \{ y_1, y_2, \dots, y_n \}$, representados por uma amostra aleatória de uma população com densidade f , são utilizados na análise

Bayesiana através da função de verossimilhança $L(y_1, \dots, y_n | \theta)$, que é a densidade conjunta destes dados.

Portanto, a partir do momento em que se opta por uma distribuição a priori, seja ela informativa ou não e obtém-se a função de verossimilhança, é possível, por meio do Teorema de Bayes, representado pela expressão (3), obter a distribuição densidade a posteriori de θ , de forma que qualquer conclusão a seu respeito é realizada a partir desta distribuição,

$$P(\theta | Y) = \frac{L(Y | \theta)P(\theta)}{\int L(Y | \theta)P(\theta)d\theta}, \quad (3)$$

sendo $Y = \{y_1, y_2, \dots, y_n\}$. O denominador, chamado de constante de integração, não depende de θ , portanto temos:

$$P(\theta | Y) \propto L(Y | \theta)P(\theta), \quad (4)$$

ou seja, a expressão (4) pode ser entendida como:

$$\text{Posteriori} \propto \text{Verossimilhança} \times \text{Priori},$$

em que \propto representa proporcionalidade.

Segundo Broemiling (1989), pode-se pensar no Teorema de Bayes como um mecanismo de atualização da opinião do estatístico sobre θ , portanto este teorema constitui a base da inferência Bayesiana, pois toda prática inferencial é realizada a partir da distribuição *a posteriori* obtida.

A distribuição *a posteriori* de um parâmetro contém toda a informação probabilística a respeito do mesmo. Assim, toda a inferência sobre o parâmetro é realizada por meio desta distribuição. Segundo Rosa (1998) para se inferir em relação a qualquer elemento de θ , a distribuição *a posteriori* conjunta dos parâmetros, $P(\theta | Y)$, deve ser integrada em relação a todos os outros elementos que a constituem. Assim, se o interesse do pesquisador se concentra sobre

determinado conjunto de θ , por exemplo θ_1 , tem-se a necessidade da obtenção da distribuição $p(\theta_1 | Y)$, denominada de marginal, a qual é dada por:

$$P(\theta_1 | Y) = \int_{\theta \neq \theta_1} P(\theta | Y) d\theta_{\theta \neq \theta_1}$$

A integração da distribuição conjunta *a posteriori* para a obtenção das marginais geralmente não é analítica, necessitando de algoritmos iterativos especializados como o Gibbs Sampler e o Metropolis-Hastings, os quais são denominados de algoritmos MCMC (Markov Chain-Monte Carlo). Portanto, para a utilização desses algoritmos, é necessário que se obtenha, a partir da distribuição *a posteriori*, um conjunto de distribuições chamadas de distribuições condicionais completas.

Recentemente, nota-se um interesse crescente em métodos estatísticos Bayesianos aplicados a diversas áreas da ciência, como em Epidemiologia, Bioestatística, Engenharia e outras. De acordo com Asai (2005), desde a segunda metade da década de noventa o maior desenvolvimento desta metodologia, especialmente no que se diz respeito aos algoritmos relacionados com o método das Cadeias de Markov – simulação Monte Carlo (MCMC), tem sido observado na área de Séries Temporais aplicadas principalmente à área de Econometria e Finanças. Isto se deve ao fato desses algoritmos fornecerem resultados mais rápidos e confiáveis, direcionados com solução numérica de expressões complexas resultantes do tratamento estatístico de modelos utilizados nessas áreas.

2.4 Algoritmos MCMC

Os métodos MCMC são uma alternativa aos métodos não iterativos tendo em vista a resolução de integrais complexas. A idéia é obter uma amostra das distribuições marginais *a posteriori* dos parâmetros de interesse por meio de um processo iterativo. Segundo Gamerman (1997) uma cadeia de Markov é um

processo estocástico em que a probabilidade de estar em um certo estado em um tempo futuro pode depender do estado atual do sistema, mas não dos estados em tempos passados, ou seja, os valores gerados no processo apresentam uma dependência. Por sua vez, estes valores gerados são considerados amostras aleatórias de uma determinada distribuição de probabilidade, caracterizando assim o método de simulação Monte Carlo. Dessa forma tem-se uma ação conjunta desses dois métodos que resulta no método MCMC, cujos principais algoritmos são o Metropolis-Hastings e o Gibbs Sampler.

O algoritmo de Metropolis-Hastings é utilizado para a obtenção da distribuição marginal *a posteriori* quando o amostrador de Gibbs não se mostra eficiente, ou seja, para parâmetros cuja distribuição condicional não se caracteriza como uma distribuição de probabilidade conhecida. Neste caso, geram-se valores do parâmetro a partir de uma distribuição proposta e esse é aceito ou não com uma certa probabilidade de aceitação (Chib & Greenberg, 1995).

Segundo Berg (2004) para descrever o algoritmo Metropolis-Hastings, suponha que a distribuição de interesse é a distribuição a posteriori, $P(\theta|Y)$, sendo θ um vetor de parâmetros a ser estimado, $\theta = (\theta_1, \theta_2, \dots, \theta_K)$, $k=1,2,\dots,K$. Considere também que todas as condicionais completas a posteriori, $P(\theta_k|\theta_{-k}, Y)$, estejam disponíveis mas não se sabe gerar amostras diretamente de cada uma, e que amostras de um novo valor de θ_k , serão geradas a partir de uma distribuição proposta, ou candidata, $C(\theta_k)$. Este valor gerado, denominado inicialmente de θ_k^c , constituirá, ou não, uma amostra da distribuição marginal a posteriori de θ_k conforme determinação de um critério probabilístico imposto pelo pesquisador.

Os valores dos parâmetros gerados por esses algoritmos, após verificação de q iterações, $q=1,2,\dots,Q$, são utilizados para formar uma amostra aleatória, $\{\theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_K^{(1)}, \theta_1^{(2)}, \theta_2^{(2)}, \dots, \theta_K^{(2)}, \dots, \theta_1^{(q)}, \theta_2^{(q)}, \dots, \theta_K^{(q)}\}$. À medida que o número de iterações aumenta, o conjunto de valores gerados aproxima de sua condição de equilíbrio. Assim assume-se que a convergência é atingida em uma iteração cuja distribuição esteja arbitrariamente próxima da distribuição de equilíbrio, ou seja, da distribuição marginal desejada (Sorensen & Gianola, 2004).

O algoritmo de Metropolis-Hastings é bastante geral, e pode, pelo menos em princípio, ser implementado com qualquer distribuição condicional completa a posteriori e para qualquer proposta. Entretanto sob o ponto de vista prático, a escolha da proposta é crucial para o bom desenvolvimento do algoritmo, ou seja, para sua convergência para a distribuição marginal a posteriori (Chib & Greenberg, 1995). No Quadro 1 é apresentado um esquema ilustrativo do algoritmo Metropolis-Hastings.

<p>I - Inicialize $\theta^{(0)} = \theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_K^{(0)}$;</p> <p>$\alpha = \min \left\{ 1, \frac{P(\theta_x^q \theta_{-x}, Y) \rightarrow \text{valor gerado: condicional completa de } \theta_x \text{ via } \theta_x^c}{P(\theta_x^{q-1} \theta_{-x}, Y) \rightarrow \text{valor aceito para } \theta_x \text{ na iteração anterior } (q-1)} \right\}$</p> <p>(ii) Aceite θ_x^c com probabilidade dada por:</p> <p>onde $\theta_{-x}^{(a)} = (\theta_1^{(a)}, \dots, \theta_{i-1}^{(a)}, \theta_{i+1}^{(a)}, \dots, \theta_S^{(a)})$.</p> <p>Se $\alpha > u$, $u \sim U[0,1]$, aceita-se o valor, caso contrário permanece o valor obtido na iteração anterior.</p> <p>III - Faça $q = q + 1$ até atingir o número estipulado (Q) de iterações.</p>

Quadro 1. Descrição do algoritmo Metropolis-Hastings.

O algoritmo Gibbs Sampler é, essencialmente, um esquema iterativo de amostragem de uma cadeia de Markov, cujo núcleo de transição é formado pelas distribuições condicionais completas. É uma técnica para gerar variáveis aleatórias de uma distribuição marginal quando se conhece a sua densidade (Gamerman, 1997). À medida que o número q de iterações aumenta, a seqüência de valores gerados se aproxima da distribuição de equilíbrio, ou seja, da densidade marginal desejada para cada parâmetro, quando se assume que a convergência foi atingida.

Para descrever o algoritmo, suponha que a distribuição de interesse seja uma distribuição a posteriori $P(\theta|Y)$, com $\theta = (\theta_1, \theta_2, \dots, \theta_K)$ e considere também que todas as condicionais completas a posteriori $P(\theta_k | \theta_{-k}, Y)$, $k=1, 2, \dots, K$, estejam disponíveis e que sabe-se gerar amostras de cada uma delas (Casella & George, 1992). No Quadro 2 é apresentado o esquema ilustrativo do algoritmo Gibbs Sampler.

<p><i>I - Inicializ: $\theta_k^{(q)} \sim P(\theta_k \theta_1^{(q-1)}, \dots, \theta_{k+1}^{(q-1)}, \dots, \theta_K^{(q-1)}, Y)$</i></p> <p><i>II – Obtenha um novo valor para $\theta_k^{(q)}$ por meio do seguinte critério:</i></p> <p style="padding-left: 40px;"><i>i) Gere um valor para $\theta_k^{(q)}$ diretamente de sua condicional completa:</i></p> <p><i>III – Faça $q = q + 1$ até atingir o número estipulado (Q) de iterações.</i></p>
--

Quadro 2. Descrição do algoritmo Gibbs Sampler.

2.5 Avaliação da convergência dos algoritmos MCMC

Os algoritmos MCMC são processos iterativos, portanto surge a seguinte questão: Quantas iterações são necessárias para que a convergência seja verificada ?

A resposta definitiva para esta questão poderá nunca ser dada, visto que a distribuição estacionária será na prática desconhecida, mas pode-se sempre avaliar a convergência das cadeias detectando problemas relacionados com a sua instabilidade. Estes problemas podem ser observados preliminarmente analisando os gráficos ou medidas descritivas dos valores simulados dos parâmetros de interesse para cada tamanho de cadeia considerado, de forma a se obter um determinado tamanho de cadeia no qual não se nota mais diferenças nos gráficos e nos valores das estatísticas descritivas. Segundo Gelman et al. (2000) os gráficos mais freqüentes são o gráfico de θ ao longo das iterações e um gráfico da estimativa da distribuição *a posteriori* de θ , por exemplo, um histograma ou uma densidade kernel, e as estatísticas usuais são a média, o desvio padrão e os quantis (2,5%; 50% e 97,5%).

Ao se analisar a convergência de uma seqüência gerada por meio de um algoritmo MCMC é comum descartar as primeiras iterações, em geral, de 40% a 50% do total, considerando-se que essa primeira parte esteja sendo influenciada pelos valores iniciais. De acordo com Gamerman (1997), este início da cadeia é chamado de período de aquecimento ou “burn-in”. Segundo este mesmo autor, outro aspecto importante relacionado com a avaliação da convergência diz respeito a dependência entre as observações subseqüentes da cadeia. Para se obter uma amostra independente, as observações devem ser espaçadas por um determinado número de iterações, ou seja, considerar saltos (“thin”) de tamanho k , usando, para compor a amostra, os valores a cada k iterações.

De acordo com Gelman et al. (2000) para avaliação da convergência dos algoritmos MCMC vários testes foram propostos, dentre eles os mais relevantes são: Geweke (1992), Gelman & Rubin (1992), Heidelberg & Welch (1983) e Raftery-Lewis (1992b).

O critério de Geweke (1992) é fundamentado em técnicas de análise espectral, e fornece um diagnóstico para a ausência de convergência. Este

propõe o diagnóstico de convergência para Cadeias de Markov baseados no teste de igualdade de médias da primeira e última parte da cadeia de Markov, geralmente dos primeiros 10% e dos últimos 50%.

O critério de Gelman & Rubin (1992) pressupõe que m cadeias tenham sido geradas em paralelo, partindo de diferentes valores iniciais, num total de $2n$ iterações, das quais n são descartadas (“burn-in”). As m seqüências rendem m possíveis inferências. Se estas inferências são similares tem-se um indicativo de que a convergência foi alcançada ou está próxima.

O critério de Raftery & Lewis (1992b) fornece estimativas do número de iterações necessárias para se obter a convergência, do número de iterações iniciais que devem ser descartadas (burn-in) e da distância mínima (k) de uma iteração à outra para se obter uma amostra independente. Esses valores são calculados mediante especificações para garantir que um quantil de uma determinada função seja estimado com uma precisão predefinida.

O critério de Heidelberger & Welch (1983) propõe testar a hipótese nula de estacionariedade da seqüência gerada, por meio de testes estatísticos. Se a hipótese nula é rejeitada para um dado valor, o teste é repetido depois de descartados os 10% valores iniciais da seqüência. 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, isto indica que é necessário um número maior de iterações. Caso contrário, o número inicial de iterações descartadas é indicado como o tamanho do “burn-in”.

Após um intenso estudo de todos estes critérios, e tendo em vista resultados obtidos via simulação de dados, Nogueira (2004) concluiu que o seguinte protocolo deve ser utilizado para que se tenha uma avaliação precisa da convergência. O procedimento recomendado é o seguinte:

- i) aplicar Raftery e Lewis em uma amostra piloto e determinar o tamanho ideal da seqüência;
- ii) 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;
- iii) determinar o tamanho do “burn-in”, ou seja, do número de iterações desprezadas para se retirar o efeito dos valores iniciais, pelo critério de Heidelberger e Welch.

2.6 Função de verossimilhança exata em modelos auto-regressivos

Ao se considerar um modelo auto-regressivo, devido à defasagem implícita no índice relativo à ordem do modelo, tem-se uma função de verossimilhança incompleta quando esta se apresenta condicionada a p primeiras observações, pois se um modelo tem ordem p , nota-se que as p primeiras observações serão ignoradas no produtório da verossimilhança. Em se tratando de dados em estrutura de painel, a soma de p observações referentes a cada indivíduo, constitui uma ausência de informação considerável para a realização da inferência estatística.

De acordo com Sowell (1992) a utilização de funções de verossimilhança aproximadas, relacionadas com ajuste de modelos auto-regressivos, faz com que o processo seja computacionalmente mais simples, mas as estimativas dos parâmetros são obtidas com menor acurácia, principalmente quando se tem um número pequeno de observações ou se verifica um comportamento da série próximo da não estacionaridade.

Uma possível solução para este problema é considerar uma função de verossimilhança denominada de exata, a qual contempla um mecanismo de utilização das observações naturalmente desconsideradas pela defasagem do

modelo. Sob o ponto de vista frequentista, Dufour & Kiviet (1998) relatam que o uso da verossimilhança exata para modelos auto-regressivos implica na necessidade de utilização de algoritmos de otimização complexos, que muitas vezes mostram-se até mesmo imprecisos.

A implementação da função de verossimilhança exata sob o enfoque Bayesiano, segundo Pai et al. (1994), torna o processo de estimação mais natural, embora este eleve a complexidade para a obtenção das marginais. Atualmente, com os recursos computacionais empregados e o advento dos métodos MCMC, a especificação de uma verossimilhança exata não inviabiliza sua utilização na inferência Bayesiana.

De acordo com Samaniego (2002) ao se utilizar a função de verossimilhança exata na inferência Bayesiana de modelos auto-regressivos percebe-se que a densidade conjunta a posteriori adquire uma forma mais complexa em relação a obtida mediante a utilização da verossimilhança aproximada. Segundo este autor um dos objetivos do trabalho em questão foi mostrar que, ao se ajustar um modelo auto-regressivo de primeira ordem, AR(1), para uma única série, pode ser usada a verossimilhança aproximada dos dados ao invés da verossimilhança exata, de forma a obter resultados similares simplificando assim vários cálculos. Porém esta conclusão não se mantém quando se trabalha com modelos auto-regressivos de ordens superiores e/ou se dispõe de dados em estrutura de painel (Dufour & Kiviet, 1998).

2.7 Informação a Priori em modelos auto-regressivos

Sendo toda a metodologia Bayesiana fundamentada em $P(\theta|Y) \propto L(\theta|Y)P(\theta)$, recentes estudos (Barreto & Andrade, 2004; Ghosh & Heo, 2003; Ni & Sun, 2003) têm sido desenvolvidos para selecionar prioris, $P(\theta)$, que sejam mais adequadas para os modelos usados na teoria de

de Séries Temporais. Dessa forma, criou-se um termo denominado de robustez, também interpretado como sensibilidade, para avaliar a qualidade das prioris adotadas. Quanto mais robusta uma priori, mais precisas serão as inferências realizadas a posteriori (Greenhouse & Wasserman, 1995). Portanto há a necessidade de se comparar diferentes classes de prioris por meio de metodologias eficientes capazes de representar com fidedignidade seus efeitos sobre a posteriori (Berger & Pericchi, 2001). Antes de descrever os métodos usados na comparação de prioris, torna-se interessante apresentar as características das distribuições a priori mais utilizadas.

Quando se espera que a informação dos dados seja dominante, no sentido de que a nossa informação a priori seja vaga, os conceitos de “conhecimento vago”, “não informação”, ou “ignorância a priori”, são claramente aceitos sob o ponto de vista Bayesiano, pois simplesmente caracterizam um tipo de priori a ser usada (Ehlers, 2005). Portanto, reconhece-se a necessidade de alguma forma de análise que consiga captar esta noção de uma priori que tenha um efeito mínimo, relativamente aos dados, na inferência final. Tal análise pode ser pensada como um ponto de partida quando não se consegue fazer uma indicação detalhada do “verdadeiro” conhecimento a priori.

A primeira idéia de “não informação” a priori que se pode ter é pensar em todos os possíveis valores de θ como igualmente prováveis, ou seja, com uma distribuição a priori uniforme. Neste caso, fazendo $P(\theta) \propto K$ para θ variando em um subconjunto da reta significa que nenhum valor particular tem preferência, ou seja, admitir uma distribuição uniforme como priori. Porém, de acordo com Ehlers (2005), esta escolha de priori pode trazer algumas dificuldades:

- i) Se o intervalo de variação de θ for ilimitado então a distribuição é imprópria:

$$\int P(\theta)d\theta = \infty$$

ii) Se $\varphi = g(\theta)$ é uma reparametrização não linear monótona de θ então $p(\varphi)$ é não uniforme uma vez que, pelo teorema de transformação de variáveis:

$$P(\varphi) = P(\theta(\varphi)) \left| \frac{d\theta}{d\varphi} \right| \propto \left| \frac{d\theta}{d\varphi} \right|$$

Uma classe de prioris não informativas proposta por Jeffreys (1961) é invariante a transformações 1 a 1, embora algumas vezes seja imprópria (Gelman et al., 2000). Portanto, de forma geral, esta classe de prioris apresenta uma maior amplitude de aplicação que a distribuição uniforme, e esta característica permite sua maior utilização em análise Bayesiana quando se deseja representar um conhecimento mínimo a respeito dos parâmetros a serem estimados (Sorensen & Gianola, 2002).

Freqüentemente, pesquisadores que utilizam a inferência Bayesiana optam pela utilização de prioris vagas quando não dispõem de informações a respeito dos parâmetros a serem estimados, porém esta conduta pode ser problemática para a análise quando a quantidade de observações amostrais é pequena (Lambert et al., 2005). Estes mesmos autores realizaram um estudo de simulação de dados para avaliar a qualidade de prioris não informativas considerando diferentes cenários, sendo as comparações efetuadas por meio de critérios freqüentistas tais como o viés e probabilidade de cobertura. Os resultados mostraram que a priori de Jeffreys mostrou-se mais robusta em relação às inferências a posteriori.

Em se tratando de modelos auto-regressivos, vários autores (Sáfadi & Morettin, 2003; Zellner, 1996; Vermaak et al., 2000) utilizaram a distribuição normal como priori para os parâmetros de auto-regressão, respeitando é claro, o espaço paramétrico relacionado com a ordem do modelo. Porém em recentes estudos (Barreto & Andrade, 2004; Falk & Roy, 2005; Zhou & Roy, 2006) a distribuição t-Student tem sido empregada com sucesso na descrição de coeficientes de modelos auto-regressivos complexos, e segundo

estes autores, se deve ao fato desta distribuição apresentar uma cauda mais pesada, a qual lhe confere uma maior capacidade em expressar o conhecimento a priori em relação aos parâmetros de interesse.

Ghosh & Heo (2003) utilizaram simulação de dados, com ênfase na probabilidade de cobertura dos intervalos de credibilidade, para comparar vários tipos de prioris para um modelo auto-regressivo de primeira ordem, AR(1), e concluíram que certas prioris informativas apresentaram resultados melhores que a priori de Jeffreys. Ni & Sun (2003), estudando modelos auto-regressivos multivariados, compararam diferentes prioris para o vetor de coeficientes. Esta comparação foi realizada com base na estabilização dos momentos da distribuição a posteriori, e indicou que a priori de referência proposta por Yang & Berger (1994) apresentou melhor resultado.

Barreto & Andrade (2004), analisando dados temporais de vazão de hidroelétricas, compararam as prioris hierárquicas Normal-Gama e t-Gama quanto à robustez, e os resultados mostraram uma superioridade dessa última. Estes autores utilizaram o fator de Bayes (Kass & Raftery, 1995) como indicador de qualidade no processo de comparação das prioris.

2.8 Comparação de modelos

A seleção de modelos é uma questão de grande relevância sobretudo quando passamos a considerar modelos mais complexos e os procedimentos de seleção usuais deixam a desejar. De acordo com Spiegelhalter et al. (2002) a estatística clássica só compara modelos encaixados, e a abordagem Bayesiana, mediante fator de Bayes, não é eficiente no caso de prioris impróprias. Nos dois casos a seleção está associada a critérios derivados da razão entre as verossimilhanças, sob o ponto de vista frequentista calculando esta razão no

ponto de máximo, e sob o Bayesiano tomando a razão entre as “médias” das densidades a posteriori.

Em se tratando de comparação de modelos auto-regressivos, observa-se que o objetivo é determinar o valor mais adequado de p , ou seja, deve-se obter a ordem do modelo. Uma das formas por meio do enfoque freqüentista é determinar a significância dos coeficientes, optando sempre que possível por um modelo mais simples. Este fato é definido por Box et al. (1994) como princípio da parcimônia. Esta idéia também pode ser adaptada para o ponto de vista Bayesiano, de forma que esta significância pode ser apresentada por meio dos intervalos de credibilidade para os parâmetros (Milani, 2000).

De acordo com Morettin & Tolo (2004) outros avaliadores de comparação de modelos utilizados em estudos de séries temporais podem ser o Critério de Informação de Akaike (Akaike, 1974) e o Critério de Informação Bayesiano de Schwarz (Schwarz, 1978). Estes critérios apresentam como característica comum um fator de ponderação relacionado com a complexidade dos modelos, e realça a idéia de que o melhor modelo nem sempre é o mais parametrizado.

Ao se utilizar diferentes prioris para o mesmo parâmetro de um modelo, tem-se, sob o ponto de vista Bayesiano, diferentes modelos, que por sua vez podem ser comparados. Esta comparação muitas vezes é realizada por meio do fator de Bayes (Kass & Raftery, 1995), o qual de forma geral corresponde a uma razão de posterioris, seguindo um processo análogo ao de razão de verossimilhanças sob o enfoque freqüentista.

De acordo com Kass & Raftery (1995), o fator de Bayes (FB) é definido por: $FB_{ij} = \frac{P(Y|M_i)}{P(Y|M_j)}$, em que: $P(Y|M_p)$ é o fator de normalização, e

corresponde ao denominador da expressão (3). Esta quantidade recebe este nome, pois é ela que garante a característica de densidade de probabilidade à

distribuição a posteriori, impedindo que esta distribuição seja considerada imprópria. Na literatura (Kass & Raftery, 1995 e Raftery, 1995) esta quantidade também é geralmente denominada de Verossimilhança Marginal, e pode ser obtida da seguinte forma:

$$P(Y|M_p) = \int_{\theta} L(Y|\theta, M_p)P(\theta|M_p)d\theta \quad (5)$$

em que: $L(Y|\theta, M_p)$ e $P(\theta|M_p)$ são, respectivamente, a função de verossimilhança e a distribuição a priori correspondentes ao modelo M_p .

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 $P(Y|M_p)$, a qual é obtida por: $\hat{P}(Y|M_p) = \frac{1}{Q} \sum_{q=1}^Q L(Y|\theta^{(q)}, M_p)$,

em que o índice q representa cada iteração dos algoritmos MCMC, sendo $q = 1, 2, \dots, Q$.

Em relação à interpretação do FB, pode-se dizer que esta é simples e direta, pois se $FB > 1$, tem-se a indicação que o modelo disposto no numerador é melhor, caso contrário, se $FB < 1$, o modelo do denominador é o preferido. Se $FB=1$, a qualidade dos dois modelos é a mesma.

Segundo Kass & Raftery (1995), quando se tem interesse em comparar modelos cujas prioris designadas para os parâmetros são impróprias, como o caso da priori de Jeffreys, o Fator de Bayes não é indicado, pois a adoção de tais prioris conduz a estimativas de Verossimilhança Marginal extremamente grandes, isto é $P(Y|M_p) \rightarrow \infty$, o que proporciona valores irreais para o Fator de Bayes.

Como alternativa, a literatura apresenta dentre outros critérios de comparação, o “Deviance Information Criteria” - DIC (Spiegelhalter et al.,

2002) e o Pseudo-Fator de Bayes (Gelfand, 1996). Este último é de aplicação mais prática, pois é calculado por meio de distribuições preditivas, as quais podem ser obtidas mediante valores gerados a cada iteração dos métodos MCMC.

Em relação ao critério Pseudo-Fator de Bayes, Silva (2006) descreve que o mesmo se aplica muito bem à comparação de modelos de Séries Temporais, pois a obtenção da distribuição preditiva, requerida por este critério, já está naturalmente implícita nos interesses dos pesquisadores.

2.9 Distribuições Preditivas

O objetivo central da análise Bayesiana de séries temporais é gerar modelos robustos de previsão. Uma observação futura é descrita por uma distribuição condicional aos dados passados, denominada distribuição preditiva (Migon & Harrison, 1985). Quando se dispõe da distribuição conjunta das previsões futuras pode-se facilmente obter a distribuição de funções desses valores futuros. Por exemplo, poderíamos estar interessados na soma das previsões da produção industrial ao longo do ano de 2006, baseada em dados mensais até dezembro de 2005.

Sob o ponto de vista freqüentista, os intervalos de previsão são particularmente sensíveis a outliers, ou seja, observações discrepantes, pois estas observações provocam um enviesamento das estimativas dos parâmetros do modelo. Contudo, as previsões não são substancialmente afetadas se os outliers não ocorrerem próximos da origem da previsão (West & Harrison, 1997).

A operação de predição é a própria razão do sistema de análise Bayesiana de modelos de séries temporais, pois fornece a distribuição de uma observação futura com base nos dados passados (Harrison & Stevens, 1976). Essa distribuição denomina-se distribuição preditiva e é obtida integrando-se sobre os parâmetros a distribuição conjunta da observação futura e dos parâmetros, condicionadas aos dados passados.

De acordo com West et al. (1998) uma forma prática de se realizar previsões sob o enfoque Bayesiano em modelos de séries temporais é utilizar o software BATS (Bayesian Analysis of Time Series) o qual produz previsões passos à frente a partir de qualquer ponto, usando a priori corrente, ou especificada pelo usuário. Isto pode ser usado também para verificar as implicações de uma priori escolhida pelo usuário, antes de se ajustar um modelo. Um empecilho para utilização deste programa é a não abordagem de dados em estruturas de painel e de funções de verossimilhanças exatas.

Barreto & Andrade (2004) utilizaram a teoria de distribuições preditivas apresentadas por de Alba (1993) para realizar previsões de vazão de hidrelétricas por meio de um modelo auto-regressivo, AR(p), e concluíram que este enfoque Bayesiano apresentou melhor desempenho se comparado com intervalos de previsão fundamentados no método da Máxima Verossimilhança. Segundo de Alba (1993) um vetor de k futuros valores a serem preditos mediante um modelo auto-regressivo de ordem p, apresentado pela expressão 1, representados por $Y_f = [Y_{n+1}, Y_{n+2}, \dots, Y_{n+k}]'$, seguem uma distribuição Normal multivariada com média dada por $E(Y_f | \Phi, \tau, Y) = -A^{-1}(\Phi)B(\Phi)Y$ e matriz de covariância por $V(Y_f | \Phi, \tau, Y) = \text{diag} \left\{ \tau^{-1} [A^{-1}(\Phi)] [A^{-1}(\Phi)]' \right\}$, em que: $\Phi = [\phi_1, \phi_2, \dots, \phi_p]'$, $\tau = 1/\sigma_e^2$ e as matrizes $A(\Phi)$ e $B(\Phi)$ definidas por:

$$A(\Phi) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi_1 & 1 & 0 & \dots & 0 \\ -\phi_2 & -\phi_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\phi_{k-1} & -\phi_{k-2} & -\phi_{k-3} & \dots & 1 \end{bmatrix}_{k \times k}$$

$$B(\Phi) = \begin{bmatrix} -\phi_p & -\phi_{p-1} & -\phi_{p-2} & \dots & -\phi_1 \\ 0 & -\phi_p & -\phi_{p-1} & \dots & -\phi_2 \\ 0 & 0 & -\phi_p & \dots & -\phi_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -\phi_{k-3} & \dots & -\phi_k \end{bmatrix}_{k \times p}$$

Barreto & Andrade (2004) estenderam as definições teóricas apresentadas por de Alba (1993) de forma a considerar algoritmos MCMC na estimação de $E(Y_f | \Phi, \tau, Y)$ e $V(Y_f | \Phi, \tau, Y)$. Os estimadores obtidos são dados por:

$$\hat{E}(Y_f | \Phi, \tau, Y) = \frac{1}{Q} \sum_{q=1}^Q [-A^{-1}(\Phi^{(q)})B(\Phi^{(q)})Y],$$

$$\hat{V}(Y_f | \Phi, \tau, Y) = \text{diag} \left\{ \frac{1}{Q} \sum_{q=1}^Q \left[\sigma_e^{2(q)} [A^{-1}(\Phi^{(q)})][A^{-1}(\Phi^{(q)})] \right] \right\} +$$

$$\frac{1}{Q} \sum_{q=1}^Q [A^{-1}(\Phi^{(q)})B(\Phi^{(q)})Y]^2 - \left[\frac{1}{Q} \sum_{q=1}^Q [A^{-1}(\Phi^{(q)})B(\Phi^{(q)})Y] \right]^2$$

Segundo Heckman & Leamer (2001), uma forma mais prática, e mais eficiente em relação aos desenvolvimentos teóricos de de Alba (1993), consiste em considerar a distribuição preditiva como sendo uma distribuição condicional a posteriori, ou seja, seus valores são atualizados simultaneamente com a geração dos demais parâmetros do modelo via metodologia MCMC.

2.10. Dados de avaliações genéticas de touros da raça Nelore

O resultado final de um processo de avaliação genética de touros por meio de modelos mistos (Henderson, 1984) caracteriza-se pela obtenção dos valores genéticos de cada touro. A metade destes valores genéticos também é denominada de DEP (Diferença Esperada na Progenie) e significa o acréscimo ou decréscimo que podemos esperar em determinada característica quando usamos o reprodutor em questão.

Estimada com base nas informações do próprio indivíduo e/ou de seus parentes, conforme Brinks (1990), a DEP é a diferença esperada na média das performances das progênies futuras de determinado touro em relação à média das diferenças esperadas das progênies futuras de todos os touros que participaram da mesma avaliação, considerando em ambos os casos acasalamentos com conjunto de vacas que tenham, entre si, o mesmo potencial genético. De acordo com Pereira (1996) a DEP é expressa na unidade de medida da característica em questão, de maneira a possibilitar comparações relativas entre os touros. Como exemplo, considerando a característica ganho de peso médio diário, este autor comenta que se um dado touro X tem uma DEP de + 30 g/dia e um touro Y tem DEP de + 60 g/dia, isto significa que espera-se que a média das performances das progênies do touro Y seja 30 g/dia superior à média das performances das progênies do touro X. A DEP de um animal, sendo uma estimativa, não é um valor estático, isto é, pode mudar em função da variação do número de informações tomadas em qualquer de seus parentes, ou seja, a cada nova avaliação genética, um touro pode ter sua DEP alterada.

Atualmente uma importante característica de bovinos de corte, cujas DEPs vêm sendo calculadas, é o Ganho de Peso ao Sobreano (apresentado em kg). Esta característica expressa o potencial de ganho em peso no período compreendido entre a desmama (205 dias) e os 18 meses (550 dias) e, portanto, refere-se ao ganho de peso em 345 dias após a desmama. As DEPs geralmente são apresentadas em tabelas anuais denominadas de sumários, dessa forma, um mesmo touro pode assumir diferentes valores de DEPs com o decorrer dos anos (Brinks, 1990), caracterizando-se assim um comportamento longitudinal típico de uma série temporal.

Diante de um conjunto de touros que apresentam características comuns, ou seja, que possam ser considerados indivíduos de uma mesma população, e ao se considerar suas respectivas DEPs computadas no decorrer dos anos, tem-se

uma estrutura de dados em painel. Assim, torna-se interessante realizar previsões de DEPs para os próximos anos visando a obtenção de informações prévias que podem auxiliar o melhorista na elaboração de programas de seleção.

Ao se promover seleção, além de ser importante ter-se uma idéia do que se espera com o processo, ou seja, com a predição dos valores genéticos, é importante também obter uma estimativa da segurança com que estes valores estão sendo preditos.

Esta segurança geralmente é definida pela acurácia da seleção, a qual depende da constatação de uma boa relação entre o fenótipo e o genótipo, denominada de herdabilidade. Desta forma, para herdabilidades altas, a seleção pelo fenótipo do indivíduo possibilita a identificação de alta acurácia no cálculo dos valores genéticos. Para herdabilidades baixas, no entanto, maiores erros serão cometidos, observando baixa acurácia. Quanto maior o número de informações relacionadas ao pedigree dos animais, mais precisos serão os cálculos das herdabilidades, e conseqüentemente das acurácias. A cada novo sumário constata-se variações na acurácia, pois a quantidade de informação aumenta devido a propagação das progênies dos animais considerados.

Uma vez que em um sumário tem-se indivíduos cujas DEPs foram obtidas sob diferentes graus de acurácia, é necessário que os mesmos sejam agrupados em classes homogêneas, denominadas de populações, para que a caracterização de dados em estrutura de painel seja evidente.

De acordo com Ministério da Agricultura (Brasil-CNPGC, 1998) os sumários referentes à avaliação nacional de touros da raça Nelore, independentemente de qual seja a empresa responsável pelo programa de avaliação, são documentos originados do protocolo de intenções assinado pelo Ministério da Agricultura, portanto verifica-se uma idoneidade fundamentada em aspectos oficiais estabelecidos a nível nacional.

Em relação aos estudos desenvolvidos com o objetivo de realizar análises longitudinais, ou temporais, de avaliações genéticas, pode-se citar o trabalho de Norman et al. (2003), o qual utilizou dados médios de DEPs de várias características leiteiras, compreendidos entre 1960 e 1998, referentes às raças Ayrshire, Brown Swiss, Guernsey, Holstein e Jersey. Como resultados relevantes os autores estimaram tendências de aumento das DEPs e comprovaram a eficiência da utilização da inseminação artificial, além de fazerem predições, reportadas em porcentagens, de alterações das DEPs para anos futuros. Powell et al. (2004) avaliaram uma série histórica de Deps, também denominada de PTA (“Predicted Transmitting Ability”), para produção de leite no período de janeiro de 1995 a agosto de 2003 (Figura 1). Estes autores utilizaram dados médios em cada período avaliado, e as predições foram feitas por meio da estimação dos desvios-padrão esperados para alterações (“Expected Standard Deviations of Change”), apresentados por Powell & Norman (1981). Os resultados obtidos mostraram que o declínio da série em determinados períodos se deve a erros de seleção, ou seja, animais que tiveram seus valores genéticos super ou subestimados, e que os valores preditos para as últimas avaliações indicam que realmente os valores tendem a se estabilizarem.

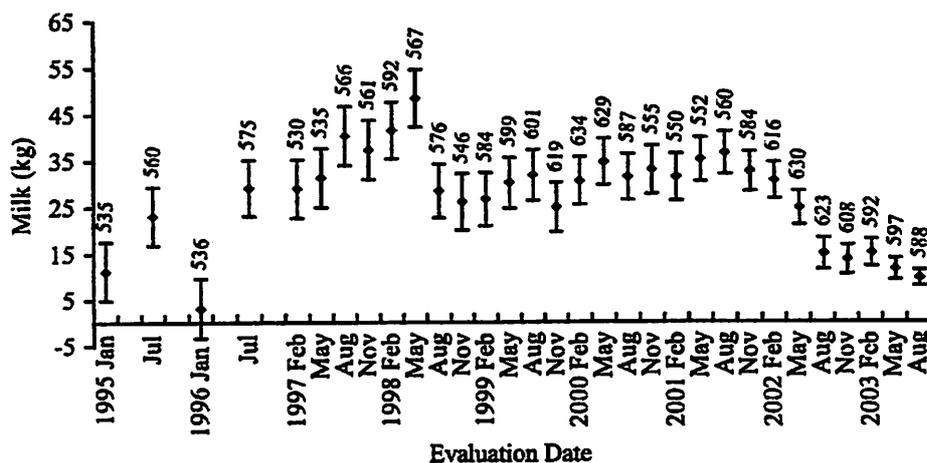


Figura 1. Série original estudada por Powell et al. (2004).

MacNeil (2005) avaliou séries temporais representadas por valores genéticos médios de touros, calculados durante os anos de 1989 e 2000 para a característica razão entre peso do bezerro aos 200 dias e peso adulto da mãe. Estudaram séries provenientes de dois tipos de modelos mistos (considerando efeito materno e considerando efeito direto) usados para a obtenção destes valores, e consideraram dois grupos de touros (provenientes de seleção para a característica estudada e um grupo controle). O autor observou a estimação de uma tendência linear positiva apenas para o grupo formado por touros selecionados com a adoção do modelo que considera o efeito materno, sendo que para as demais séries foi observado um processo estacionário. As séries estudadas são apresentadas na Figura 2.

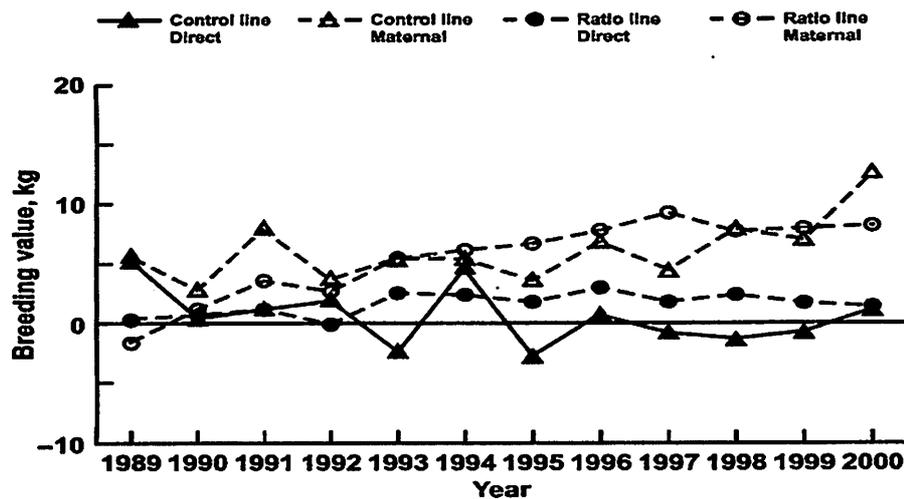


Figura 2. Série original estudada por MacNeil (2005).

3 MATERIAL E MÉTODOS

De acordo com a metodologia Bayesiana, para que seja possível realizar inferências fundamentadas em um determinado modelo é necessário obter a função de verossimilhança e especificar as distribuições a priori para os parâmetros de interesse.

3.1 Função de verossimilhança exata

No que se segue, será considerado $n = n_1 = n_2 = \dots = n_m$, ou seja, o mesmo número de observações longitudinais para cada indivíduo $i, i=1,2,\dots,m$.

Respeitando o modelo auto-regressivo de ordem p para dados em painel, o qual já foi apresentado na expressão (2), tem-se, para um único indivíduo, a seguinte função de verossimilhança exata:

$$L_i(Y_i | \phi, \sigma_e^2) \propto \Psi_i(\phi, \sigma_e^2 | Y_{ip}) \sigma_e^{-(\frac{n-p}{2})} \exp\left\{-\frac{1}{2\sigma_e^2} \sum_{t=p+1}^n e_{it}^2\right\}$$

$$L_i(Y_i | \phi, \sigma_e^2) \propto \Psi_i(\phi, \sigma_e^2 | Y_{ip}) \sigma_e^{-(\frac{n-p}{2})} \exp\left\{-\frac{1}{2\sigma_e^2} \sum_{t=p+1}^n (Y_{it} - \sum_{j=1}^p \phi_{ij} Y_{i(t-j)})^2\right\},$$

em que: $\Psi_i(\phi, \sigma_e^2 | Y_{ip}) \propto \sigma_e^{-(\frac{p}{2})} |V_{ip}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_e^2} Y_{ip}' V_{ip} Y_{ip}\right\},$

$$Y_i = [y_{i1}, y_{i2}, \dots, y_{ip}, y_{i(p+1)}, y_{i(p+2)}, \dots, y_{in}]'_{n \times 1} \quad \text{e} \quad Y_{ip} = [y_{i1}, y_{i2}, \dots, y_{ip}]'_{p \times 1}$$

O componente $\Psi_i(\phi, \sigma_e^2 | Y_{ip})$ representa a inserção das p observações que foram desprezadas devido à defasagem implícita no modelo auto-regressivo. Mediante a especificação deste componente, é que se caracteriza uma função de verossimilhança como exata ou incondicional. A matriz V_{ip} pode ser obtida

diretamente pelo método de Yule-Walker (Box et al., 1994). Como exemplo, para modelos auto-regressivos de primeira, AR(1), e segunda, AR(2), ordens, têm-se, respectivamente, as seguintes representações de V_{ip} :

$$V_{ip} = 1 - \phi_{i1}^2$$

$$V_{ip} = \begin{bmatrix} 1 - \phi_{i2}^2 & -\phi_{i1}(1 + \phi_{i2}) \\ -\phi_{i1}(1 + \phi_{i2}) & 1 - \phi_{i2}^2 \end{bmatrix}$$

Se reunirmos as observações referentes a todos os indivíduos, supostas independentes, ou seja, utilizando a combinação de informações individuais relacionada com análise de dados em estrutura de painel, pode-se reescrever a função de verossimilhança como:

$$L(Y | \Phi, \sigma_\epsilon^2) = \prod_{i=1}^m \Psi_i(\Phi, \sigma_\epsilon^2 | Y_p) \sigma_\epsilon^{2 \left(\frac{n-p}{2} \right)} \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \sum_{t=p+1}^n e_{it}^2 \right\},$$

$$L(Y | \Phi, \sigma_\epsilon^2) = \Psi(\Phi, \sigma_\epsilon^2 | Y_p) \sigma_\epsilon^{2 \left(\frac{m(n-p)}{2} \right)} \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^m \sum_{t=p+1}^n (Y_{it} - \sum_{j=1}^p \phi_{ij} Y_{i(t-j)})^2 \right\},$$

em que: $\Psi(\Phi, \sigma_\epsilon^2 | Y_p) = \sigma_\epsilon^{2 \left(\frac{mp}{2} \right)} |V_p|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} Y_p' V_p Y_p \right\}$

$$Y_p = [y_{11}, y_{12}, \dots, y_{1p}, y_{21}, y_{22}, \dots, y_{2p}, \dots, y_{m1}, y_{m2}, \dots, y_{mp}]'$$

A matriz V_p assume agora, sob esta condição, uma estrutura bloco diagonal, a qual é capaz de contemplar todos possíveis valores de V_{ip} . Esta estrutura é dada, respectivamente para modelos AR(1) e AR(2), pela seguinte expressão:

$$V_p = \begin{bmatrix} 1 - \phi_{i1}^2 & 0 & 0 & 0 \\ 0 & 1 - \phi_{i2}^2 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 1 - \phi_{m1}^2 \end{bmatrix}_{m \times m}$$

$$V_p = \begin{bmatrix} 1 - \phi_{12}^2 & -\phi_{11}(1 + \phi_{12}) & 0 & 0 & 0 & 0 & 0 \\ -\phi_{11}(1 + \phi_{12}) & 1 - \phi_{12}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 - \phi_{22}^2 & -\phi_{21}(1 + \phi_{22}) & 0 & 0 & 0 \\ 0 & 0 & -\phi_{21}(1 + \phi_{22}) & 1 - \phi_{22}^2 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 - \phi_{m2}^2 & -\phi_{m1}(1 + \phi_{m2}) \\ 0 & 0 & 0 & 0 & 0 & -\phi_{m1}(1 + \phi_{m2}) & 1 - \phi_{m2}^2 \end{bmatrix}$$

Reescrevendo a função de verossimilhança para todos os indivíduos, em forma matricial, têm-se:

$$L(Y | \Phi, \sigma_e^2) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \sigma_e^{2 \left(\frac{m(n-p)}{2} \right)} \exp \left\{ -\frac{1}{2\sigma_e^2} (Y_1 - X\Phi)' (Y_1 - X\Phi) \right\}$$

em que:

$$Y_1 = [y_{1p+1}, y_{1p+2}, \dots, y_{1n}, y_{2p+1}, y_{2p+2}, \dots, y_{2n}, \dots, y_{mp+1}, y_{mp+2}, \dots, y_{mn}]'$$

$$X = \begin{bmatrix} X_1 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & X_m \end{bmatrix}_{m(n-p) \times mp} \quad X_i = \begin{bmatrix} y_{ip} & \dots & y_{i1} \\ y_{ip+1} & \dots & y_{i2} \\ \vdots & \vdots & \vdots \\ y_{in-1} & \dots & y_{in-p} \end{bmatrix}_{(n-p) \times p}$$

$$\Phi = [\phi_{11}, \phi_{12}, \dots, \phi_{1p}, \phi_{21}, \phi_{22}, \dots, \phi_{2p}, \dots, \phi_{m1}, \phi_{m2}, \dots, \phi_{mp}]' \in \mathbb{R}^{mp}$$

Os vetores Y_1 e Φ , considerados nas expressões, apresentam, respectivamente, as dimensões $m(n-p) \times 1$ e $mp \times 1$.

3.2 Distribuições a priori

Neste estudo optou-se por comparar três diferentes distribuições a priori para os parâmetros de interesse, Φ e σ_e^2 . Desta forma foram consideradas duas prioris informativas (priori hierárquica Normal multivariada - Gama Inversa e priori independente t-Student multivariada – Gama Inversa) e uma não informativa (priori de Jeffreys).

3.2.1 Priori hierárquica Normal multivariada - Gama Inversa

Esta distribuição a priori é expressa da seguinte forma:

$$P(\Phi, \sigma_e^2) = P(\Phi | \sigma_e^2) P(\sigma_e^2)$$

em que: $(\Phi | \sigma_e^2) \sim N(\mu, \sigma_e^{-2} P)$ e $\sigma_e^2 \sim GI(\alpha, \beta)$ (Gama Inversa) .

Descrevendo implicitamente estas prioris tem-se:

$$P(\Phi | \sigma_e^2) \propto \sigma_e^{-\left(\frac{mp}{2}\right)} \exp\left\{-\frac{1}{2\sigma_e^2} [(\Phi - \mu)' P^{-1} (\Phi - \mu)]\right\},$$

$$P(\sigma_e^2) \propto (\sigma_e^2)^{-(\alpha+1)} \exp\left\{-\frac{\beta}{\sigma_e^2}\right\}.$$

Portanto, distribuição conjunta das prioris, $P(\Phi, \sigma_e^2)$, é dada por:

$$P(\Phi, \sigma_e^2) \propto \sigma_e^{-\left(\frac{mp+2\alpha}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma_e^2} [2\beta + (\Phi - \mu)' P^{-1} (\Phi - \mu)]\right\}.$$

Os componentes μ, P, α e β , são denominados hiperparâmetros, e representam os valores assumidos pelos parâmetros das distribuições a priori.

3.2.2 Priori independente t-Student multivariada - Gama Inversa

Neste caso optou-se pela utilização de prioris independentes, sendo considerada uma distribuição t-Student multivariada para o parâmetro Φ e uma distribuição Gama-Inversa para σ_e^2 . Então, tem-se:

$$P(\Phi, \sigma_e^2) = P(\Phi)P(\sigma_e^2),$$

$\Phi \sim t\text{-Student}(\mu, P)$ com v graus de liberdade e $\sigma_e^2 \sim GI(\alpha, \beta)$, ou seja:

$$P(\Phi) \propto \left[1 + (\Phi - \mu)' P^{-1} (\Phi - \mu) \right]^{-\frac{v+mp}{2}} \text{ e } P(\sigma_e^2) \propto (\sigma_e^2)^{-(\alpha+1)} \exp\left\{-\frac{\beta}{\sigma_e^2}\right\}.$$

Assim, tem-se a seguinte distribuição conjunta:

$$P(\Phi, \sigma_e^2) \propto \left[1 + (\Phi - \mu)' P^{-1} (\Phi - \mu) \right]^{-\frac{v+mp}{2}} (\sigma_e^2)^{-(\alpha+1)} \exp\left\{-\frac{\beta}{\sigma_e^2}\right\}.$$

3.2.3 Priori não informativa de Jeffreys

Esta priori foi utilizada para representar a situação em que o pesquisador não tem uma informação definida a respeito dos parâmetros a serem estimados.

Utilizou-se a abordagem descrita por Broemiling & Cook (1993) na descrição da priori de Jeffreys para modelos auto-regressivos. Esta descrição está fundamentada na matriz de informação de Fisher, e de acordo com o desenvolvimento destes autores, pode ser apresentada pela seguinte expressão:

$$P(\Phi, \sigma_e^2) \propto \frac{1}{\sigma_e^2}$$

3.3 Distribuições conjuntas a posteriori

De acordo com a teoria Bayesiana, para cada distribuição a priori especificada deve-se obter uma distribuição a posteriori, considerando no presente estudo a mesma função de verossimilhança.

3.3.1 Distribuição conjunta a posteriori (Priori hierárquica Normal multivariada - Gama Inversa)

Partindo do princípio, $Posteriori \propto Verossimilhança \times Priori$, pode-se apresentar a seguinte relação:

$$P(\Phi, \sigma_e^2 | Y) \propto L(Y | \Phi, \sigma_e^2) P(\Phi, \sigma_e^2),$$

$$P(\Phi, \sigma_e^2 | Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \sigma_e^2^{-\frac{m(n-p)}{2}} \exp\left\{-\frac{1}{2\sigma_e^2} (Y_1 - X\Phi)' (Y_1 - X\Phi)\right\} \\ \times \sigma_e^2^{-\left(\frac{mp+2\alpha}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma_e^2} [2\beta + (\Phi - \mu)' P^{-1} (\Phi - \mu)]\right\},$$

$$P(\Phi, \sigma_e^2 | Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \sigma_e^2^{-\left(\frac{mn+2\alpha}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma_e^2} [2D + (\Phi - \hat{\Phi}_B)' \Sigma^{-1} (\Phi - \hat{\Phi}_B)]\right\},$$

em que:

$$D = \beta + \frac{(Y_1' Y_1 + \mu' P^{-1} \mu) - (X' Y_1 + P^{-1} \mu)' (X' X + P^{-1})^{-1} (X' Y_1 + P^{-1} \mu)}{2},$$

$$\hat{\Phi}_B = (X' X + P^{-1})^{-1} (X' Y_1 + P^{-1} \mu) \quad \text{e} \quad \Sigma = X' X + P^{-1}.$$

3.3.2 Distribuição conjunta a posteriori (Priori independente t-Student multivariada - Gama Inversa)

Tendo em vista a mesma idéia apresentada no item anterior, tem-se agora a seguinte distribuição conjunta a posteriori:

$$P(\Phi, \sigma_e^2 | Y) \propto L(Y | \Phi, \sigma_e^2) P(\Phi, \sigma_e^2)$$

$$P(\Phi, \sigma_e^2 | Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \sigma_e^{-2 \left[\frac{m(n-p)}{2} \right]} \exp \left\{ -\frac{1}{2\sigma_e^2} (Y_1 - X\Phi)' (Y_1 - X\Phi) \right\} \\ \times \left[1 + (\Phi - \mu)' P^{-1} (\Phi - \mu) \right]^{\left(\frac{\nu+mp}{2} \right)} (\sigma_e^2)^{-(\alpha+1)} \exp \left\{ -\frac{\beta}{\sigma_e^2} \right\},$$

$$P(\Phi, \sigma_e^2 | Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \sigma_e^{-2 \left[\frac{m(n-p)+2\alpha+1}{2} \right]} \\ \times \exp \left\{ -\frac{1}{\sigma_e^2} \left[\left(\frac{(Y_1 - X\Phi)' (Y_1 - X\Phi)}{2} \right) + \beta \right] \right\} \\ \times \left[1 + (\Phi - \mu)' P^{-1} (\Phi - \mu) \right]^{\left(\frac{\nu+p}{2} \right)},$$

$$P(\Phi, \sigma_e^2 | Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \sigma_e^{-2 \left[\frac{m(n-p)+2\alpha+1}{2} \right]} \\ \times \exp \left\{ -\frac{1}{\sigma_e^2} \left[\left(\frac{(\Phi - \hat{\Phi})' (X'X)(\Phi - \hat{\Phi}) + (Y_1 - \hat{Y}_1)' (Y_1 - \hat{Y}_1)}{2} \right) + \beta \right] \right\} \\ \times \left[1 + (\Phi - \mu)' P^{-1} (\Phi - \mu) \right]^{\left(\frac{\nu+p}{2} \right)},$$

em que:

$$\hat{\Phi} = (X'X)^{-1} (X'Y_1) \quad \text{e} \quad \hat{Y}_1 = X\hat{\Phi} = X(X'X)^{-1} (X'Y_1).$$

3.3.2 Distribuição conjunta a posteriori (Priori não informativa de Jeffreys)

De acordo com o teorema de Bayes, é possível expressar esta distribuição como sendo:

$$P(\Phi, \sigma_e^2 | Y) \propto L(Y | \Phi, \sigma_e^2) P(\Phi, \sigma_e^2),$$

$$P(\Phi, \sigma_e^2 | Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) (\sigma_e^2)^{-\left(\frac{m(n-p)}{2}\right)} \exp\left\{-\frac{1}{2\sigma_e^2} (Y_1 - X\Phi)' (Y_1 - X\Phi)\right\} \times \frac{1}{\sigma_e^2},$$

$$P(\Phi, \sigma_e^2 | Y) \propto \sigma_e^2^{-\left(\frac{mn}{2}+1\right)} |V_p|^{-1/2} \exp\left\{-\frac{1}{2\sigma_e^2} [Y_p' V_p^{-1} Y_p + \right. \\ \left. + (\Phi - (X'X)^{-1} X'Y_1)' (X'X) (\Phi - (X'X)^{-1} X'Y_1) + \right. \\ \left. + (X'Y_1)' (X'X)^{-1} X'Y_1 + Y_1'Y_1]\right\}.$$

3.4 Distribuições condicionais completas a posteriori

Conforme elucidado anteriormente, para que se obtenham as distribuições marginais a posteriori, é necessário realizar integrações na distribuição conjunta a posteriori. Neste estudo verificou-se que estas integrais não apresentam soluções analíticas, e por isso serão utilizados algoritmos MCMC visando a obtenção de amostras dessas distribuições marginais. Portanto, é necessário apresentar as distribuições condicionais completas a posteriori, para que estes algoritmos possam ser implementados.

3.4.1 Distribuições condicionais completas a posteriori (Priori hierárquica Normal multivariada - Gama Inversa)

Uma vez obtida a distribuição completa a posteriori para a priori em questão, pode-se apresentar as seguintes distribuições condicionais completas:

$$P(\Phi | \sigma_e^2, Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \exp \left\{ -\frac{1}{2\sigma_e^2} [(\Phi - \hat{\Phi}_B)' \Sigma (\Phi - \hat{\Phi}_B)] \right\}, \text{ou seja,}$$

$$P(\Phi | \sigma_e^2, Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \text{Normal Multivariada}(\hat{\Phi}_B, \sigma_e^{-2} \Sigma^{-1}),$$

$$P(\sigma_e^2 | \Phi, Y) \propto \sigma_e^{-2 \left(\frac{mp}{2} \right)} \exp \left\{ -\frac{1}{2\sigma_e^2} Y_p' V_p Y_p \right\} \times \\ \sigma_e^{-2 \left(\frac{mn+2\alpha}{2} + 1 \right)} \exp \left\{ -\frac{1}{\sigma_e^2} \left[D + \frac{1}{2} (\Phi - \hat{\Phi}_B)' \Sigma (\Phi - \hat{\Phi}_B) \right] \right\},$$

$$P(\sigma_e^2 | \Phi, Y) \propto \sigma_e^{-2 \left(\frac{mp+mn+2\alpha}{2} + 1 \right)} \exp \left\{ -\frac{1}{\sigma_e^2} \left[\frac{1}{2} (Y_p' V_p Y_p) + D + \frac{1}{2} (\Phi - \hat{\Phi}_B)' \Sigma (\Phi - \hat{\Phi}_B) \right] \right\},$$

ou seja:

$$\sigma_e^2 | \Phi, Y \sim \text{Gama Inversa} \left(\frac{mp+mn+2\alpha}{2}, \frac{1}{2} (Y_p' V_p Y_p) + D + \frac{1}{2} (\Phi - \hat{\Phi}_B)' \Sigma (\Phi - \hat{\Phi}_B) \right),$$

ou ainda: $\sigma_e^2 | \Phi, Y \sim \text{Gama Inversa}(\alpha_1, \beta_1)$.

3.4.2 Distribuições condicionais completas a posteriori (Priori independente t-Student multivariada - Gama Inversa)

Ao se utilizar estas prioris têm-se as seguintes distribuições condicionais completas a posteriori:

$$P(\Phi | \sigma_e^2, Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \times \exp \left\{ -\frac{1}{2\sigma_e^2} [(\Phi - \hat{\Phi})' (X' X) (\Phi - \hat{\Phi})] \right\} \\ \times \left[1 + (\Phi - \mu)' P (\Phi - \mu) \right]^{-\frac{v+mp}{2}},$$

$$P(\Phi | \sigma_e^2, Y) \propto \Psi(\Phi, \sigma_e^2 | Y_p) \times \text{Normal multivariada}(\hat{\Phi}, (X'X)^{-1}) \times \\ t - \text{Student multivariada}(\mu, P^{-1})$$

$$P(\sigma_e^2 | \Phi, Y) \propto \sigma_e^{-2 \left[\frac{m(n-p) + mp + 2\alpha + 1}{2} \right]} \exp \left\{ -\frac{1}{\sigma_e^2} \left[\frac{1}{2} (Y_p' V_p Y_p) + \beta + \right. \right. \\ \left. \left. \frac{(\Phi - \hat{\Phi})' (X'X) (\Phi - \hat{\Phi}) + (Y_1 - \hat{Y}_1)' (Y_1 - \hat{Y}_1)}{2} \right] \right\}, \text{ou seja,}$$

$$\sigma_e^2 | \Phi, Y \sim \text{Gama Inversa} \left(\frac{mn + 2\alpha}{2}, \frac{1}{2} (Y_p' V_p Y_p) + \beta + \right. \\ \left. \frac{(\Phi - \hat{\Phi})' (X'X) (\Phi - \hat{\Phi}) + (Y_1 - \hat{Y}_1)' (Y_1 - \hat{Y}_1)}{2} \right)$$

$$\sigma_e^2 | \Phi, Y \sim \text{Gama Inversa}(\alpha_2, \beta_2)$$

3.4.3 Distribuições condicionais completas a posteriori (Priori não informativa de Jeffreys)

Ao se utilizar a priori não informativa de Jeffreys, torna-se possível a obtenção das seguintes distribuições condicionais completas a posteriori:

$$P(\Phi | \sigma_e^2, Y) \propto (\sigma_e^2)^{-\left(\frac{mn}{2} + 1\right)} |V_p|^{-1/2} \exp \left\{ -\frac{1}{2\sigma_e^2} \left[Y_p' V_p Y_p + \right. \right. \\ \left. \left. + (\Phi - (X'X)^{-1} X'Y_1)' (X'X) (\Phi - (X'X)^{-1} X'Y_1) \right] \right\}$$

De acordo com as distribuições condicionais completas para o parâmetro Φ apresentadas nos itens 3.4.1 e 3.4.2, nota-se que as mesmas não se caracterizam como distribuições de probabilidade conhecidas, mas pelo menos parte dessas podem ser identificadas como tais, como, por exemplo, a normal e a t-Student multivariadas, as quais são de extrema relevância quando se pretende utilizar o algoritmo Metropolis-Hastings. Por outro lado, ao se trabalhar com a priori não informativa de Jeffreys esta característica não é verificada, o que nos leva a utilização de um arranjo de forma a verificar a existência de alguma distribuição conhecida. Para tanto, considere a função densidade de probabilidade da distribuição normal multivariada. Seja $\mathbf{x} \sim \text{Normal multivariada}(\mathbf{m}, \mathbf{V})$, então $f(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} |\mathbf{V}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[(\mathbf{x} - \mathbf{m})' \mathbf{V} (\mathbf{x} - \mathbf{m})]\right\}$, em que n é a dimensão da matriz de covariância \mathbf{V} .

$$P(\Phi | \sigma_e^2, Y) \propto |V_p|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_e^2} [Y_p' V_p Y_p]\right\} \times \\ \sigma_e^{-2\left(\frac{m}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma_e^2} [(\Phi - (X'X)^{-1} X'Y_1)' (X'X) (\Phi - (X'X)^{-1} X'Y_1)]\right\},$$

$$P(\Phi | \sigma_e^2, Y) \propto \text{Normal multivariada}(0, V_p^{-1}) \times \\ \sigma_e^{-2\left(\frac{m}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma_e^2} [(\Phi - (X'X)^{-1} X'Y_1)' (X'X) (\Phi - (X'X)^{-1} X'Y_1)]\right\}.$$

Para σ_e^2 tem-se:

$$P(\sigma_e^2 | \Phi, Y) \propto \sigma_e^{-2\left(\frac{m}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma_e^2} [Y_p' V_p Y_p + \right. \\ \left. + (\Phi - (X'X)^{-1} X'Y_1)' (X'X) (\Phi - (X'X)^{-1} X'Y_1) + \right. \\ \left. + (X'Y_1)' (X'X)^{-1} X'Y_1 + Y_1' Y_1]\right\}.$$

$$\sigma_e^2 | \Phi, Y \sim \text{Gama Inversa} \left(\frac{mn}{2}, \frac{1}{2} [Y_p' V_p Y_p + (\Phi - (X'X)^{-1} X'Y_1)'(X'X)(\Phi - (X'X)^{-1} X'Y_1) + (X'Y_1)'(X'X)^{-1} X'Y_1 + Y_1'Y_1] \right), \text{ ou seja,}$$

$$\sigma_e^2 | \Phi, Y \sim \text{Gama Inversa}(\alpha_3, \beta_3).$$

3.5 Implementação dos algoritmos MCMC

Nota-se nos itens 3.4.1, 3.4.2 e 3.4.3 que a distribuição condicional completa para o parâmetro σ_e^2 é dada por uma distribuição Gama-Inversa, ou seja, ela apresenta uma forma conhecida, portanto passível ao uso do algoritmo Gibbs Sampler.

O mesmo não acontece para a distribuição condicional do parâmetro Φ , a qual, para todas as prioris especificadas, não apresenta uma forma definida, devendo-se então utilizar, nesta situação, o algoritmo Metropolis-Hastings.

Os algoritmos Gibbs Sampler e Metropolis-Hastings foram implementados matricialmente no software estatístico R (R Development Core Team, 2006). A função *mnormt* (“multivariate Normal and t-Student distributions”) foi utilizada para a geração de números aleatórios implícita nos algoritmos MCMC.

Estes algoritmos foram implementados, na análise dos dados simulados, considerando uma cadeia de 50.000 iterações, das quais a primeira metade foi eliminada (“burn-in”) para evitar os efeitos dos valores iniciais adotados. O tamanho final das cadeias de cada parâmetro, bem como a especificação do “burn-in” e do intervalo de amostragem foram avaliados por meio do critério de Raftery & Lewis (1992b). A constatação final da convergência foi realizada

conforme o protocolo sugerido por Nogueira (2004). Todos os critérios que compõem este protocolo foram avaliados mediante o pacote *BOA* (“Bayesian Output Analysis”) do software R, e os gráficos confeccionados pelo software livre *Gnuplot*, o qual realiza uma leitura direta dos dados fornecidos pelo software R.

Considerando a análise dos dados reais de DEPs de touros da raça Nelore, optou-se por uma cadeia com 20.000 iterações e um “bur in” de 10.000. Nesta situação, foram utilizados os critérios de Gelman & Rubin (1992) e Raftery & Lewis (1992) também por meio do pacote *BOA*.

3.6 Distribuições preditivas

A distribuição preditiva de um dado futuro, relacionada com cada indivíduo i listado em uma estrutura de dados em painel, é representada por uma distribuição condicional aos dados passados, ou seja, $P(Y_{i(n+1)} | Y_i)$. Esta distribuição é dada por:

$$P(Y_{i(n+1)}, \phi_i, \sigma_e^2 | Y_i) \propto L(Y_{i(n+1)} | \phi_i, \sigma_e^2, Y_i) P(\phi_i, \sigma_e^2 | Y_i),$$

$$P(Y_{i(n+1)} | Y_i) \propto \int \int P(Y_{i(n+1)}, \phi_i, \sigma_e^2 | Y_i) d\phi_i d\sigma_e^2.$$

Portanto, torna-se interessante começar o desenvolvimento pelo termo $L(Y_{i(n+1)} | \phi_i, \sigma_e^2, Y_i)$, que representa a função de verossimilhança para 1 dado futuro. Seja o modelo:

$$Y_{i(n+1)} = \phi_{i1} Y_{in} + \phi_{i2} Y_{i(n-1)} + \phi_{i3} Y_{i(n-2)} + \dots + \phi_{ip} Y_{i(n+1-p)} + e_{i(n+1)},$$

que pode ser reescrito como:

$$Y_{i(n+1)} = \sum_{j=1}^p \phi_{ij} Y_{i(n+1-j)} + e_{i(n+1)}.$$

Assumindo $e_{i(n+1)} \sim N(0, \sigma_e^2)$, a função de verossimilhança, para cada indivíduo i , pode ser especificada como:

$$L(Y_{i(n+1)} | \phi_i, \sigma_e^2, Y_i) \propto (\sigma_e^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} \left[Y_{i(n+1)} - \sum_{j=1}^p \phi_{ij} Y_{i(n+1-j)} \right]^2 \right\}.$$

Generalizando para m indivíduos tem-se:

$$L(Y_{(n+1)} | \Phi, \sigma_e^2, Y) \propto \prod_{i=1}^m L(Y_{i(n+1)} | \phi_i, \sigma_e^2, Y_i),$$

$$L(Y_{(n+1)} | \Phi, \sigma_e^2, Y) \propto (\sigma_e^2)^{-\frac{m}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} \sum_{i=1}^m \left[Y_{i(n+1)} - \sum_{j=1}^p \phi_{ij} Y_{i(n+1-j)} \right]^2 \right\}.$$

Sob forma matricial tem-se:

$$L(Y_{(n+1)} | \Phi, \sigma_e^2, Y) \propto (\sigma_e^2)^{-\frac{m}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} \left[(Y_{(n+1)} - X\Phi)' (Y_{(n+1)} - X\Phi) \right] \right\},$$

sendo:

$$Y_{(n+1)} = \begin{bmatrix} y_{1(n+1)} \\ y_{2(n+1)} \\ \vdots \\ y_{m(n+1)} \end{bmatrix}_{m \times 1}, \quad X = \begin{bmatrix} X_1 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & X_m \end{bmatrix}_{m \times mp}, \quad X_i = [y_{in} \ y_{i(n-1)} \ \dots \ y_{i(n+1-p)}]_{1 \times p}.$$

Portanto, $P(Y_{(n+1)}, \Phi, \sigma_e^2 | Y) \propto L(Y_{(n+1)} | \Phi, \sigma_e^2, Y) P(\Phi, \sigma_e^2 | Y)$, em que

$P(\Phi, \sigma_e^2 | Y)$ representa a distribuição a posteriori. Logo,

$$P(Y_{(n+1)} | Y) \propto \int \int (\sigma_e^2)^{-\frac{m}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} \left[(Y_{(n+1)} - X\Phi)' (Y_{(n+1)} - X\Phi) \right] \right\} P(\Phi, \sigma_e^2 | Y) d\Phi d\sigma_e^2.$$

A integral acima não apresenta solução analítica, porém conforme a teoria descrita por Heckman & Leamer (2001), é possível demonstrar que mediante a utilização da técnica MCMC, tem-se:

$$Y_{(n+1)}^{(q)} | Y \sim N(X\Phi^{(q)}, \sigma_e^{2(q)}I),$$

em que: I é uma matriz de ordem $mp \times mp$.

Assim, o conjunto de valores gerados para esta distribuição Normal multivariada, provenientes de cada q iteração dos algoritmos Metropolis-Hastings e Gibbs Sampler, constituem a distribuição preditiva para um dado futuro, cuja estimativa, $\hat{P}(Y_{(n+1)} | Y)$, é representada pela média desta distribuição. Caso seja de interesse, pode-se generalizar esta metodologia para a predição de k dados futuros, porém para esta implementação é necessário obedecer um processo iterativo, fundamentado na ordem de geração dos valores, ou seja, para gerar a distribuição de $Y_{(n+2)}$, deve-se anteriormente gerar a distribuição de $Y_{(n+1)}$, e assim sucessivamente até a predição $Y_{(n+k)}$.

3.7 Fator de Bayes e Pseudo-Fator de Bayes: comparação de priors

Para comparar os modelos 1 (priori hierárquica Normal multivariada - Gama Inversa) e 2 (priori independente t-Student multivariada - Gama Inversa) utilizou-se o Fator de Bayes (FB) sob o enfoque apresentado por Barreto & Andrade (2004). Este utiliza valores gerados pelos métodos MCMC para obter as estimativas do fator de normalização, $P(Y | M_p)$, também denominado de Verossimilhança Marginal, o qual compõe a expressão do Fator de Bayes:

$$FB_{ij} = \frac{\hat{P}(Y | M_i)}{\hat{P}(Y | M_j)} = \frac{\frac{1}{Q} \sum_{q=1}^Q L(Y | \theta^{(q)}, M_i)}{\frac{1}{Q} \sum_{q=1}^Q L(Y | \theta^{(q)}, M_j)}.$$

Como frisado anteriormente, $\theta^{(q)}$ indica os valores gerados para os parâmetros na q-ésima iteração ($q = 1, 2, \dots, Q$) para cada um dos modelos comparados. Assim, o termo $L(Y | \theta^{(q)}, M_p)$ corresponde a valores da função de verossimilhança obtidos pela substituição dos valores atuais dos parâmetros.

Usando a função de verossimilhança adotada neste estudo, tem-se a seguinte estimativa da Verossimilhança Marginal de um modelo p:

$$\begin{aligned} \hat{P}(Y | M_p) &= \frac{1}{Q} \sum_{q=1}^Q L(Y | \Phi^{(q)}, \sigma_e^{2(q)}, M_p) = \\ &= \frac{1}{Q} \sum_{q=1}^Q \Psi(\Phi^{(q)}, \sigma_e^{2(q)} | Y_p) \sigma_e^{2(q)^{-m} \binom{n-p}{2}} \exp \left\{ -\frac{1}{2\sigma_e^{2(q)}} [Y_1 - X\Phi^{(q)}][Y_1 - X\Phi^{(q)}] \right\} \end{aligned}$$

Em relação a comparação dos modelos 1 e 2 com o modelo 3 (Priori não informativa de Jeffreys), optou-se pela utilização do Pseudo-Fator de Bayes, uma vez que o modelo 3 considera uma priori imprópria. A aplicação deste critério considerou a definição apresentada por Gelfand (1996), que o definem como sendo o produtório de razões de densidades preditivas para k observações entre os modelos i e j, ou seja:

$$PsdFB_{ij} = \frac{\prod_{i=1}^k \hat{P}(Y_{(nk)} | Y, M_i)}{\prod_{i=1}^k \hat{P}(Y_{(nk)} | Y, M_j)}$$

Diferentemente do fator de Bayes, o Pseudo-Fator de Bayes não pode ser analisado pela sua grandeza, pois não é uma verdadeira razão de chances a posteriori, dessa forma, Gelfand (1996) propôs a seguinte quantidade, a qual pode ser avaliada simplesmente pelo seu sinal:

$$PsdFB_{ij} = \ln \left(\frac{\prod_{i=1}^k \hat{P}(Y_{(n+k)} | Y, M_i)}{\prod_{i=1}^k \hat{P}(Y_{(n+k)} | Y, M_j)} \right).$$

Se $PsdFB_{ij} > 0$ seleciona-se o modelo i, caso contrário seleciona-se o modelo j. Também é importante ressaltar que no presente trabalho adotou-se $k=1$, portanto:

$$PsdFB_{ij} = \ln \left(\frac{\hat{P}(Y_{(n+1)} | Y, M_i)}{\hat{P}(Y_{(n+1)} | Y, M_j)} \right),$$

Pode-se notar que a estimativa de uma observação futura corresponde na verdade a um vetor $Y_{(n+1)}$ de observações futuras, uma vez que se trata de dados em estrutura de painel. Assim, conforme estudo apresentado por Ansari et al. (2002), foi necessário generalizar o pseudo-Fator de Bayes para esta situação, de forma a obter:

$$PsdFB_{ij} = \ln \left(\frac{\prod_{i=1}^m \hat{P}(Y_{i(n+1)} | Y, M_i)}{\prod_{i=1}^m \hat{P}(Y_{i(n+1)} | Y, M_j)} \right),$$

o que leva esta expressão a considerar o produto das estimativas de uma observação futura de cada indivíduo i.

3.8 Simulação de dados

Um estudo de simulação de dados foi conduzido com o objetivo de avaliar a metodologia desenvolvida, ou seja, foram simulados dados em estrutura de painel, cujos valores dos parâmetros são conhecidos (Thompson, 1999), os quais na literatura são denominados valores paramétricos. Por se tratar de inferência Bayesiana, e principalmente havendo o interesse em comparações de

prioris, estes valores paramétricos devem ser gerados por distribuições de probabilidade com parâmetros conhecidos.

Este estudo de simulação, além de permitir uma comparação entre os valores paramétricos e os valores estimados, também possibilita a verificação da eficiência do Fator de Bayes e do Pseudo-Fator de Bayes como recursos de comparação de prioris, pois ao analisar dados gerados por uma distribuição Normal multivariada – Gama Inversa (modelo 1), espera-se que na comparação, os resultados apontem para este modelo como sendo o preferido em relação aos modelos caracterizados pelas prioris t-Student multivariada – Gama Inversa (modelo 2) e Jeffreys (modelo 3). Este procedimento também foi utilizado para os dados gerados pela distribuição t-Student multivariada – Gama Inversa. Por opção, adotou-se um modelo auto-regressivo de segunda ordem, AR(2), o qual é dado por: $Y_{it} = \phi_{i1}Y_{i(t-1)} + \phi_{i2}Y_{i(t-2)} + e_{it}; \quad i = 1, 2, \dots, 10, \quad t = 1, 2, \dots, 12$.

O espaço paramétrico do referido modelo é representado por:

$$\phi = [\phi_{i1}, \phi_{i2}]' \in \mathbb{R}^2 \text{ se } \phi_{i1} + \phi_{i2} < 1; \quad \phi_{i2} - \phi_{i1} < 1; \quad -1 < \phi_{i2} < 1$$

Os Quadros 3 e 4 ilustram o sistema de simulação utilizado, respectivamente, para os modelos 1 e 2.

$\phi \sim N \left(\begin{bmatrix} 0,5 \\ -0,5 \end{bmatrix}, \begin{bmatrix} 0,025 & 0 \\ 0 & 0,010 \end{bmatrix} \right)$	
$e_{it} \sim N(0, \sigma_e^2)$	$\sigma_e^2 \sim GI(3, 2)$

Quadro 3. Distribuições de probabilidade utilizadas no processo de simulação dos dados por meio do modelo 1.

$\phi \sim t\text{-Student} \left(\begin{bmatrix} 0,5 \\ -0,5 \end{bmatrix}, \begin{bmatrix} 0,025 & 0 \\ 0 & 0,010 \end{bmatrix}, gl = m(n-1) \right)$	
$e_{it} \sim N(0, \sigma_e^2)$	$\sigma_e^2 \sim GI(3, 2)$

Quadro 4. Distribuições de probabilidade utilizadas no processo de simulação dos dados por meio do modelo 2.

Uma observação importante está relacionada com a ausência de um estudo de simulação de dados utilizando a priori de Jeffreys (modelo 3), e isto se deve ao fato desta priori não ser caracterizada como uma distribuição de probabilidade, o que a impede de gerar valores para os parâmetros de auto-regressão. Portanto, este modelo será comparado com os demais apenas para testar a eficiência do Pseudo-Fator de Bayes.

As Figuras 3 e 4 mostram as distribuições de probabilidade utilizadas no processo de simulação dos dados.

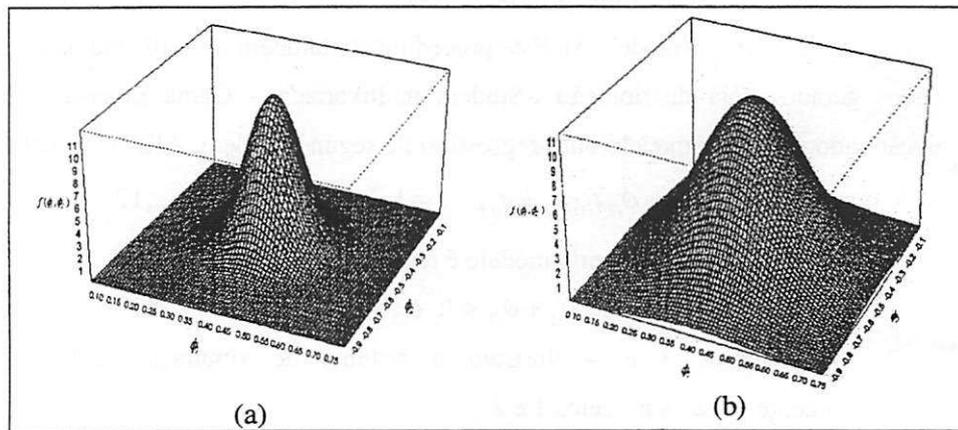


Figura 3. Distribuição Normal multivariada (a) e Distribuição t-Student multivariada (b) utilizadas na simulação dos dados.

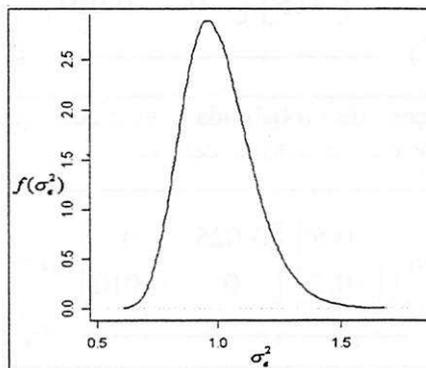


Figura 4. Distribuição Gama Inversa utilizada na simulação dos dados.

As comparações dos modelos via Fator de Bayes e Pseudo-Fator de Bayes foram estruturadas de acordo com a Tabela 2.

Tabela 2. Esquema de comparação das prioris consideradas.

Modelos (simulação)	Critério
	$FB_{12} = \frac{\hat{P}(Y M_1)}{\hat{P}(Y M_2)}$
Normal multivariada – Gama Inversa (modelo 1)	$PsdFB_{13} = \ln \left(\frac{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_1)}{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_3)} \right)$ $PsdFB_{23} = \ln \left(\frac{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_2)}{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_3)} \right)$
	$FB_{21} = \frac{\hat{P}(Y M_2)}{\hat{P}(Y M_1)}$
t-Student multivariada – Gama Inversa. (modelo 2)	$PsdFB_{23} = \ln \left(\frac{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_2)}{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_3)} \right)$ $PsdFB_{13} = \ln \left(\frac{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_1)}{\prod_{i=1}^m \hat{P}(Y_{i(n+l)} Y, M_3)} \right)$

Para avaliar a capacidade preditiva de cada modelo, ou seja, verificar os valores preditos para um dado futuro por meio de suas distribuições preditivas, utilizou-se o recurso apresentado por Liu (1980), o qual consiste na remoção da

última observação de cada indivíduo. Assim os parâmetros dos modelos serão estimados sem a presença destas observações, as quais serão preditas pela metodologia adotada. Uma forma prática e precisa de avaliar a eficiência destas predições está relacionada com o fato dos intervalos de credibilidade conter o verdadeiro valor ocultado na análise.

Um fato que merece destaque na descrição metodológica é referente a possibilidade de se observar, em uma estrutura de dados em painel, a presença de indivíduos cujas séries apresentam diferentes comportamentos, por exemplo, modelos auto-regressivos de diferentes ordens. Além disso, deve ser relatado que os indivíduos não precisam, necessariamente, de apresentar o mesmo número de observações longitudinais. Portanto, o presente trabalho se restringe a considerar, para a análise dos dados reais e dos dados simulados, o mesmo número de observações individuais, respectivamente dez e doze, e o mesmo modelo a ser ajustado, auto-regressivo de segunda ordem, AR(2).

Também é importante frisar que a configuração adotada para os dados simulados, dez indivíduos com doze observações cada um, segue a sugestão de Hsiao (2003), cuja justificativa está fundamentada no uso comum de arquivos com medições mensais, portanto o doze, de um grupo teoricamente pequeno, de tamanho dez, que reflete bem o aproveitamento de observações neste tipo de estrutura.

3.9 Análise dos dados reais

Foram utilizados dados cedidos pelo Grupo de Melhoramento Animal da Faculdade de Zootecnia e Engenharia de Alimentos (USP-Pirassununga,SP), GMA/FZEA, os quais correspondem aos sumários de touros Nelore compreendidos entre os anos de 2000 e 2005. Constam no arquivo dados de DEPs para a característica ganho de peso entre a desmama (205 dias de idade) e o

sobreano (550 dias de idade) de 117 touros, divididos em três grupos, conforme o valor da acurácia referente a última observação, isto é, aos dados de 2005.

Estes grupos foram divididos como seguem: baixa (0 a 40%), contendo 31 touros, média (41 a 60%), contendo 63 touros e alta (acima de 60%), com a frequência de 23 touros. Também foi considerada a análise de todos os animais, sem a separação em grupos de acordo com a acurácia.

Foram utilizados, portanto, quatro arquivos de dados, e todos eles submetidos a análises considerando os três modelos adotados (modelo 1, 2 e 3).

Em todos os arquivos a última observação foi suprimida visando a comparação entre os valores estimados por meio da distribuição preditiva e os verdadeiros valores observados. Os valores iniciais para os parâmetros foram obtidos de um ajuste médio, ou seja, obteve-se uma série cujos valores são as médias de DEPs em cada ano, e a esta se ajustou um modelo AR(2) por meio do método da máxima verossimilhança. Este ajuste foi realizado mediante o pacote *ar.mle* (“Autoregressive Maximum Likelihood Estimate”) do software R.

4 RESULTADOS E DISCUSSÃO

4.1. Apresentação dos dados simulados

O conjunto de dados em estrutura de painel, simulados por meio dos modelos 1 e 2, são apresentados respectivamente nas Tabelas 3 e 4.

Tabela 3. Dados simulados por meio do modelo 1.

		Tempo (t)											
Ind.	(i)	1	2	3	4	5	6	7	8	9	10	11	12
1		1,62	0,52	-0,40	-1,19	0,51	-0,79	0,73	0,21	-0,97	-0,35	1,57	0,69
2		1,07	1,54	0,08	-1,44	-2,39	1,16	1,20	-0,48	-0,42	-0,75	0,32	0,28
3		0,85	0,50	-1,10	-0,65	0,43	1,61	0,59	-0,77	-0,17	0,12	-0,97	0,42
4		0,08	0,53	0,39	-1,30	-0,70	0,92	0,63	-0,60	-1,21	-0,83	0,69	0,66
5		0,02	0,61	0,68	0,03	0,63	1,49	1,64	-0,28	-2,57	-1,20	-0,61	1,18
6		-0,39	-0,36	0,07	0,47	0,53	-0,48	-0,82	0,55	0,62	-0,91	-1,70	1,18
7		0,17	1,75	0,70	-1,40	-2,30	-0,43	1,10	0,47	-0,56	-1,18	0,37	1,43
8		-1,20	-0,16	0,88	-1,19	-0,55	-0,34	-0,23	0,19	0,61	-0,74	-0,56	-0,55
9		-1,11	0,46	0,59	-1,44	-1,15	-0,41	1,10	0,42	-1,57	-1,40	0,32	1,27
10		0,98	-0,15	2,18	-0,65	2,22	0,79	-2,62	-3,54	0,33	2,95	2,73	0,63

Tabela 4. Dados simulados por meio do modelo 2.

		Tempo (t)											
Ind.	(i)	1	2	3	4	5	6	7	8	9	10	11	12
1		-0,40	0,87	0,56	-0,09	-0,69	-0,41	1,53	0,45	0,88	-0,73	-0,74	0,60
2		2,14	-0,52	-2,07	-0,77	0,58	1,10	-0,46	0,16	0,03	-1,37	-0,84	-1,58
3		1,23	0,41	-0,25	-1,29	-0,50	-0,23	-0,80	-0,24	1,68	2,56	1,28	-0,56
4		0,77	1,89	0,44	-2,32	0,14	1,78	0,31	-2,55	-1,84	-0,28	2,23	1,49
5		1,40	0,24	-0,66	1,50	3,89	1,87	-0,94	-2,33	-0,48	-0,94	-0,71	-0,70
6		1,22	0,29	-0,73	-0,86	-0,18	0,09	0,67	-0,95	-1,42	0,24	0,60	0,03
7		0,34	-0,06	-0,92	0,31	1,00	1,50	-0,68	-3,44	-2,00	0,92	1,48	-0,54
8		1,36	0,00	0,53	0,47	-0,49	-0,27	2,34	0,17	0,28	-0,66	-1,71	-0,94
9		1,96	-1,01	0,92	0,00	-0,98	-2,51	-1,41	0,88	0,92	-1,63	-1,13	-0,04
10		0,26	-0,07	0,15	0,08	-0,33	-0,28	0,56	0,27	0,57	0,46	-0,56	-0,96

Os gráficos representativos de algumas séries simuladas, respectivamente pelos modelos 1 e 2, são mostrados nas Figuras 5 e 6.

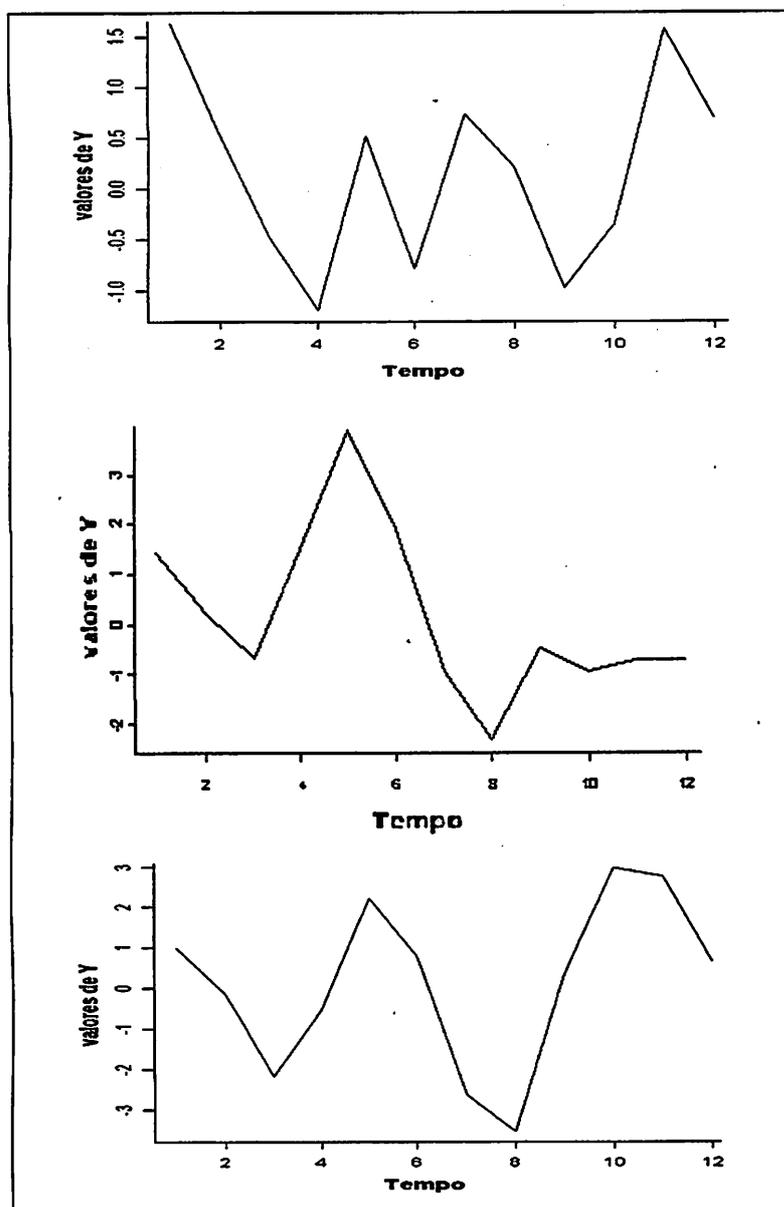


Figura 5. Séries obtidas para os indivíduos 1, 5 e 10, modelo 1.

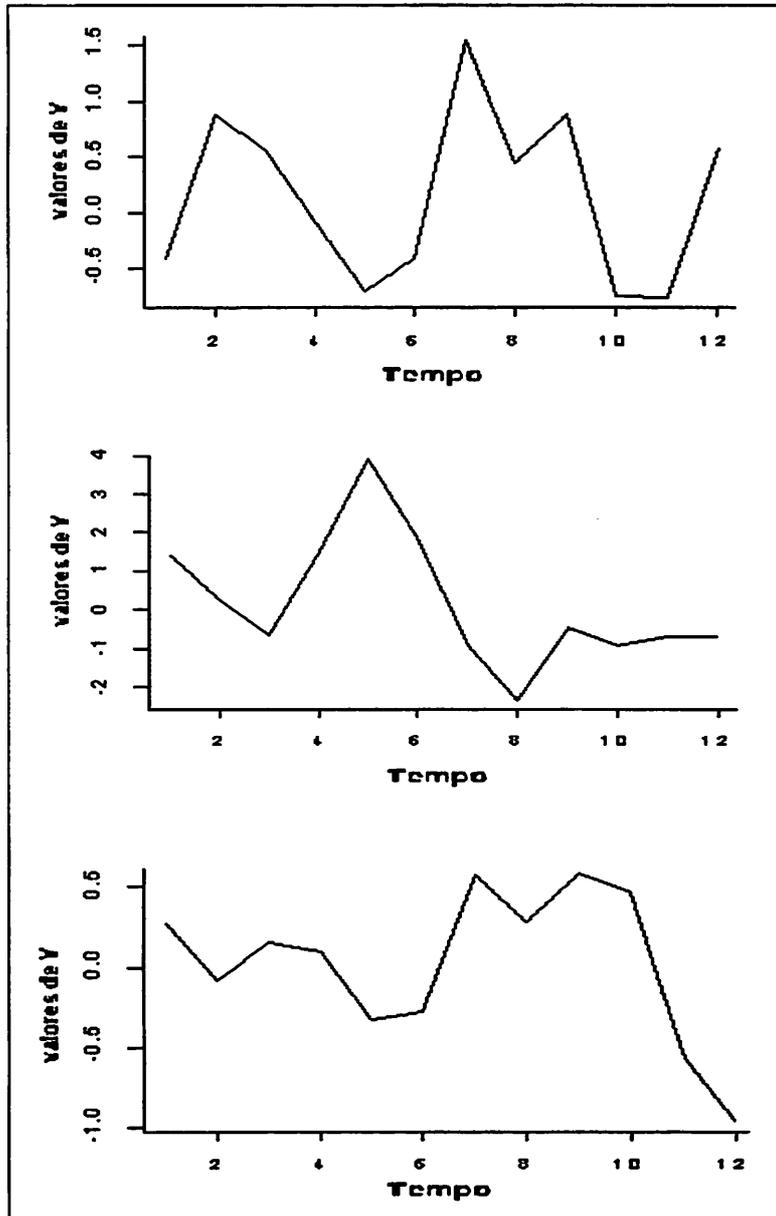


Figura 6. Séries obtidas para os indivíduos 1, 5 e 10, modelo 2.

4.2 Resultados obtidos para os dados simulados

Nas Tabelas 5 e 6 são apresentadas as estimativas (média a posteriori) para o parâmetro de auto-regressão de primeira ordem (ϕ_1), obtidas respectivamente para os modelos 1 e 2, juntamente com os resultados provenientes da análise de convergência das cadeias geradas pelo método MCMC.

Esta mesma configuração é mostrada para o parâmetro de auto-regressão de segunda ordem (ϕ_2) nas Tabelas 7 e 8, e na Tabela 9 encontra-se as estimativas obtidas para a variância do erro (σ_e^2) para ambos os modelos.

Na Figura 7 é apresentado um gráfico “box-plot” que contempla simultaneamente os valores das cadeias MCMC de todos os indivíduos para cada parâmetro (ϕ_1 e ϕ_2) levando-se em conta os dois modelos considerados.

Tabela 5. Valor do Parâmetro ϕ_1 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke e “burn-in” considerando o modelo 1.

Ind. (i)	ϕ_1	$\hat{\phi}_1$	LI*	LS	$\sqrt{\hat{R}}$	Z	Burn-in
1	0,24	0,28	0,22	0,33	1,0001	0,95	2
2	0,31	0,36	0,30	0,42	1,0000	0,84	2
3	0,59	0,69	0,55	0,82	1,0003	0,55	6
4	0,36	0,33	0,27	0,38	1,0000	0,39	2
5	0,34	0,40	0,28	0,53	1,0001	-0,57	3
6	0,41	0,48	0,39	0,58	1,0000	-0,58	2
7	0,39	0,36	0,27	0,44	0,9990	-0,90	2
8	0,60	0,70	0,61	0,82	1,0024	1,15	196
9	0,65	0,61	0,54	0,67	1,0000	0,33	3
10	0,61	0,71	0,63	0,87	1,0013	1,38	208

*LI: Limite inferior; LS: Limite superior

Tabela 6. Valor do Parâmetro ϕ_1 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke e “burn-in” considerando o modelo 2.

Ind. (i)	ϕ_1	$\hat{\phi}_1$	LI	LS	$\sqrt{\hat{R}}$	Z	Burn-in
1	0,47	0,51	0,32	0,68	0,9932	0,42	3
2	0,65	0,67	0,59	0,74	0,9916	0,63	3
3	0,39	0,45	0,41	0,55	1,1041	1,42	250
4	0,47	0,49	0,35	0,65	0,9940	0,33	3
5	0,84	0,89	0,78	1,1	0,9190	1,31	105
6	0,60	0,69	0,59	0,78	0,9914	0,63	2
7	0,33	0,36	0,24	0,50	0,9220	0,58	2
8	0,49	0,51	0,33	0,67	0,9970	0,47	3
9	0,30	0,34	0,24	0,45	0,9902	0,97	2
10	0,41	0,41	0,34	0,49	0,9989	0,45	5

Tabela 7. Valor do Parâmetro ϕ_2 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke e “burn-in” considerando o modelo 1.

Ind. (i)	ϕ_2	$\hat{\phi}_2$	LI	LS	$\sqrt{\hat{R}}$	Z	Burn-in
1	-0,60	-0,52	-0,62	-0,44	0,9955	0,85	4
2	-0,57	-0,49	-0,60	-0,39	0,9900	-0,64	3
3	-0,34	-0,39	-0,46	-0,33	0,9941	0,35	3
4	-0,68	-0,58	-0,72	-0,43	0,9930	0,54	3
5	-0,60	-0,51	-0,65	-0,36	0,9399	1,17	205
6	-0,50	-0,43	-0,56	-0,32	0,9811	0,14	3
7	-0,52	-0,45	-0,58	-0,33	0,9900	0,78	2
8	-0,36	-0,71	-0,66	-0,49	1,0450	-1,05	505
9	-0,64	-0,75	-0,91	-0,57	0,9802	0,37	5
10	-0,81	-0,75	-0,92	-0,59	1,0159	0,56	320

Tabela 8. Valor do Parâmetro ϕ_2 usado na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke, e “burn-in” considerando o modelo 2.

Ind. (i)	ϕ_2	$\hat{\phi}_2$	LI	LS	$\sqrt{\hat{R}}$	Z	Burn-in
1	-0,78	-0,82	-0,88	-0,75	0,9901	0,14	2
2	-0,58	-0,66	-0,79	-0,53	0,9990	0,25	2
3	-0,61	-0,61	-0,78	-0,53	0,9991	0,33	1
4	-0,56	-0,61	-0,72	-0,49	0,9990	-0,18	2
5	-0,57	-0,60	-0,66	-0,54	0,9839	0,32	3
6	-0,40	-0,38	-0,55	-0,22	0,9911	0,08	3
7	-0,60	-0,65	-0,71	-0,59	0,9905	-0,96	3
8	-0,59	-0,62	-0,69	-0,54	0,9787	1,25	102
9	-0,31	-0,32	-0,39	-0,24	0,9920	-0,23	2
10	-0,38	-0,37	-0,42	-0,33	0,9959	0,38	2

Tabela 9. Variâncias residuais usadas na simulação, suas estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, escore Z de Geweke, e “burn-in” considerando os modelos 1 e 2.

Modelo	σ_e^2	$\hat{\sigma}_e^2$	LI	LS	$\sqrt{\hat{R}}$	Z	Burn-in
1	1,9905	2,7023	1,9414	4,3202	0,9458	1,25	306
2	1,0541	1,6061	0,9599	3,0250	0,9874	1,33	450

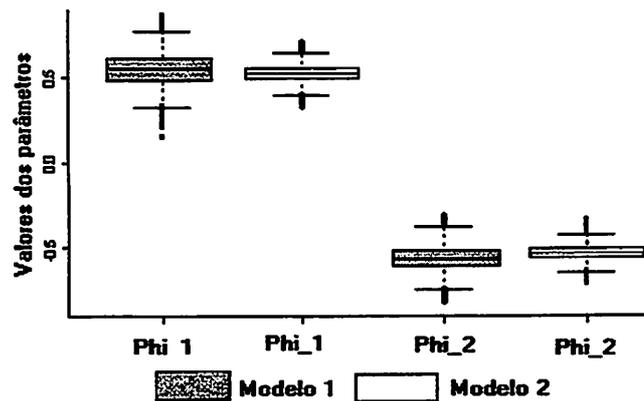


Figura 7. Box-plot para os parâmetros de auto-regressão de todos os indivíduos.

Tendo em vista os resultados obtidos pelos critérios de convergência utilizados, é possível dizer que todas as cadeias geradas pelo método MCMC apresentaram comportamentos semelhantes, e que todas elas, sem exceção, atingiram a convergência requerida com as 50.000 iterações utilizadas.

Esta convergência foi constatada devido ao fato dos valores do Fator de Gelman-Rubin se conservarem próximos a 1. Além disso, os valores do escore Z foram todos menores que 1,96, o que segundo Geweke (1992) aponta para a estacionaridade da cadeia gerada, que por sua vez está relacionada com a verificação da convergência. Os valores indicados para o “burn-in” por meio do critério de Raftery & Lewis (1992) também foram todos inferiores aos utilizados, os quais foram de 25.000, que representa 50% das observações. Este último resultado certifica que as inferências a posteriori foram realizadas desconsiderando possíveis influências dos valores iniciais.

Em relação aos resultados apresentados nas Tabelas 5 e 6 pode-se afirmar que, de forma geral, a metodologia utilizada foi eficiente, pois os intervalos de credibilidade a posteriori incluem, respectivamente para os modelos 1 e 2, 80% e 90% dos valores paramétricos de ϕ_1 considerados na simulação. Ao se tratar do parâmetro ϕ_2 , cujos resultados encontram-se nas Tabelas 7 e 8, estas porcentagens assumem, respectivamente para os modelos 1 e 2, os valores de 90% e 100%.

Estes resultados podem ser interpretados de maneira prática como indicativos de uma maior qualidade do modelo 2, visto que foi verificada uma superação deste modelo em relação ao modelo 1 na porcentagem de inclusão dos verdadeiros valores dos parâmetros nos intervalos de credibilidade a posteriori. Além disso, a Figura 7 mostra um gráfico “box-plot” no qual se verifica uma menor variação para todos os parâmetros estimados por meio do modelo 2, o que indica uma maior precisão por parte deste modelo. Resultados semelhantes foram obtidos por Barreto & Andrade (2004), os quais constataram uma maior

precisão das estimativas na presença de uma distribuição a priori t-multivariada para os parâmetros do modelo auto-regressivo considerado.

Na Tabela 9 pode-se perceber que para ambos os modelos os intervalos de credibilidade a posteriori incluem os verdadeiros valores da variância do erro, indicando que na estimação deste parâmetro a metodologia também se mostrou eficiente.

Diante desses resultados iniciais, torna-se relevante ressaltar que, embora estas discussões apontem para a priori t-multivariada como sendo a de maior qualidade, a avaliação da qualidade dos modelos será discutida posteriormente mediante a utilização de critérios específicos, resguardando as informações contidas nas Tabelas 5, 6, 7, 8 e 9 apenas para a verificação da eficiência da metodologia utilizada.

As Figuras 8 e 9 mostram, respectivamente para os modelos 1 e 2, as cadeias e as densidades marginais a posteriori para os parâmetros ϕ_1 e ϕ_2 de alguns indivíduos. Na Figura 10 são apresentados, sob o mesmo conceito, os gráficos para a variância do erro.

A Figura 11 apresenta as distribuições preditivas a posteriori referentes às últimas observações (y_{12}) para alguns indivíduos. Na Tabela 10 estão listadas, para ambos os modelos, as estimativas de y_{12} para todos os indivíduos. Nesta tabela são mostrados também os intervalos de credibilidade (95%) a posteriori, os quais permitem uma comparação simples e objetiva em relação a capacidade de predição de cada modelo.

Na Figura 12 tem-se um gráfico representativo das amplitudes dos intervalos de credibilidade para y_{12} referentes a cada um dos dez indivíduos, considerando as estimativas obtidas pelos modelos 1 e 2.

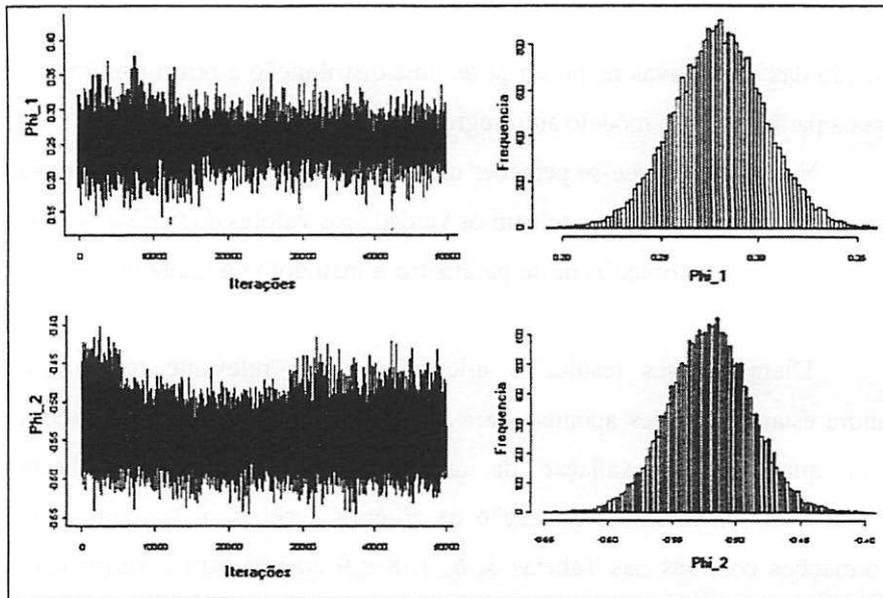


Figura 8. Cadeias geradas pelo método MCMC e densidade a posteriori para os parâmetros ϕ_1 e ϕ_2 referentes ao indivíduo 1 (modelo 1).

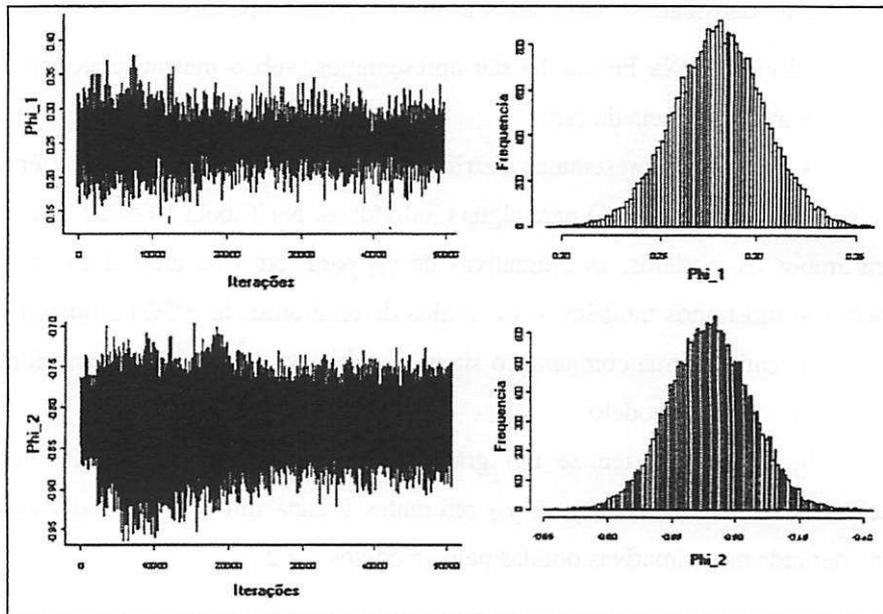


Figura 9. Cadeias geradas pelo método MCMC e densidade a posteriori para os parâmetros ϕ_1 e ϕ_2 referentes ao indivíduo 1 (modelo 2).

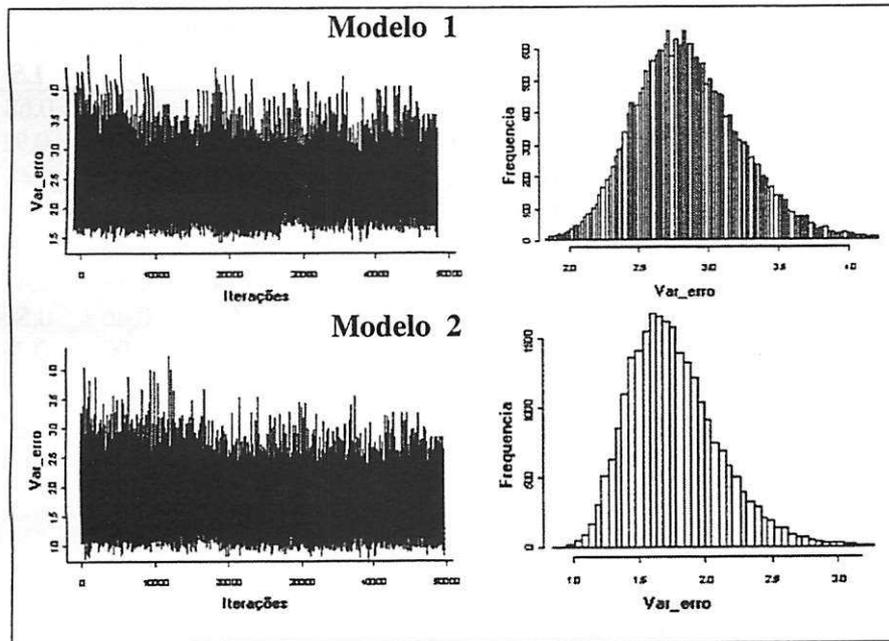


Figura 10. Cadeias geradas pelo método MCMC e densidade a posteriori para a variância do erro.

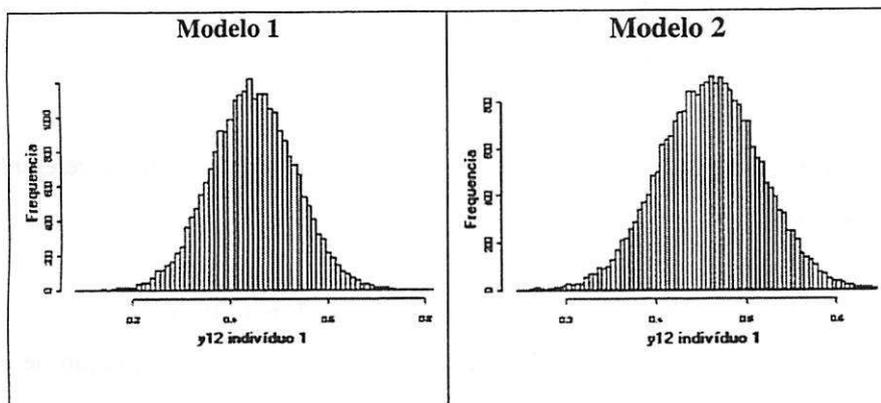


Figura 11. Distribuições preditivas de y_{12} para o indivíduo 1.

Tabela 10. Última observação (y_{12}), suas estimativas (médias a posteriori) e intervalos de credibilidade (95%) para ambos os modelos.

Ind. (i)	Modelo 1				Modelo 2			
	y_{12}	\hat{y}_{12}	LI	LS	y_{12}	\hat{y}_{12}	LI	LS
1	0,69	0,43	0,12	0,74	0,60	0,46	0,24	0,68
2	0,28	0,12	-0,13	0,46	-1,58	-1,17	-1,63	-0,91
3	0,42	0,25	0,06	0,44	-0,56	0,02	-0,44	0,48
4	0,66	1,04	0,69	1,39	1,49	1,92	1,35	2,49
5	1,18	1,36	1,13	1,59	-0,70	-0,92	-1,24	-0,60
6	1,18	1,41	0,89	1,93	0,03	0,14	-0,19	0,47
7	1,43	1,02	0,75	1,29	-0,54	0,06	-0,46	0,51
8	-0,55	-0,08	-0,42	-0,26	-0,94	-0,69	-1,00	-0,38
9	1,27	0,97	0,65	1,30	-0,04	0,13	-0,06	0,33
10	0,63	1,25	0,95	1,55	-0,96	-1,19	-1,54	-0,84

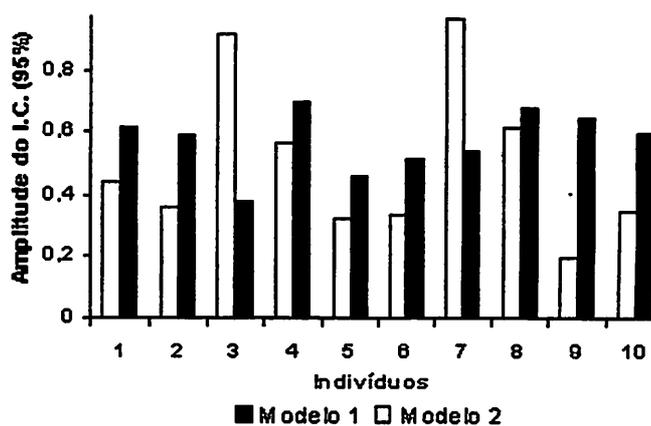


Figura 12. Amplitudes do Intervalo de Credibilidade de 95% para y_{12} referentes a todos os indivíduos.

De acordo com os resultados mostrados na Tabela 10 percebe-se que, de forma geral, a metodologia utilizada mostrou-se eficiente na estimação de um dado futuro, uma vez que ao se considerar o modelo 1, 60 % dos indivíduos têm inclusos em seus Intervalos de Credibilidade os verdadeiros valores da última

observação (y_{12}), e ao se considerar o modelo 2, 80% dos indivíduos apresentam esta característica. Portanto ao se avaliar conjuntamente os dois modelos, constata-se uma eficiência de 70% na predição de um dado futuro. Esta marca é semelhante àquelas apresentadas na literatura, cujos autores também discutem a capacidade preditiva de modelos mediante avaliação dos intervalos de credibilidade a posteriori. Sob este enfoque, de Alba (1993), ao simular quatro séries independentes sob um modelo auto-regressivo de ordem quatro, AR (4), obteve 75% de eficiência na predição de um dado futuro. Hay & Pettitt (2001) obtiveram uma eficiência de 58% ao analisar doze séries independentes de incidência de pneumonia por meio de um modelo auto-regressivo de primeira ordem, AR(1), generalizado para dados de contagem. Recentemente, Brandt & Freeman (2006), reanalisaram dados de política externa de Israel, compreendidos entre abril de 1979 e dezembro de 1988, por meio de modelos auto-regressivos multivariados, e concluíram que, em média, as regiões de credibilidade para um vetor de dados futuros continham os verdadeiros valores dos parâmetros em 68% dos casos.

A Tabela 10 também fornece informações relacionadas com a comparação entre a capacidade preditiva dos modelos 1 e 2, pois nota-se que o primeiro apresentou uma eficiência de 60%, ou seja, seis dos dez intervalos referentes aos indivíduos incluía o verdadeiro valor da última observação, enquanto que o segundo assegurou uma eficiência de 80%. Estes resultados estão de acordo com aqueles obtidos por Barreto & Andrade (2004), os quais concluíram que a adoção de uma priori mais robusta, no caso a distribuição t-multivariada, fornece estimativas mais precisas para dados futuros, fato este que também pode ser verificado pela Figura 12, na qual se observa intervalos de credibilidade mais estreitos para o modelo 2 em 80% dos casos.

A comparação entre os modelos por meio dos critérios Fator de Bayes e Pseudo-Fator de Bayes é apresentada na Tabela 11.

Tabela 11. Modelos utilizados na simulação dos dados e respectivos critérios de comparação.

Modelos (simulação)	Critério
	$FB_{12} = \frac{1.336.129}{5.133.331} = 0,2602$
Normal multivariada – Gama Inversa (modelo 1)	$PsdFB_{13} = \ln\left(\frac{-0,0025}{-0,0013}\right) = 0,6539$
	$PsdFB_{23} = \ln\left(\frac{-0,4321}{-0,0013}\right) = 5,8062$
	$FB_{21} = \frac{2.973.740,179}{6.263,167} = 474.798$
t-Student multivariada – Gama Inversa (modelo 2)	$PsdFB_{23} = \ln\left(\frac{-0,00032}{-0,0000222}\right) = 2,6682$
	$PsdFB_{13} = \ln\left(\frac{-0,0000917}{-0,0000222}\right) = 1,4184$

FB_{ij} : Fator de Bayes para os modelos i e j; $PsdFB_{ij}$: Pseudo-Fator de Bayes para os modelos i e j.

Nota-se na Tabela 11 que para ambos os modelos considerados na simulação, a utilização de uma priori t-multivariada para o vetor de parâmetros auto-regressivos mostrou-se mais eficiente. Em relação ao modelo 1, a discussão deste resultado é bem interessante, pois teoricamente os parâmetros foram gerados de uma distribuição Normal-multivariada, o que pressupõe a indicação desta como a mais adequada. Uma possível explicação está no fato da distribuição t-multivariada utilizada apresentar um elevado número de graus de liberdade (110), o que faz com que a mesma se aproxime de forma significativa da distribuição Normal-multivariada. Além disso, Kass & Raftery (1995) apresentam uma escala de valores para o Fator de Bayes, a qual permite julgar de forma mais detalhada a superioridade de um modelo em relação ao seu

contrastante, e esta admite que valores entre 1 e 3,2 indicam uma ligeira superioridade do modelo disposto no numerador, a qual não deve ser discutida de forma decisiva a respeito da qualidade de um dos modelos. Resultados semelhantes foram discutidos por Kizilkaya et al. (2003), os quais utilizaram a simulação de dados para avaliar a eficiência entre modelos probit usando funções de ligação t-Student e Normal acumuladas. Estes autores simularam três conjuntos de dados fundamentados na distribuição Normal, e em dois deles o Fator de Bayes apontou para a t-Student como sendo a mais adequada.

De forma geral, a literatura relata em vários trabalhos a superioridade da distribuição a priori t-Student, e a enfatiza como sendo a mais adequada para modelos auto-regressivos de séries temporais. Dentre estes destacam-se Barreto & Andrade (2004), os quais usaram o fator de Bayes e as estimativas de densidades preditivas como critério de comparação no estudo de uma série de vazões de usina hidrelétrica, e Huerta & West (1999) que estudaram séries históricas na área de climatologia usando a capacidade preditiva como parâmetro de comparação. Segundo estes autores a forma mais achatada desta distribuição, ou seja, sua cauda mais pesada, faz uma melhor cobertura de valores especialmente importantes para o ajuste dos modelos auto-regressivos.

Ainda em relação aos resultados apresentados na Tabela 11, é importante realçar que em todas as comparações, considerando ambos os modelos, que a utilização da priori não-informativa de Jeffreys apresentou baixa eficiência se comparadas com as demais, e este fato vem sustentar uma recente tendência na área da inferência Bayesiana relacionada com a comparação de prioris, a qual conduz na maioria das vezes a indicação de prioris informativas (Lambert et al., 2005).

4.3 Resultados obtidos para os dados reais

Os gráficos do tipo Box-plot para os dados de avaliações genéticas de gado Nelore compreendidas entre os anos de 2000 a 2005, para cada conjunto (baixa, média, alta e geral) determinado de acordo com acurácia, podem ser visualizados nas Figuras 13 a 16.

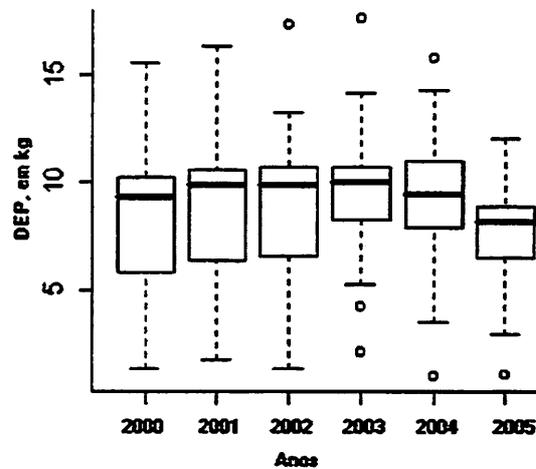


Figura 13.Box-plot para valores do conjunto representado pela acurácia baixa.

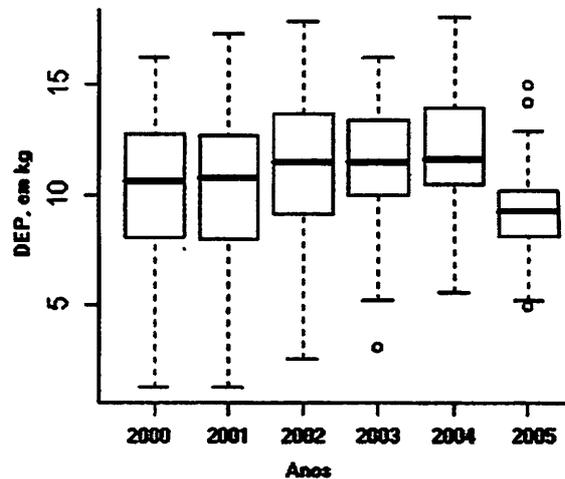


Figura 14.Box-plot para valores do conjunto representado pela acurácia média.

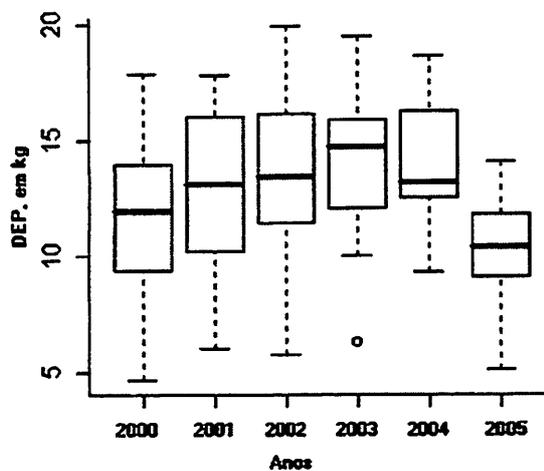


Figura 15.Box-plot para valores do conjunto representado pela acurácia alta.

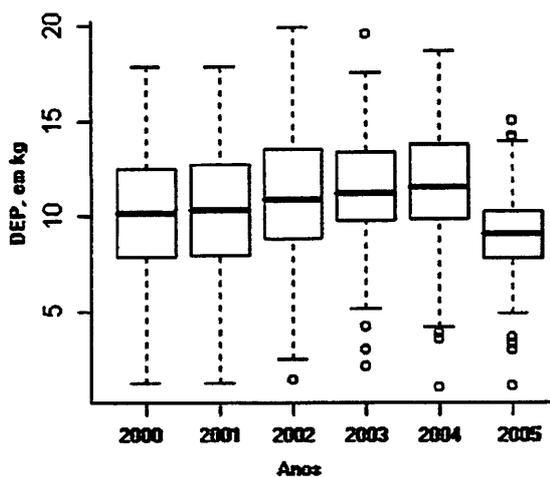


Figura 16. Box-plot para valores do conjunto representado pela ausência da divisão por classes de acurácia.

De acordo com as Figuras 13 a 16 nota-se que a ausência da divisão por classes de acurácia (Figura 16) proporciona uma maior variabilidade dos dados em cada ano estudado, uma vez que o número de observações, respectivamente abaixo e acima dos quantis 2,5% e 97,5%, é consideravelmente maior em

relação aos demais conjuntos. Esta informação pode ser utilizada como uma justificativa preliminar para a separação das observações de acordo com os valores da acurácia.

Para otimizar o processamento dos dados dentro do sistema MCMC, tornou-se útil atribuir valores iniciais, para os parâmetros a serem estimados, mais condizentes com a realidade. Para tanto, utilizou-se como valores iniciais as estimativas de ϕ_1 , ϕ_2 e σ_e^2 obtidas pelo método da Máxima Verossimilhança. Este método foi aplicado para analisar as séries representadas pelas médias das observações em cada ano, considerando cada conjunto de dados. Estes resultados são apresentados na Tabela 12 e na Figura 17 são mostradas as séries, cujo modelo AR(2) foi ajustado.

Tabela 12. Estimativas de Máxima Verossimilhança e desvio-padrão (DP).

Estrutura dos dados	$\hat{\phi}_1$	$DP(\hat{\phi}_1)$	$\hat{\phi}_2$	$DP(\hat{\phi}_2)$	$\hat{\sigma}_e^2$
Acurácia baixa	0,1703	0,3650	-0,2444	0,1233	0,7938
Acurácia média	1,1859	0,3190	-0,5471	0,0246	0,2391
Acurácia alta	1,8299	0,3766	-0,9614	0,1510	0,2526
Geral	1,6874	0,3680	-0,9302	0,1214	0,8240

Nota-se na Tabela 12 que as estimativas para os parâmetros ϕ_1 e ϕ_2 mostraram-se menos precisas no conjunto de dados representados por animais com baixa acurácia, uma vez que seus desvios-padrão são proporcionalmente maiores em relação aqueles apresentados para as outras estruturas consideradas. Uma possível explicação está relacionada com a ocorrência de uma maior variação no comportamento das séries que formam este conjunto de dados médios, o que leva a uma falta de padronização que reflete na precisão das estimativas obtidas.

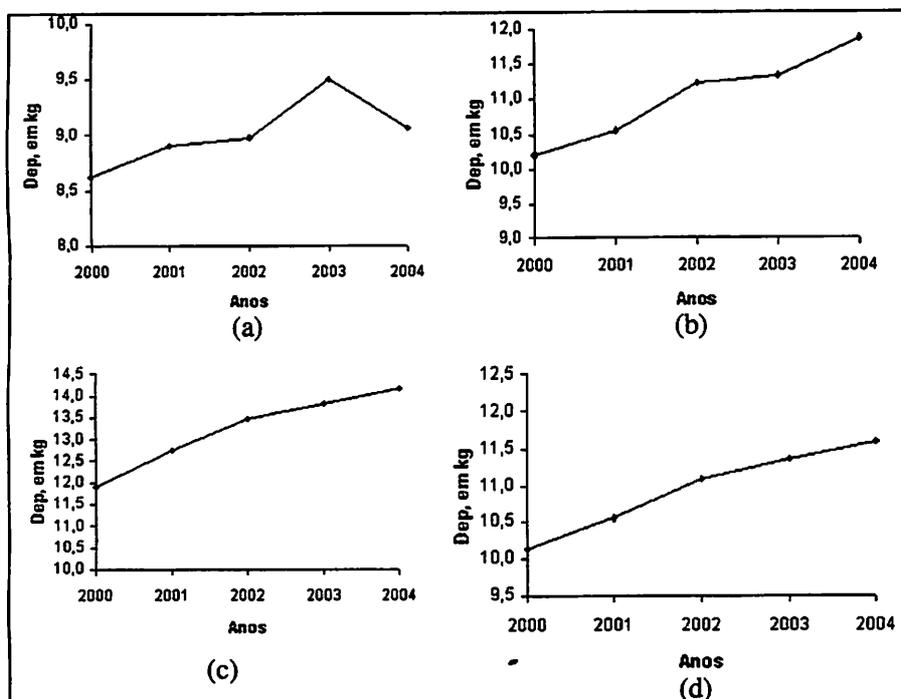


Figura 17. Série obtida com os dados médios para conjunto classificado de acordo com a acurácia: baixa (a), média (b), alta (c) e geral (d).

Os dados reais foram analisados por meio dos três modelos apresentados anteriormente. Portanto, foram calculados os avaliadores (Fator de Bayes e Pseudo-Fator de Bayes) da qualidade de ajuste, os quais possibilitaram a escolha de um modelo como sendo o mais adequado. Estes resultados referentes a comparação dos modelos, dentro de cada estrutura de dados, encontram-se na Tabela 13.

A Tabela 14 apresenta um resumo dos resultados obtidos no ajuste do modelo 2 aos dados reais de DEP's. As Tabelas 1A, 2A, 3A e 4A em anexo fornecem as observações utilizados para calcular as porcentagens listadas na Tabela 14.

Tabela 13. Modelos considerados na análise dos dados reais referentes às Dep's calculadas no decorrer dos anos de 2000 a 2004 e respectivos critérios de comparação.

Estrutura dos dados	Critério
Acurácia baixa	$FB_{21} = \frac{105.542,540}{52.502,635} = 2,0102$
	$PsdFB_{23} = \ln\left(\frac{-0,0875}{-0,0035}\right) = 3,2188$
	$PsdFB_{13} = \ln\left(\frac{-0,0541}{-0,0035}\right) = 2,7380$
Acurácia média	$FB_{21} = \frac{206.664,210}{62.402,354} = 3,3118$
	$PsdFB_{23} = \ln\left(\frac{-0,1255}{-0,0011}\right) = 4,7369$
	$PsdFB_{13} = \ln\left(\frac{-0,0902}{-0,0011}\right) = 4,4067$
Acurácia alta	$FB_{21} = \frac{176.001,435}{58.213,034} = 3,0234$
	$PsdFB_{23} = \ln\left(\frac{-0,0987}{-0,0023}\right) = 3,7591$
	$PsdFB_{13} = \ln\left(\frac{-0,0857}{-0,0023}\right) = 3,6179$
Geral	$FB_{21} = \frac{112.040,331}{48.111,012} = 2,3287$
	$PsdFB_{23} = \ln\left(\frac{-0,0898}{-0,0021}\right) = 3,7556$
	$PsdFB_{13} = \ln\left(\frac{-0,0741}{-0,0021}\right) = 3,5634$

1: priori Normal multivariada - gama inversa, 2: priori t Student multivariada - gama inversa, 3: priori de Jeffreys

Os resultados contidos na Tabela 13 informam que, analogamente aqueles obtidos no estudo de simulação, a distribuição t-Student multivariada proporcionou uma melhor qualidade de ajuste em relação às outras duas prioris utilizadas em todas as estruturas de dados consideradas, embora a superioridade em relação à distribuição Normal multivariada seja de pequena magnitude se avaliada mediante o critério de interpretação do Fator de Bayes apresentado por Kass & Raftery (1995).

Também foi observado na Tabela 13 que o pior desempenho foi apresentado pela priori de Jeffreys, indicando que mesmo quando não se tem informações a priori, prioris informativas podem ser consideradas visando a comparação de sua eficiência em relação a priori de Jeffreys (Lambert et al., 2005)

Para complementar esta discussão, é viável salientar que os fatos abordados nos dois parágrafos anteriores são menos evidentes na estrutura de baixa acurácia, uma vez que os valores obtidos para o fator de Bayes e para o Pseudo-Fator de Bayes são de menor magnitude. Isto talvez seja devido a uma maior heterogeneidade no comportamento das séries referentes a esta estrutura.

Tabela 14. Resultados referentes à porcentagem de significância dos parâmetros do modelo AR(2) ajustado às séries individuais de Dep's.

Estrutura dos dados	Significância (%)	
	Parâmetro ϕ_1	Parâmetro ϕ_2
Acurácia baixa	43,47	43,47
Acurácia média	73,01	65,07
Acurácia alta	65,21	65,21
Geral	61,53	57,26

De acordo com a Tabela 14 é possível inferir que as séries referentes à população de indivíduos com baixa acurácia apresentam um comportamento diferenciado das demais, pois a porcentagem de significância dos parâmetros foi

menor. Nota-se também, que o parâmetro ϕ_2 apresentou uma porcentagem de significância um pouco menor que aquela obtida para o parâmetro ϕ_1 em duas estruturas de dados (média e geral), indicando que as séries de DEP's de alguns indivíduos não apresentam um comportamento auto-regressivo de segunda ordem, podendo estas serem descritas por uma ordem inferior, AR(1), ou ainda se apresentarem como um processo estacionário, que indica a ausência de autocorrelação.

Para explicar de forma mais detalhada os resultados apresentados na Tabela 14, foram confeccionados gráficos das funções de autocorrelação e autocorrelação parcial para as séries dos indivíduos cujos parâmetros ϕ_1 e ϕ_2 não se mostraram significativos. Estes gráficos estão nas Figuras 1A, 2A e 3A (Ver Anexo). Para os indivíduos cujos gráficos da função de autocorrelação indicaram a presença de “ruídos”, obteve-se os gráficos de autocorrelação parcial para verificar a ordem indicada para o modelo auto-regressivo, e aqueles cujos “ruídos” não foram detectados, assumiu-se a caracterização de um processo estacionário.

Por meio da interpretação destes gráficos, nota-se que na estrutura relacionada com baixa acurácia (Figura 1A) a função de autocorrelação parcial apresentou lag's significativos para alguns indivíduos, indicando que modelos de ordem superior, AR(3), devem ser considerados em 38% dos casos. Além disso, também foram identificados indivíduos (62%) cujas séries já são caracterizadas como um processo estacionário, pois em nenhum lag da função de autocorrelação observaram-se pontos fora do intervalo limítrofe. Esta variedade de processos identificados nesta estrutura de dados não foi observada nos conjuntos relacionados com a acurácia média e alta, pois o número de indivíduos cujas séries demonstraram seguir um modelo AR(3) foram respectivamente 23,80% e 11,11%, observando-se portanto uma quase totalidade de processos

estacionários (Figuras 2A e 3A). Esta discussão também evidencia o fato da acurácia baixa proporcionar maior variabilidade entre as séries, o que justifica a separação dos indivíduos em grupos. Porém, não foram observadas evidências a respeito de diferenças na variabilidade do comportamento de séries representadas pelas estruturas de acurácia média e alta, o que pode indicar uma possível união desses indivíduos em uma só estrutura.

O ponto mais marcante em relação aos resultados da Tabela 14 está relacionado com a seleção de indivíduos cujas densidades preditivas serão apresentadas, pois, não há razão em discutir o processo de previsão de valores futuros de séries individuais que não apresentam um processo auto-regressivo de primeira e segunda ordem, uma vez que a densidade preditiva está fundamentada no modelo AR(2) para dados em painel. Assim, serão discutidas sob o aspecto preditivo apenas séries cujos parâmetros mostraram-se significativos. As Tabelas 15, 16 e 17 apresentam os verdadeiros valores da última observação (y_6) correspondentes às DEPs do ano de 2005, os quais não foram considerados na análise para possibilitar a comparação entre valor observado e estimado via densidades preditivas.

Tabela 15. Verdadeiros valores da última observação (y_6), suas estimativas e intervalos de credibilidade (95%) considerando a estrutura de acurácia baixa.

Ind.	y_6	\hat{y}_6	LI	LS	Ind.	y_6	\hat{y}_6	LI	LS
6	8,78	9,60	7,11	12,02	20	8,41	9,40	5,72	12,18
8	8,71	10,85	5,34	15,35	21	9,05	8,44	6,82	12,32
9	8,48	10,89	4,88	16,09	22	8,40	8,97	3,21	12,25
10	3,28	9,39	4,42	10,57	23	9,15	12,77	10,07	16,28
12	3,61	4,03	-3,99	8,14	25	9,88	9,57	1,53	14,45
13	6,31	7,90	3,76	10,66	27	9,10	7,28	1,99	12,91
14	10,41	16,97	14,98	19,95	28	7,60	4,44	-1,61	6,53
15	5,62	4,38	-4,75	7,14	29	8,17	8,73	0,99	16,57
16	7,75	5,99	-2,05	11,02	31	2,97	0,68	-2,64	4,96

Tabela 16. Verdadeiros valores da última observação (y_6), suas estimativas e intervalos de credibilidade (95%) considerando a estrutura de acurácia média.

Ind	y_6	\hat{y}_6	LI	LS	Ind	y_6	\hat{y}_6	LI	LS
2	8,49	9,87	-0,78	11,02	41	12,66	11,29	7,83	16,15
5	6,55	5,25	1,09	17,74	42	8,95	8,67	2,72	14,00
6	9,11	9,57	6,93	12,89	43	10,23	10,73	6,06	15,96
8	12,32	10,04	8,79	17,72	44	6,84	5,52	2,79	13,54
10	9,06	12,75	9,28	15,24	45	8,80	8,05	2,26	17,54
11	9,77	7,18	2,28	15,86	46	6,43	9,29	7,06	12,07
16	10,99	14,20	11,03	17,84	47	9,90	7,95	2,03	18,55
17	7,21	5,20	3,81	8,75	48	9,18	13,12	10,81	17,39
21	7,66	7,05	3,29	9,59	51	11,25	14,32	2,09	19,32
22	8,92	9,42	3,96	18,98	52	5,28	5,82	2,65	9,44
24	4,94	8,58	5,26	17,31	53	12,49	13,22	7,74	18,56
25	8,22	8,63	3,05	12,12	54	8,10	13,76	3,02	18,02
26	9,39	11,35	7,48	14,68	55	8,70	6,71	3,21	11,33
28	8,55	8,55	4,96	12,94	56	9,29	7,17	3,86	15,51
30	9,67	7,86	-0,07	13,36	57	9,61	8,61	5,68	12,93
31	12,87	12,78	5,52	18,96	58	9,91	11,51	6,13	14,66
33	10,06	15,86	11,44	20,13	59	9,00	10,50	4,88	16,46
35	14,95	12,01	6,47	18,32	60	9,65	10,79	2,92	18,97
37	8,36	10,21	2,66	17,91	61	10,73	10,31	6,20	15,72
38	8,58	7,01	2,21	18,01	62	5,22	7,19	0,02	12,68
40	10,86	10,44	6,10	15,40	63	9,59	15,15	11,04	21,07

Tabela 17. Verdadeiros valores da última observação (y_6), suas estimativas e intervalos de credibilidade (95%) considerando a estrutura de acurácia alta.

Ind	y_6	\hat{y}_6	LI	LS	Ind	y_6	\hat{y}_6	LI	LS
1	9,20	10,38	3,10	12,01	12	14,10	14,24	2,39	23,83
3	9,06	10,94	7,12	12,08	14	10,84	14,06	6,40	16,90
6	10,55	17,46	12,30	23,07	15	8,38	6,69	1,40	13,95
7	10,03	17,62	11,99	25,09	17	9,71	10,82	4,12	15,62
8	12,55	11,25	2,22	25,13	19	13,91	13,64	4,05	23,69
9	9,12	7,51	3,19	16,31	20	11,41	12,55	2,98	19,16
10	10,85	12,93	4,90	18,65	22	11,79	12,91	4,06	17,92

Os resultados apresentados nas Tabelas 15, 16 e 17 permitem afirmar que a metodologia utilizada para realizar previsões de dados futuros individuais com base na obtenção das distribuições preditivas foi eficiente, uma vez que a porcentagem de intervalos de credibilidade que continham os verdadeiros valores das DEP's referentes ao ano de 2005 variou entre 77,78% e 85,71%. Estes resultados também nos conduzem a discussões relacionadas com as classes de acurácia consideradas, pois nota-se que a capacidade preditiva do modelo é menor quando a acurácia é baixa, 77,78%, e para as acurácias médias e altas, respectivamente 83,33% e 85,71%, não há grandes diferenças em relação a esta capacidade. Este fato também é explicado mediante estimativas obtidas para a variância do erro (σ_e^2) em cada uma das estruturas de dados consideradas, as quais são mostradas na Tabela 18.

Tabela 18. Estimativas (médias a posteriori), intervalos de credibilidade (95%), fator de Gelman-Rubin, e “burn-in” para cada estrutura de dados.

Estrutura dos dados	$\hat{\sigma}_e^2$	LI	LS	$\sqrt{\hat{R}}$	Burn-in
Baixa	84,32	53,52	128,32	0,9918	562
Média	31,00	19,25	47,33	0,9994	320
Alta	37,99	26,35	58,97	0,9999	128

Realmente nota-se na Tabela 18 que a variância do erro foi maior na estrutura de dados relacionada com animais de baixa acurácia, e esta revelação apenas vem a confirmar discussões anteriores relacionadas com a maior variabilidade entre o comportamento das séries dos indivíduos desse arquivo. Os gráficos das cadeias MCMC para este parâmetro e suas respectivas densidades a posteriori são apresentados na Figura 18.

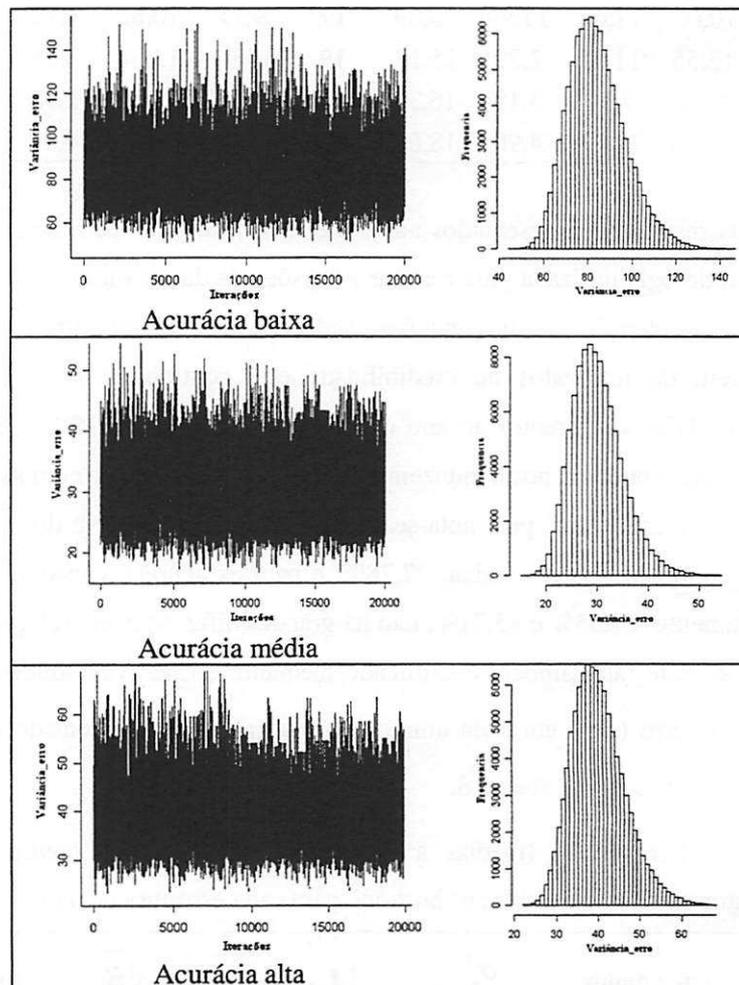


Figura 18. Cadeias geradas pelo método MCMC e densidade a posteriori para a variância do erro em cada estrutura de dados considerada.

5 CONCLUSÕES

De forma geral, a metodologia apresentada neste trabalho demonstrou qualidades que podem torná-la uma ferramenta importante na predição de DEP's de bovinos para sumários de anos futuros. Esta afirmação está fundamentada nos resultados obtidos via simulação de dados e análise de dados reais.

Os estudos de simulação demonstraram que é importante testar distribuições a priori mesmo quando o conhecimento a respeito dos parâmetros é mínimo, uma vez que prioris informativas (hierárquica Normal multivariada-Gama inversa e independentes t-Student multivariada-Gama inversa) superam a priori não informativa de Jeffreys em relação à qualidade de ajuste do modelo AR(2) para dados em painel.

A análise de dados de DEP's de touro Nelore em anos sucessivos confirmou os resultados discutidos na simulação, e foram constatadas boas estimativas para o valor de um dado futuro obtidas a partir da distribuição preditiva. Além disso, este processo possibilitou a detecção de diferenças nos comportamentos de séries referentes a indivíduos pertencentes a grupos distintos de acurácia, indicando que este fator deve ser levado em consideração. Mas estas diferenças não impõem barreiras para a não adoção desta metodologia na presença de dados provenientes de animais com baixa acurácia para DEP's, visto que a porcentagem de acerto na previsão de um dado futuro foi próxima a 80%.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions Automatic Control**, Tokyo, v. 19, n. 6, p. 716–723, Dec. 1974.
- ANSARI, A.; JEDIDI, K.; DUBE, L. Heterogeneous factor analysis models: a Bayesian approach. **Psychometrika**, New York, v. 67, n. 1, p. 49-78, Mar. 2002.
- ASAI, M. Comparison of MCMC Methods for Estimating Stochastic Volatility Models. **Computational Economics**, New York, v. 25, n. 4, p. 281-301, June 2005.
- BALTAGI, B. **Econometric analysis of panel data**. 2. ed. New York: USA: John Wiley and Sons, 2000. 314 p.
- BARRETO, G.; ANDRADE, M. G. Robust Bayesian Approach for AR(p) Models Applied to Streamflow Forecasting. **Journal Applied Statistical Science**, New York, v. 12, n. 3, p. 269-292, Mar. 2004.
- BERG, B. A. **Markov Chain Monte Carlo Simulations and Their Statistical Analysis**: With web-based fortran code. Singapore: World Scientific, 2004. 35 p.
- BERGER, J.; PERICCHI, L. Objective Bayesian methods for model selection: Introduction and comparison. In: LAHIRI, P. (Ed.). **Model selection**. Beachwood, HO: IMS, 2001. p. 135-207.
- BRINKS, J. S. **Expected progeny differences**. Denver, USA: Don-Arts Printers Colorado, 1990. 38 p.
- BOX, G. E. P.; JENKINS, G. M.; G. C. REINSEL. **Time series analysis: forecasting and control**. 3. ed. San Francisco, USA: Holden-Day, 1994. 500 p.
- BRANDT, P. T.; FREEMAN, J. R. Advances in Bayesian Time Series Modeling and the Study of Politics: Theory Testing, Forecasting, and Policy Analysis. **Political Analysis**, Oxford, v. 4, n. 1, p. 20–36, Jan. 2006.

- BRASIL. Ministério da Agricultura/Secretaria de Desenvolvimento Rural. Departamento de Tecnologia e Produção Animal. Coordenadoria de Melhoramento Animal. **Sumário de touros-arquivo zootécnico nacional – Gado de Corte**. Brasília, 1998.
- BROEMILING, L. D. **Bayesian Analysis of linear models**. New York, USA: John Wiley, 1989. 412 p.
- BROEMILING, L. D.; COOK, P. Bayesian estimation of the mean of an autoregressive process. **Journal of Applied Statistics**, Oklahoma, v. 20, n. 1, p. 25-38, Jan. 1993.
- CASELLA, G.; GEORGE, E. I. Explaining the Gibbs sampler. **The American Statistician**, Salt Lake, v. 46, n. 3, p. 167-174, Aug. 1992.
- CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hastings algorithm. **The American Statistician**, Salt Lake, v. 49, n. 4, p. 327-345, Nov. 1995.
- DAVIES, A.; LAHIRI, K. **Re-examining the rational expectations hypothesis using panel data on multi-period forecasts, analysis of panels and limited dependent variable models**. Cambridge, UK: Cambridge University Press, 2000, 154 p.
- de ALBA, E. Constrained forecasting in autoregressive time series models: A Bayesian analysis. **International Journal of Forecasting**, New York, v. 9, n. 1, p. 95-108, Apr. 1993.
- DUFOUR, J. M.; KIVIET, J. F. Exact Inference Methods for First-Order Autoregressive Distributed Lag Models. **Econometrica**, Tel Aviv, v. 66, n. 1, p. 79-104, Jan. 1998.
- EHLERS, R. S. **Introdução à inferência bayesiana**. Departamento de Estatística, UFPR. Disponível em: <<http://www.est.ufpr.br/~ehlers/notas>>. Acesso em: 21 nov. 2005.
- FALK, B.; ROY, A. Forecasting using the trend model with autoregressive errors. **International Journal of Forecasting**, London, v. 21, n. 4, p. 291-302, Oct./Dec. 2005.
- FREES, E. **Longitudinal and panel data**. Cambridge, UK: Cambridge University Press, 2004. 320 p.

GAMERMAN, D. Markov Chain Monte Carlo - stochastic simulation for bayesian inference. Londres, UK: Chapman & Hall, 1997. 245 p.

GELFAND, A. E. Model determination using sampling based methods. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Ed.). **Markov Chain Monte Carlo in Practice.** London, UK: Chapman and Hall, 1996. p. 145–162.

GELFAND, A. E.; SMITH, A. F. M. Sampling based approaches to calculating marginal densities. **Journal of the American Statistical Association,** Alexandria, v. 85, n. 4, p. 398-409, June 1990.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. Bayesian data analysis. Boca Raton, USA: Chapman and Hall/CRC, 2000. 526 p.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequence. **Statistical Science,** Hayward, v. 7, n. 4, p. 457-511, May 1992.

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.** New York, USA: Oxford University Press, 1992. p. 625-631.

GHOSH, M.; HEO, J. Default Bayesian priors for regression models with first-order autoregressive residuals. **Journal of Time Series Analysis,** New York, v. 24, n. 3, p. 269-282, May 2003.

GREENHOUSE, J. B.; WASSERMAN, L., Robust Bayesian analysis of clinical trials, **Statistics in Medicine,** Hoboken, v. 14, n. 12, p. 1379–1391, June 1995.

HAY, J. L.; PETTTTT, A. N. Bayesian analysis of a time series of counts with covariates: an application to the control of an infectious disease. **Biostatistics,** Oxford, v. 2, n. 4, p. 433-444, Dec. 2001.

HARRISON, P. J.; STEVENS, C. F. Bayesian forecasting. **Journal of the Royal Statistical Society (Series B),** London, v. 38, n. 3, 205-247, May 1976.

HECKMAN, J.; LEAMER, E. Handbook of econometrics. Amsterdam, Netherlands: Elsevier Science, 2001. v. 5, p. 744.

- HEIDELBERGER, P.; WELCH, P. Simulation run length control in the presence of an initial transient. **Operations Research**, Maryland, v. 31, n. 6, p. 1109-44, Nov./Dec. 1983.
- HENDERSON, C. R. **Applications of linear models in animal breeding**. Ontario, Canada: University of Guelph Press, 1984. 462 p.
- HIRANO, K. Semiparametric Bayesian inference in autoregressive panel data models. **Econometrica**, Tel Aviv, v. 70, n. 2, p. 781-799, Mar. 2002.
- HSIAO, C. **Analysis of panel data**. 2. ed. Cambridge, UK: Cambridge University Press, 2003. 382 p.
- HSIAO, C.; SUN, B. H. To pool or not to pool panel data. panel data econometrics: future directions. In: KRISHNAKUMAR, J.; RONCHETTI, E. (Ed.). **Papers in Honour of Professor Pietro Balestra**. Amsterdam: North Holland, 2000. p. 181-198.
- HUERTA, G.; WEST, M. Priors and component structures in autoregressive time series models. **Journal of the Royal Statistical Society (Serie B)**, London, v. 61, n. 4, p. 881-899, 1999.
- JEFFREYS, H. **Theory of Probability**. Oxford, UK: Clarendon Press, 1961.
- KASS, R. E.; RAFTERY, A. E. Bayes factors. **Journal of the American Statistical Association**, Alexandria, v. 90, n. 430, p. 773-795, June 1995.
- KIZILKAYA, K.; CARNIER, P.; ALBERA, A.; BITTANTE, G.; TEMPELMAN, R. Cumulative t-link threshold models for the genetic analysis of calving ease scores. **Genetic, Selection and Evolution**, Berlin, v. 35, n. 5, p. 489-512, May 2003.
- LAMBERT, P. C.; SUTTON, A. J.; BURTON, P. R.; ABRAMS, K. R.; JONES, D. R. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. **Statistics in Medicine**, Hoboken, v. 24, n. 15, 2401-2428, Aug. 2005.
- LIU, L. M.; TIAO, G. C. Random Coefficient first-order autoregressive model. **Journal of Econometrics**, New York, v. 13, n. 3, p. 305-325, 1980.
- MacNEIL, M. D. Genetic evaluation of the ratio of calf weaning weight to cow weight. **Journal of Animal Science**, Savoy, v. 83, n. 4, p. 794-802, Apr. 2005.

MARQUES, L. D. **Modelos dinâmicos com dados em painel: revisão de literatura.** Porto: Faculdade de Economia do Porto, 2000. 85 p. (Relatório Técnico).

MIGON, H. S.; HARRISON P. J. An application of non-linear Bayesian forecasting to television advertising. In: BERNARDO, J. M. et al. (Ed.). **Bayesian statistics.** 2 ed. New York: Oxford University Press, 1985. p. 681-696.

MILANI, L. L. **Análise Bayesiana do modelo de intervenção com erro arma.** 2000. 142 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de series temporais.** São Paulo: Edgard Blucher, 2004. 535 p.

NI, S.; SUN, D. Noninformative priors and frequentist risks of Bayesian estimators of vector-autoregressive models. **Journal of Econometrics,** New York, v. 115, n. 1, p. 159 -197, July 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. 142 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.

NORMAN, H. D.; POWELL, R. L.; WRIGHT, J. R.; SATTLER, C. G. Timeliness and Effectiveness of Progeny Testing Through Artificial Insemination. **Journal of Dairy Science,** Champaign v. 86, n. 4, p. 1513-1525, Apr. 2003.

O'HAGAN, A. **Kendall's advanced theory of statistics: bayesian inference.** New York, USA: Edward Arnold Press, 1994. v. 2b, 402 p.

PAI, J. S.; RAVISHANKER, N.; GELFAND, A. E. Bayesian analysis of concurrent time series with application to regional IBM revenue data. **Journal of Forecasting,** London, v. 13, n. 5, p. 463-479, Sept. 1994.

PEREIRA, J. C. C. **Melhoramento genético aplicado à produção animal.** Belo Horizonte: Imprensa Universitária-UFMG, 1996. 416 p.

POWELL, R. L.; NORMAN, H. D. Changes in predicted difference milk with increased reliability. **Journal of Dairy Science**, Champaign, v. 63, n. 10, p. 1972–1977, Oct. 1981.

POWELL, R. L.; SANDERS, A. H.; NORMAN, H. D. Stability of Genetic Evaluations for Active Artificial Insemination Bulls. **Journal of Dairy Science**, Champaign, v. 87, n. 8, p. 2614-2620, Aug. 2004.

R Development Core Team (2006). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<http://www.R-project.org>>. Acesso em: 2006.

RAFTERY, A. E.; LEWIS, S. How many iterations in the Gibbs sampler? In: BERNARDO, J. M. et al. (Ed.). **Bayesian statistics**. Oxford, USA: University Press, 1992. p. 763-773.

RAFTERY, A. Hypothesis testing and model selection via posterior simulation. In: GILKS, W. R.; SPIEGELHALTER, D. J.; RICHARDSON, S. (Ed.). **Practical Markov Chain Monte Carlo**. Oxford, USA: University Press, 1995.

ROSA, G. J. M. **Análise Bayesiana de Modelos Mistos Robustos via Amostrador de Gibbs**. 1998. 57 p. Tese (Doutorado em Estatística e Experimentação Agrônômica) – Universidade de São Paulo, Piracicaba, SP.

SAFADI, T.; MORETTIN, P. A Bayesian analysis of autoregressive models with normal random coefficients. **Journal of Statistical Computation and Simulation**, Blakburg, v. 73, n. 8, p. 563-573, Aug. 2003.

SAMANIEGO, J. R. L. **Abordagem Clássica e Bayesiana em Modelos Auto-Regressivos com Transformação de Box-Cox**. 2002. Dissertação (Mestrado em Estatística) - Universidade Federal de São Carlos, São Carlos, SP.

SCHWARZ, G. Estimating the dimension of a model. **Annals of Statistics**, Minneapolis, v. 6, n. 2, p. 461–464, May 1978.

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

SILVA, N. M. A.; MUNIZ, J. A.; SILVA, F. F.; AQUINO, L. H. 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.

SILVA, W. S. **Inferência Bayesiana para os modelos da classe ARCH com Potência Assimétrica**. 2006. 81 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.

SOWELLI, F. B. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. **Journal of Econometrics**, New York, v. 53, n. 1/3, p. 165-188, July/Sept. 1992.

SORENSEN, D.; GIANOLA, D. **Likelihood, bayesian and MCMC methods in quantitative genetics**. New York, USA: Springer, 2002. 770 p.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; van der LINDE, A. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society (Serie B)**, London, v. 64, n. 4, p. 583-639, 2002.

THOMPSON, J. R. **Simulation: a modeler's approach**. New York, USA: Wiley-Interscience Press, 1999. 328 p.

VERMAAK, J.; ANDRIEU, C.; DOUCET, A.; GODSILL, S. J. **Bayesian model selection of autoregressive processes**. Cambridge: Cambridge University Engineering Department, 2000. (Technical Report CUED/F-INFENG/TR, 360).

WEST, M.; HARRISON, J. **Bayesian forecasting and dynamic models**. 2 ed. New York, USA: Springer, 1997. 430 p.

WEST, M.; HARRISON, P. J.; POLE, A. **Bats - a user guide – bayesian analysis of time series - release 1. 3**. Warwick, USA: University of Warwick Press, 1998. 120 p.

YANG, R. Y.; BERGER, J. O. Estimation of a covariance matrix using the reference prior. **Annals of Statistics**, Minneapolis, v. 22, n. 3, p. 1195-1211, Sept. 1994.

ZELLNER, A. Models, prior information, and Bayesian analysis. **Journal of Econometrics**, New York, v. 75, n. 1, p. 51-68, Nov. 1996.

ZHOU, Y. Y.; ROY, A. Effect of Tapering on Accuracy of Forecasts Made With Stable Estimators of Vector Autoregressive Processes. **International Journal of Forecasting**, London, v. 22, n. 1, p. 169-180, Jan./Mar. 2006.

ANEXO A

Tabela 1A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “burn-in” (acurácia baixa).....	87
Tabela 2A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “burn-in” (acurácia média).....	88
Tabela 3A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “burn-in” (acurácia alta).....	90
Tabela 4A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “burn-in” (geral).....	91
Figura 1A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial (Partial ACF) para os indivíduos da estrutura de acurácia baixa.....	95
Figura 2A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial (Partial ACF) para os indivíduos da estrutura de acurácia média.....	97
Figura 3A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial (Partial ACF) para os indivíduos da estrutura de acurácia alta.....	100

Tabela 1A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “Burn-in” (acurácia baixa).

Ind.	$\hat{\phi}_1$	LI	LS	\sqrt{R}	Burn-in	$\hat{\phi}_2$	LI	LS	\sqrt{R}	Burn-in
1	-0,52	-0,67	0,23	0,9055	502	-0,46	-0,59	0,21	0,9955	600
2	-0,34	-0,54	0,44	0,9940	345	-0,51	-0,81	0,14	0,9900	401
3	-0,49	-0,80	0,12	0,9901	900	-0,68	-0,87	-0,12	0,9911	850
4	-0,34	-0,44	0,01	0,9970	300	-0,20	-0,71	0,20	0,9900	255
5	0,15	0,07	0,25	0,9999	204	-0,33	-0,69	0,02	0,9812	201
6	0,39	0,23	0,49	1,0811	50	-0,55	-0,77	-0,19	0,9811	35
7	-0,52	-0,63	0,27	0,9910	230	-0,25	-0,80	0,12	0,9900	203
8	-0,24	-0,39	-0,18	0,9921	508	-0,19	-0,49	-0,04	0,9874	405
9	-0,31	-0,45	-0,19	0,9902	400	-0,46	-0,71	-0,22	0,9802	390
10	0,37	0,08	0,62	1,0012	5	-0,18	-0,50	-0,04	0,9956	50
11	0,35	0,01	0,74	0,9912	2	-0,29	-0,54	0,21	0,9955	80
12	-0,60	-0,86	-0,42	0,9800	3	-0,52	-0,72	-0,22	0,9911	92
13	0,35	0,12	0,52	0,9944	14	-0,17	-0,79	-0,17	0,9931	132
14	0,03	-0,09	0,14	0,9932	901	-0,37	-0,82	-0,19	0,9920	400
15	0,44	0,23	0,53	0,9391	25	-0,55	-0,81	-0,26	0,9999	625
16	0,11	0,02	0,22	0,9956	20	-0,75	-0,75	-0,19	0,9911	700
17	-0,35	-0,46	0,40	0,9912	714	-0,49	-0,86	0,14	0,9847	987
18	0,11	-0,14	0,23	0,9945	500	-0,32	-0,76	0,23	1,0050	502
19	-0,45	-0,72	0,15	0,9912	896	-0,12	-0,47	0,27	0,9902	365
20	-0,31	-0,62	-0,16	1,0059	589	-0,21	-0,56	-0,01	1,0059	802
21	-0,26	-0,35	-0,09	0,9915	652	-0,31	-0,63	-0,17	0,9955	704
22	0,29	0,05	0,52	0,9945	12	-0,13	-0,42	-0,01	0,9900	605
23	-0,59	-0,68	-0,04	0,9919	4	-0,54	-0,78	-0,20	0,9964	80
24	-0,25	-0,49	0,07	0,9931	625	-0,43	-0,74	0,22	0,9927	741
25	-0,45	-0,66	-0,23	0,9312	2	-0,13	-0,29	-0,02	0,9805	47
26	-0,24	-0,46	0,03	0,9814	380	-0,31	-0,69	0,10	0,9911	502
27	0,34	0,04	0,68	0,9923	147	-0,25	-0,51	-0,04	0,9887	645
28	0,16	-0,99	0,36	0,9874	502	-0,26	-0,90	-0,23	0,98740	507
29	-0,38	-0,75	-0,04	0,9835	101	-0,14	-0,34	-0,01	0,9902	80
30	-0,23	-0,52	0,33	1,0113	700	-0,32	-0,60	0,02	0,9825	904
31	0,25	0,04	0,51	0,9918	201	-0,23	-0,79	-0,12	0,9905	328

Tabela 2A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “Burn-in” (acurácia média).

Ind.	$\hat{\phi}_1$	LI	LS	\sqrt{R}	Burn-in	$\hat{\phi}_2$	LI	LS	\sqrt{R}	Burn-in
1	0,83	-1,14	1,67	0,9874	530	-0,85	-1,82	0,31	0,9842	787
2	1,28	0,69	1,83	0,9994	606	-0,71	-0,85	-0,34	1,0024	878
3	0,81	-0,19	1,59	0,9922	648	-0,61	-0,93	0,09	1,0097	811
4	0,56	-0,53	1,34	0,9819	488	-0,45	-0,81	0,02	0,9935	852
5	0,57	0,08	1,42	1,0226	490	-0,68	-0,89	-0,38	1,0339	724
6	1,42	0,09	1,85	1,0214	580	-0,67	-0,88	-0,39	0,9998	1022
7	0,61	-0,36	1,04	1,0022	564	-0,48	-1,04	0,03	0,9968	1065
8	0,76	0,41	1,71	1,0052	532	-0,66	-0,83	-0,34	1,0081	667
9	1,15	0,42	1,80	1,0152	508	-0,44	-0,89	0,04	1,0103	692
10	1,03	0,50	1,52	1,0113	468	-0,67	-0,80	-0,39	1,0057	785
11	1,09	0,60	1,48	1,0001	463	-0,69	-0,84	-0,46	1,0341	720
12	1,33	-0,18	1,98	1,0146	412	-0,36	-0,83	0,02	0,9867	804
13	0,71	-0,68	0,95	0,9877	450	-0,41	-0,96	0,12	1,0075	819
14	1,08	-0,69	1,26	1,0271	593	-0,64	-1,03	0,04	0,9952	760
15	0,59	0,02	1,07	1,0171	474	-0,67	-1,09	0,18	0,9889	782
16	1,17	0,77	0,75	0,9922	521	-0,63	-0,91	-0,28	1,0196	763
17	0,60	0,33	1,73	1,0170	416	-0,85	-1,01	-0,33	0,9695	703
18	1,10	-0,81	1,83	0,9862	707	-0,21	-0,84	0,37	0,9982	899
19	0,94	0,72	1,21	1,0056	655	-0,60	-0,95	-0,41	0,9937	723
20	0,86	0,65	1,18	0,9656	523	0,45	-0,12	0,95	0,9872	810
21	0,95	0,04	1,13	0,9913	664	-0,78	-0,95	-0,48	1,0125	841
22	0,99	0,55	1,33	0,9770	353	-0,60	-0,86	-0,37	1,0087	793
23	0,95	-0,51	1,55	1,0207	446	-0,21	-0,92	0,48	0,9787	755
24	1,20	-0,44	1,51	1,0044	460	-0,57	-0,80	-0,22	0,9771	795
25	0,83	0,64	1,36	0,9802	499	-0,64	-0,84	-0,27	0,9956	717
26	1,43	1,07	1,75	0,9906	471	-0,65	-0,89	-0,39	0,9936	820
27	1,30	0,40	1,61	1,0055	543	-0,08	-0,96	0,50	1,0190	953
28	1,29	1,16	1,37	0,9999	472	-0,61	-0,90	-0,34	1,0198	868
29	1,06	0,80	1,14	1,0055	530	-0,07	-0,77	0,45	0,9837	689
30	0,64	0,26	1,40	1,0176	777	-0,66	-0,96	-0,48	0,9909	873
31	0,61	0,33	1,27	0,9930	525	-0,49	-0,89	-0,32	1,0130	833
32	1,09	0,44	1,48	0,9536	305	-0,35	-0,97	0,02	1,0257	827
33	0,57	0,38	1,62	0,9722	463	-0,48	-0,93	-0,02	0,9952	820
34	0,78	-0,85	1,20	0,9988	565	-0,04	-0,48	0,79	1,0058	732

Continuação da Tabela 2A

35	0,90	0,62	1,26	0,9912	496	-0,83	-0,76	-0,53	0,9795	831
36	-0,05	-0,50	1,10	1,0295	532	-0,23	-0,75	0,26	1,0235	825
37	0,63	0,48	1,00	1,0022	465	-0,58	-0,92	-0,33	1,0270	956
38	0,48	0,00	0,94	1,0282	533	-0,65	-0,95	-0,45	0,9983	780
39	0,92	-0,07	0,62	1,0344	343	-0,01	-0,96	0,32	1,0196	785
40	0,58	0,22	0,84	0,9944	465	-0,51	-0,85	-0,34	0,9979	970
41	1,31	0,02	1,99	0,9970	361	-0,59	-0,84	-0,43	0,9921	722
42	1,02	0,31	1,74	0,9887	454	-0,65	-0,91	-0,36	1,0037	761
43	-0,13	-0,21	-0,02	1,0126	421	-0,40	-0,90	-0,02	0,9880	756
44	1,53	1,05	1,94	0,9788	351	-0,33	-0,81	-0,07	0,9803	615
45	0,70	0,01	1,60	1,0013	465	-0,41	-0,85	-0,14	1,0062	721
46	0,70	0,08	1,78	0,9760	473	-0,55	-0,96	-0,13	1,0383	761
47	0,45	0,00	1,14	1,0134	550	-0,3	-0,92	-0,48	1,0036	952
48	0,64	0,33	1,06	1,0345	603	-0,67	-0,99	-0,44	0,9925	619
49	0,74	-0,45	1,37	0,9915	675	-0,12	-0,84	0,49	1,0194	972
50	0,68	-0,06	1,61	0,9847	540	-0,09	-0,84	0,51	0,9476	703
51	0,78	0,28	1,45	0,9981	435	-0,51	-0,82	-0,32	0,9919	846
52	1,01	0,39	1,41	1,0118	424	-0,41	-0,83	-0,19	1,0285	644
53	0,97	0,32	1,80	1,0174	581	-0,47	-0,91	-0,11	0,9841	794
54	0,77	-0,47	1,70	0,9960	666	-0,45	-0,93	-0,09	1,0191	795
55	1,04	0,37	1,45	1,0081	376	-0,55	-0,93	-0,30	1,0170	845
56	0,90	0,71	1,16	1,0067	405	-0,59	-0,89	-0,23	1,0262	911
57	1,33	-1,32	1,23	0,9735	626	-0,43	-0,98	-0,02	1,0143	762
58	0,97	0,21	2,07	1,0259	743	-0,62	-1,07	-0,41	0,9941	939
59	0,78	0,23	1,94	1,0161	682	-0,63	-0,82	-0,35	0,9638	725
60	0,96	0,66	1,14	1,0252	495	-0,60	-0,76	-0,32	1,0060	1002
61	0,42	0,08	1,55	1,0202	485	-0,45	-0,83	-0,34	0,9702	800
62	0,27	0,09	0,61	1,0026	385	-0,65	-0,92	-0,42	0,9821	572
63	1,37	0,92	1,72	1,0160	562	-0,68	-0,88	-0,44	0,9886	1028

Tabela 3A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “Burn-in” (acurácia alta).

Ind.	$\hat{\phi}_1$	LI	LS	\sqrt{R}	Burn-in	$\hat{\phi}_2$	LI	LS	\sqrt{R}	Burn-in
1	1,49	0,66	2,64	1,0376	200	-0,68	-0,96	-0,38	1,0034	239
2	1,06	-0,39	1,57	1,0115	469	-0,13	-0,51	0,27	0,9937	307
3	1,55	0,81	2,04	1,0098	59	-0,22	-0,62	-0,03	0,9785	288
4	1,01	0,29	1,15	1,0108	72	-0,48	-0,93	0,07	0,9707	209
5	1,32	-0,05	1,98	1,0229	115	-0,41	-0,65	-0,23	0,9882	240
6	1,07	0,73	1,56	1,0724	70	-0,55	-0,90	-0,23	0,9687	200
7	0,96	0,14	1,35	0,9894	239	-0,57	-0,85	-0,26	0,9929	248
8	1,47	0,91	1,85	1,0357	128	-0,52	-0,90	0,38	1,0319	315
9	0,64	0,24	0,96	0,9961	285	-0,18	-0,45	-0,07	1,0048	134
10	1,00	0,58	1,24	1,0238	165	-0,41	-0,98	-0,15	1,0048	169
11	0,22	-0,43	1,71	1,0483	459	-0,12	-0,61	0,40	1,0033	383
12	1,02	0,85	1,23	1,0312	196	-0,44	-0,85	-0,11	1,0203	7
13	0,23	-0,26	0,85	1,0086	238	-0,24	-0,49	-0,01	1,0090	334
14	1,01	0,82	1,29	1,0538	111	-0,32	-0,58	-0,18	0,9769	240
15	0,89	0,53	1,03	1,0037	163	-0,38	-0,57	-0,26	0,9967	143
16	1,03	-0,41	1,53	1,0501	156	-0,20	-0,30	-0,02	1,0029	231
17	0,58	0,13	0,98	1,0084	179	-0,61	-0,94	-0,23	0,9966	38
18	1,03	-0,49	1,96	0,9985	303	-0,38	-0,71	0,00	1,0462	128
19	0,27	0,03	0,51	1,0216	123	-0,57	-0,95	-0,03	1,0007	219
20	0,73	0,08	1,02	1,0476	231	-0,22	-0,63	-0,04	1,0037	230
21	0,13	-0,19	0,53	1,0126	395	-0,22	-0,91	0,28	1,0071	402
22	0,38	0,03	0,96	1,0750	115	-0,51	-0,91	-0,13	0,9958	154
23	0,35	-0,14	0,81	1,0374	231	-0,38	-0,74	-0,14	0,9776	150

Tabela 4A. Estimativas para os parâmetros ϕ_1 e ϕ_2 , intervalo de credibilidade (95%), fator de Gelman e Rubin e “burn-in” (geral).

Ind.	$\hat{\phi}_1$	LI	LS	\sqrt{R}	Burn-in	$\hat{\phi}_2$	LI	LS	\sqrt{R}	Burn-in
1	-0,51	-0,66	0,23	0,8918	742	-0,45	-0,58	0,21	0,9805	949
2	-0,33	-0,53	0,43	0,9739	507	-0,50	-0,79	0,14	0,9699	649
3	-0,45	-0,74	0,11	0,9140	1246	-0,63	-0,80	-0,11	0,9149	1595
4	-0,38	-0,49	0,01	1,1152	503	-0,22	-0,79	0,22	1,1074	644
5	0,13	0,06	0,22	0,8984	275	-0,30	-0,62	0,02	0,8816	352
6	0,35	0,21	0,44	0,9806	68	-0,50	-0,70	-0,17	0,8899	87
7	-0,53	-0,64	0,27	1,0066	350	-0,25	-0,81	0,12	1,0056	449
8	-0,25	-0,40	-0,18	1,0179	782	-0,19	-0,50	-0,04	1,0131	1001
9	-0,30	-0,44	-0,19	0,9679	587	-0,45	-0,69	-0,22	0,9582	751
10	0,35	0,08	0,59	0,9449	7	-0,17	-0,47	-0,04	0,9396	9
11	0,40	0,01	0,84	1,1253	3	-0,33	-0,61	0,24	1,1302	4
12	-0,54	-0,77	-0,38	0,8814	4	-0,47	-0,65	-0,20	0,8913	5
13	0,38	0,13	0,57	1,0928	23	-0,19	-0,87	-0,19	1,0914	30
14	0,03	-0,08	0,12	0,8789	1196	-0,33	-0,73	-0,17	0,8779	1531
15	0,42	0,22	0,51	0,9051	36	-0,53	-0,78	-0,25	0,9637	46
16	0,11	0,02	0,22	0,9816	30	-0,74	-0,74	-0,19	0,9772	38
17	-0,33	-0,44	0,38	0,9429	1019	-0,47	-0,82	0,13	0,9367	1304
18	0,11	-0,14	0,23	0,9836	742	-0,32	-0,75	0,23	0,9940	949
19	-0,43	-0,69	0,14	0,9558	1296	-0,12	-0,45	0,26	0,9548	1659
20	-0,32	-0,64	-0,17	1,0429	916	-0,22	-0,58	-0,01	1,0429	1172
21	-0,25	-0,34	0,09	0,9591	946	-0,30	-0,61	0,16	0,9630	1211
22	0,31	0,95	-0,21	1,0631	19	-0,14	-0,45	0,11	1,0583	25
23	-0,52	-0,59	-0,03	0,8667	5	-0,47	-0,68	-0,17	0,8707	7
24	-0,27	-0,54	0,08	1,0868	1026	-0,47	-0,81	0,24	1,0864	1313
25	-0,47	-0,69	-0,24	0,9730	3	-0,14	-0,30	-0,02	1,0245	4
26	-0,26	-0,49	0,03	1,0453	607	-0,33	-0,73	-0,11	1,0557	777
27	0,33	0,04	0,65	0,9528	212	-0,24	-0,49	-0,04	0,9494	271
28	0,14	-0,84	0,31	0,8372	638	-0,22	-0,76	-0,20	0,8372	817
29	-0,36	-0,72	-0,04	0,9410	145	-0,13	-0,33	-0,01	0,9474	186
30	-0,24	-0,54	0,34	1,0490	1089	-0,33	-0,62	0,02	1,0191	1394

Continuação da Tabela 4A

31	0,26	0,04	0,52	1,0144	308	-0,24	-0,81	-0,12	1,0130	395
32	0,79	-1,08	1,59	0,9393	756	-0,81	-1,73	0,29	0,9363	968
33	1,18	0,64	1,69	0,9230	840	-0,66	-0,79	-0,31	0,9258	1075
34	0,76	-0,18	1,49	0,9297	911	-0,57	-0,87	0,08	0,9461	1166
35	0,57	-0,54	1,36	0,9937	741	-0,46	-0,82	0,02	1,0055	948
36	0,57	0,08	1,42	1,0220	735	-0,68	-0,89	-0,38	1,0333	940
37	1,74	0,11	2,27	1,2539	1068	-0,82	-1,08	-0,48	1,2273	1367
38	0,59	-0,35	1,01	0,9766	824	-0,47	-1,01	0,03	0,9713	1055
39	0,79	0,42	1,77	1,0405	826	-0,68	-0,86	-0,35	1,0435	1057
40	0,92	0,34	1,44	0,8122	610	-0,35	-0,71	0,03	0,8083	780
41	1,04	0,51	1,54	1,0244	711	-0,68	-0,81	-0,40	1,0187	910
42	1,10	0,60	1,49	1,0071	699	-0,69	-0,85	-0,46	1,0413	895
43	1,36	-0,18	2,02	1,0338	630	-0,37	-0,85	0,02	1,0054	806
44	0,66	-0,63	0,89	0,9214	630	-0,38	-0,90	0,11	0,9399	806
45	1,08	-0,69	1,26	1,0275	890	-0,64	-1,03	0,04	0,9956	1139
46	0,52	0,02	0,94	0,8897	622	-0,59	-0,95	0,16	0,8651	796
47	1,02	0,67	0,66	0,8666	683	-0,55	-0,79	-0,24	0,8905	874
48	0,58	0,32	1,68	0,9895	607	-0,83	-0,98	-0,32	0,9433	777
49	0,95	-0,70	1,58	0,8492	913	-0,18	-0,72	0,32	0,8595	1169
50	0,87	0,67	1,12	0,9332	912	-0,56	-0,88	-0,38	0,9222	1167
51	0,81	0,61	1,11	0,9048	735	0,42	-0,11	0,89	0,9251	941
52	1,04	0,04	1,24	1,0899	1095	-0,86	-1,04	-0,53	1,1132	1402
53	0,98	0,54	1,32	0,9678	524	-0,59	-0,85	-0,37	0,9992	671
54	0,93	-0,50	1,51	0,9950	652	-0,20	-0,90	0,47	0,9541	835
55	1,16	-0,43	1,46	0,9740	669	-0,55	-0,78	-0,21	0,9476	856
56	0,96	0,74	1,57	1,1331	865	-0,74	-0,97	-0,31	1,1509	1108
57	1,40	1,05	1,71	0,9690	691	-0,64	-0,87	-0,38	0,9719	885
58	1,22	0,38	1,51	0,9448	765	-0,08	-0,90	0,47	0,9575	980
59	1,34	1,21	1,43	1,0413	737	-0,64	-0,94	-0,35	1,0620	944
60	1,15	0,87	1,24	1,0917	863	-0,08	-0,84	0,49	1,0680	1105
61	0,67	0,27	1,47	1,0697	1225	-0,69	-1,01	-0,50	1,0416	1568
62	0,58	0,31	1,20	0,9414	747	-0,46	-0,84	-0,30	0,9603	956
63	0,98	0,39	1,32	0,8532	409	-0,31	-0,87	0,02	0,9177	524
64	0,56	0,37	1,59	0,9570	684	-0,47	-0,92	0,02	0,9796	875

Continuação da Tabela 4A

65	0,75	-0,82	1,15	0,9587	814	-0,04	-0,46	0,76	0,9655	1041
66	0,90	0,62	1,26	0,9906	744	-0,83	-0,76	-0,53	0,9789	952
67	-0,06	-0,60	1,33	1,2454	965	-0,28	-0,91	0,31	1,2382	1236
68	0,55	0,42	0,87	0,8756	609	-0,51	-0,80	-0,29	0,8973	780
69	0,51	0,00	1,00	1,0985	854	-0,69	-1,01	-0,48	1,0665	1093
70	1,02	-0,08	0,69	1,1465	570	-0,01	-1,06	0,35	1,1301	730
71	0,62	0,23	0,89	1,0567	741	-0,54	-0,90	-0,36	1,0605	949
72	1,31	0,02	1,98	0,9944	540	-0,59	-0,84	-0,43	0,9895	691
73	1,09	0,33	1,86	1,0571	728	-0,69	-0,97	-0,38	1,0731	932
74	-0,12	-0,25	0,08	0,9197	574	-0,36	-0,75	0,03	0,8974	734
75	1,58	1,09	2,01	1,0125	545	-0,34	-0,84	-0,07	1,0141	697
76	0,78	0,01	1,77	1,1102	773	-0,45	-0,94	-0,16	1,1156	990
77	0,73	0,08	1,87	1,0236	744	-0,58	-0,95	-0,14	1,0889	952
78	0,49	-0,06	1,25	1,1086	903	-0,33	-0,98	0,13	1,0979	1155
79	0,61	0,32	1,01	0,9887	864	-0,64	-0,95	-0,42	0,9486	1107
80	0,73	-0,44	1,34	0,9719	992	-0,12	-0,82	0,48	0,9992	1270
81	0,67	-0,06	1,60	0,9756	803	-0,09	-0,83	0,51	0,9388	1027
82	0,85	0,30	1,57	1,0824	708	-0,55	-0,89	-0,35	1,0757	906
83	1,05	0,40	1,46	1,0507	660	-0,43	-0,86	-0,20	1,0680	845
84	0,98	0,32	1,82	1,0283	881	-0,48	-0,92	-0,11	0,9947	1128
85	0,70	-0,43	1,54	0,9046	907	-0,41	-0,84	-0,08	0,9255	1161
86	1,00	0,35	1,39	0,9659	540	-0,53	-0,89	-0,29	0,9744	692
87	0,97	0,76	1,24	1,0801	652	-0,63	-0,95	-0,25	1,1011	834
88	1,30	-1,29	1,20	0,9517	918	-0,42	-0,96	-0,02	0,9916	1175
89	0,98	0,21	2,10	1,0404	1130	-0,63	-1,09	-0,42	1,0081	1447
90	0,69	0,20	1,70	0,8928	899	-0,55	-0,72	-0,31	0,8469	1151
91	0,90	0,62	1,07	0,9587	694	-0,56	-0,71	-0,30	0,9407	889
92	0,42	-0,08	1,54	1,0106	721	-0,45	-0,82	0,34	0,9611	922
93	0,25	0,08	0,56	0,9279	534	-0,60	-0,85	-0,39	0,9089	684
94	1,39	0,93	1,75	1,0321	856	-0,69	-0,89	-0,45	1,0042	1096
95	1,58	0,70	2,81	1,1025	319	-0,72	-1,02	-0,40	1,0661	408
96	1,10	-0,41	1,63	1,0533	733	-0,14	-0,53	0,28	1,0348	938
97	1,38	0,72	1,82	0,8986	79	-0,20	-0,55	-0,03	0,8707	101

Continuação da Tabela 4A

98	1,02	0,29	1,16	1,0180	109	-0,48	-0,94	0,07	0,9776	139
99	1,28	-0,05	1,92	0,9940	168	-0,40	-0,63	-0,22	0,9602	215
100	0,95	0,65	1,39	0,9568	94	-0,49	-0,80	-0,21	0,8642	120
101	0,99	0,14	1,40	1,0228	371	-0,59	-0,88	-0,27	1,0264	474
102	1,46	0,90	1,84	1,0296	191	-0,52	-0,89	0,38	1,0258	244
103	0,69	0,26	1,04	1,0741	461	-0,19	-0,49	-0,08	1,0835	590
104	0,95	0,55	1,18	0,9769	236	-0,39	-0,94	-0,14	0,9588	302
105	0,21	-0,42	1,65	1,0144	666	-0,12	-0,59	0,39	0,9709	853
106	0,97	0,81	1,17	0,9796	279	-0,42	-0,81	-0,10	0,9693	357
107	0,23	-0,26	0,85	1,0102	358	-0,24	-0,49	-0,01	1,0106	458
108	0,94	0,76	1,19	0,9756	154	-0,30	-0,54	-0,17	0,9044	197
109	0,72	0,43	0,84	0,8146	198	-0,31	-0,46	-0,21	0,8089	254
110	1,13	-0,45	1,68	1,1504	256	-0,22	-0,33	-0,02	1,0987	328
111	0,65	0,15	1,10	1,1333	302	-0,69	-1,06	-0,26	1,1200	386
112	1,16	-0,55	2,20	1,1219	511	-0,43	-0,80	0,00	1,1755	654
113	0,26	-0,03	0,68	0,9649	174	-0,54	-0,90	0,08	0,9452	223
114	0,69	0,08	0,96	0,9877	327	-0,21	-0,59	-0,04	0,9463	418
115	0,14	-0,20	0,56	1,0719	627	-0,23	-0,96	0,30	1,0660	803
116	0,35	-0,13	0,89	0,9916	159	-0,47	-0,84	0,02	0,9185	204
117	0,34	-0,13	0,78	0,9978	333	-0,37	-0,71	-0,13	0,9403	427

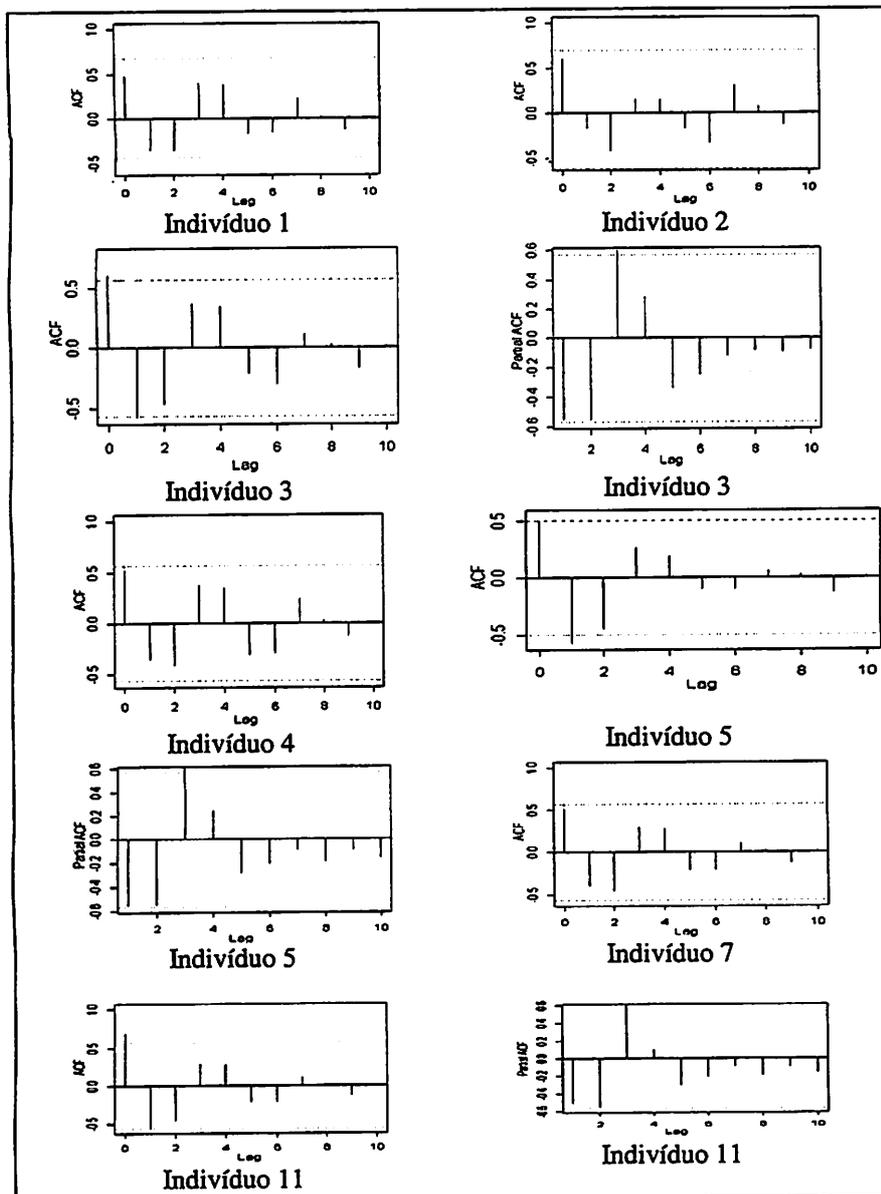


Figura 1A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial ("Partia. ACF"), esta última para as séries em que não se constatou o ruído branco, considerando a estrutura de acurácia baixa.

Continuação da Figura 1A

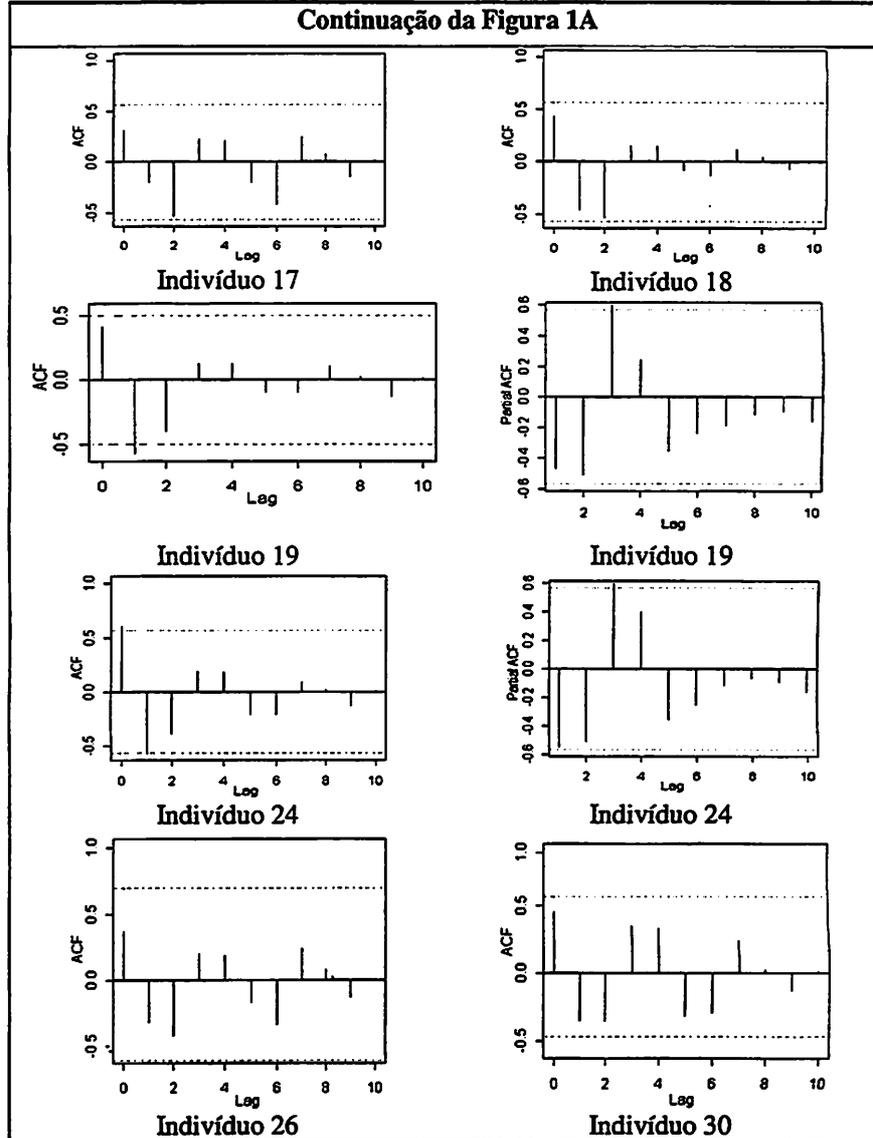


Figura 1A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial ("Partial ACF"), esta última para as séries em que não se constatou o ruído branco, considerando a estrutura de acurácia baixa.

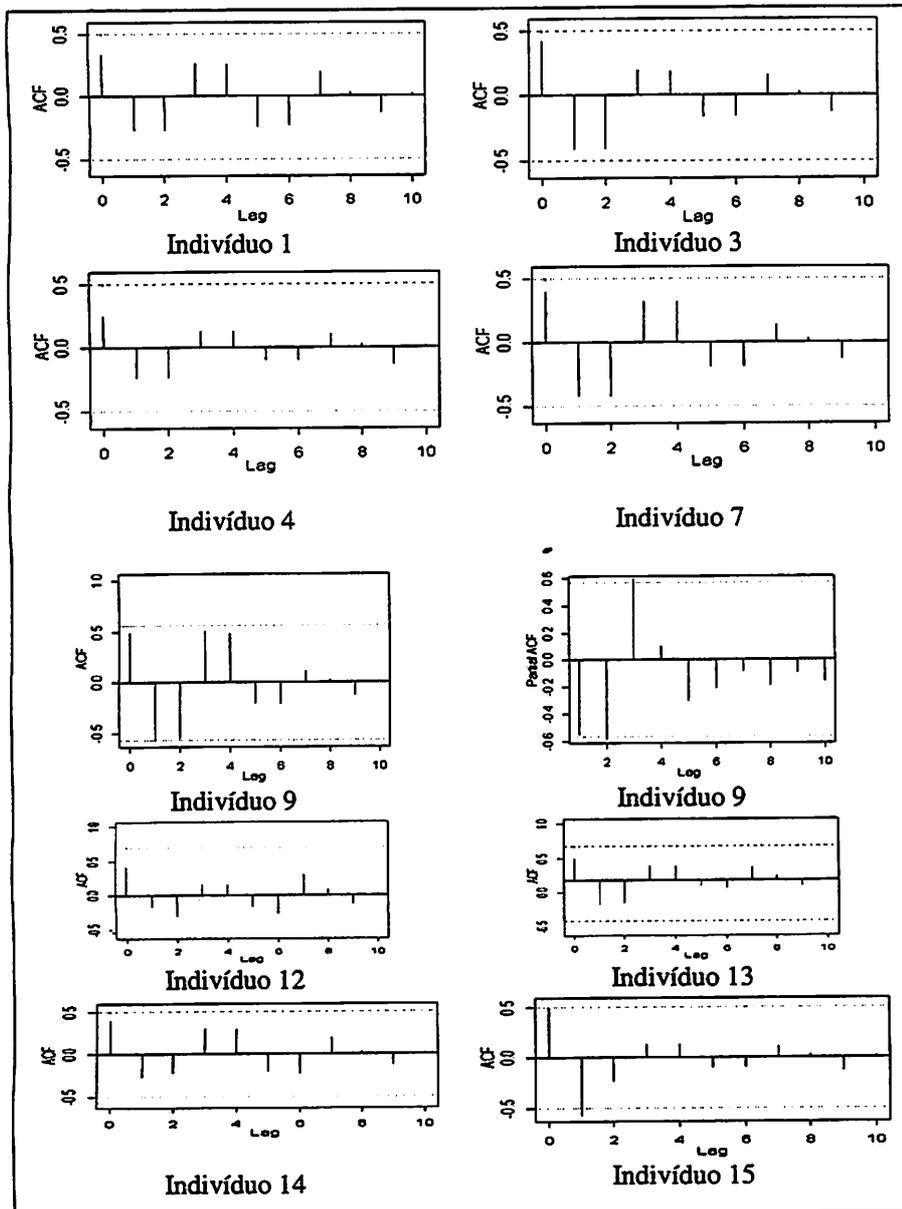


Figura 2A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial ("Partial ACF"), esta última para as séries em que não se constatou o ruído branco, considerando a estrutura de acurácia média.

Continuação da Figura 2A

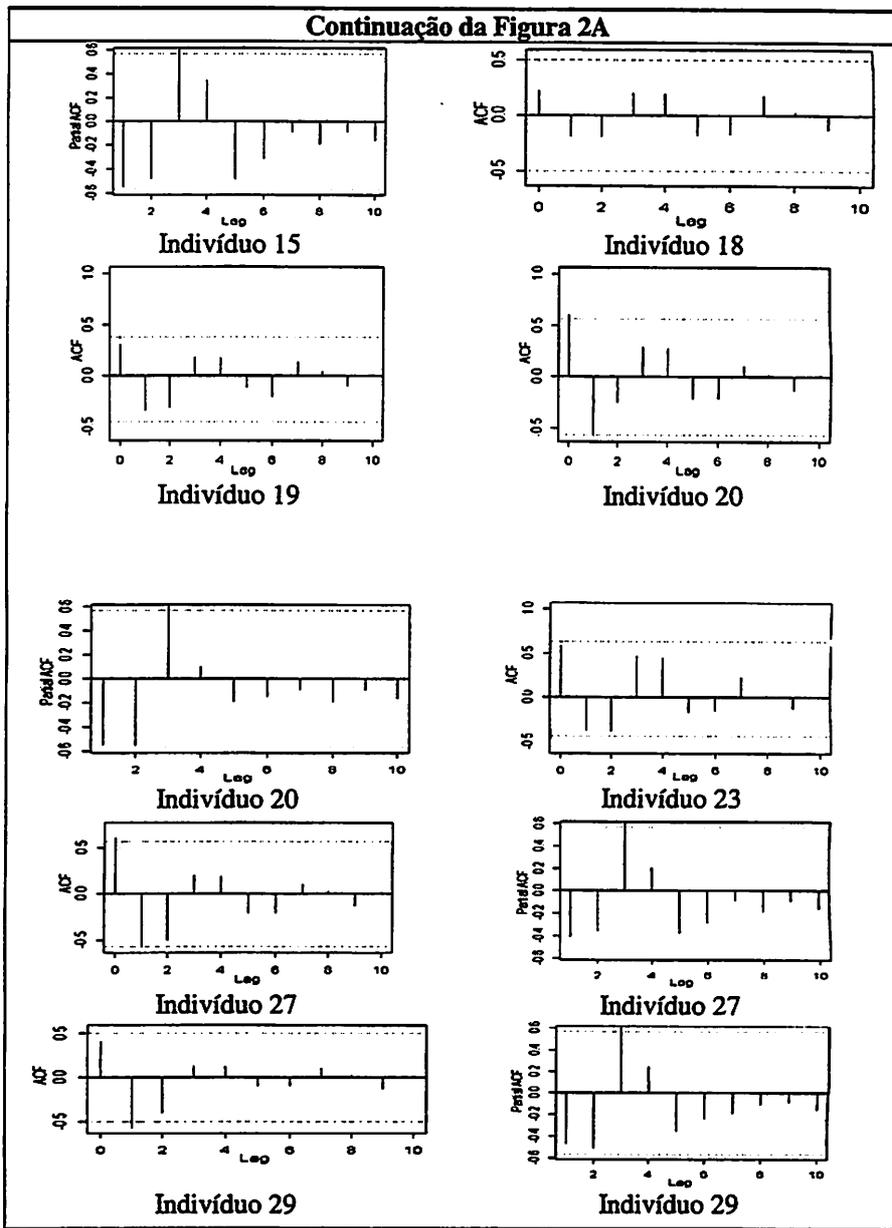


Figura 2A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial ("Partial ACF"), esta última para as séries em que não se constatou o ruído branco, considerando a estrutura de acurácia média.

Continuação da Figura 2A

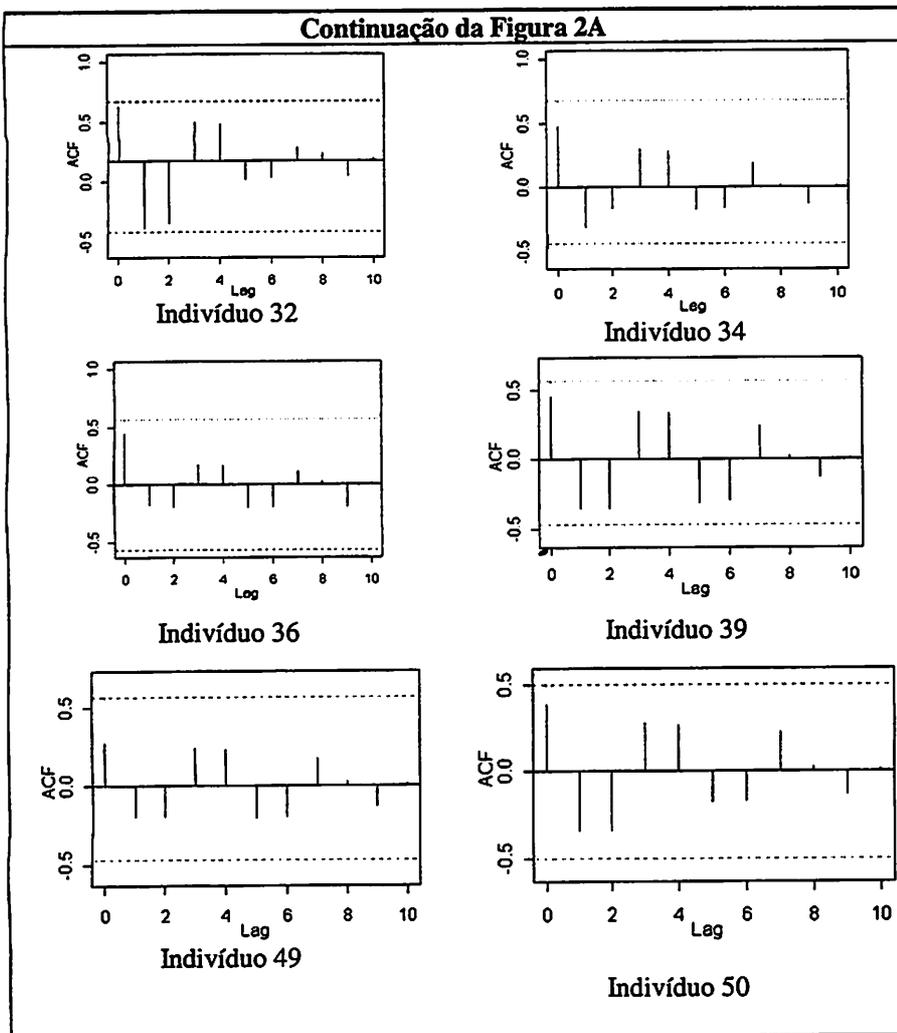


Figura 2A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial (“Partial ACF”), esta última para as séries em que não se constatou o ruído branco, considerando a estrutura de acurácia média.

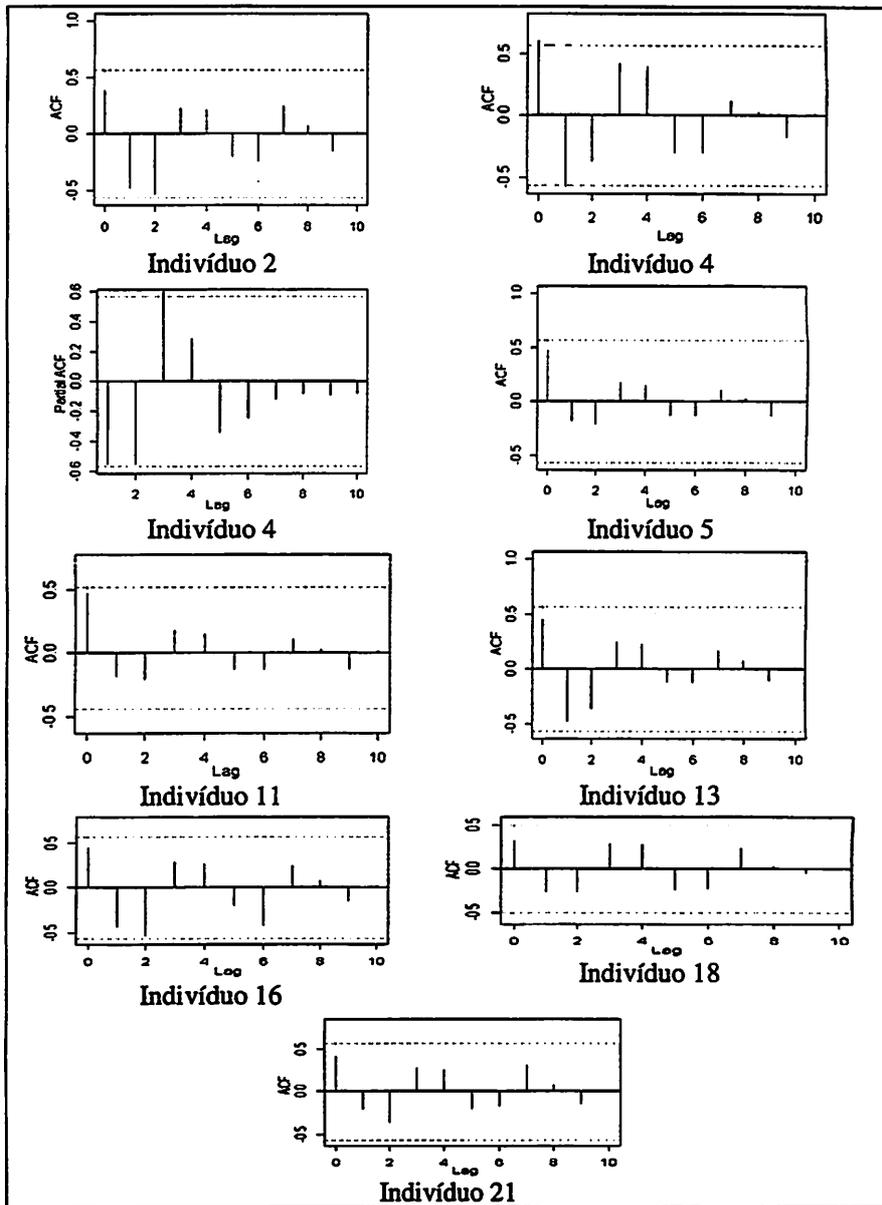


Figura 3A. Gráficos das Funções de autocorrelação (ACF) e autocorrelação parcial (“Partial ACF”), esta última para as séries em que não se constatou o ruído branco, considerando a estrutura de acurácia alta.