



DEDALUS - Acervo - ICMSC



30300036889

**ANÁLISE PARA MODELOS DE SÉRIES
TEMPORAIS AR(p) USANDO OS ALGORITMOS
GIBBS SAMPLING E METROPOLIS-HASTINGS**

**MARINHO G. ANDRADE
CLAUDIA F.F. HUTTER
JOSÉ ROBERTO T. OLIVEIRA**

Nº 64

RELATÓRIOS TÉCNICOS DO ICMSC

São Carlos
Dez./1997

RT

Class.	- SCE
Cutt.	- me 64
	A553ap
Tembo-	21.568

Análise para Modelos de Séries Temporais $AR(p)$ Usando os Algoritmos Gibbs Sampling e Metropolis-Hastings

Marinho G. Andrade, Claudia F.F. Hutter, José Roberto T. Oliveira
USP - ICMSC - SCE - CP: 668-13.560-970 - São Carlos - SP

Resumo

Neste trabalho apresentamos a análise Bayesiana de modelos de séries temporais autorregressivos de ordem p $AR(p)$, considerando três alternativas de densidades a priori para os parâmetros. Primeiramente consideramos a priori não informativa de Jeffreys, onde a densidade a posteriori marginal, para os parâmetros do modelo, pode ser calculada analiticamente. No segundo caso, adotamos uma função densidade a priori conjugada normal-gama. No terceiro caso adotamos uma densidade a priori informativa t -Student, p -dimensional, para os parâmetros, e uma priori gama para o inverso da variância dos resíduos. Neste caso a análise a posteriori só pode ser feita usando-se algoritmos de simulação em cadeia de Markov, MCMC. Nos três casos estudados, utilizamos o algoritmo de simulação MCMC para fazer a previsão de valores futuros da série para vários passos a frente e também para aplicar o critério de seleção de modelos utilizando a densidade preditiva.

1. Introdução

A análise de modelos de séries temporais é uma ferramenta muito importante em várias áreas, tais como, engenharia, economia e outras. A inferência desses modelos é feita geralmente maximizando-se a função de verossimilhança para os dados [Box, Jenkins e Reinsel, 1994]. Uma abordagem alternativa proposta na literatura é a análise bayesiana [Broemeling e Land - 1988]. A vantagem do uso dos métodos bayesianos é permitir a consideração de informações a priori sobre os parâmetros do modelo na análise. Apesar disso a análise bayesiana geralmente leva a problemas numéricos de difícil solução. No entanto com o surgimento de algoritmos de simulação de Monte-Carlo baseado em Cadeias de Markov, essa técnica tomou novo impulso [Casela e George-1992].

O uso desses algoritmos para ajuste de modelos de séries temporais vem sendo feito com sucesso [McCulloch e Tsay -1994], [de Alba -1993].

O objetivo deste trabalho é fazer a análise Bayesiana de modelos de séries temporais quando diferentes densidades a priori são adotadas destacando as características particulares de cada modelo. Na análise bayesiana estudamos os modelos auto-regressivos considerando três alternativas de densidades a priori para os parâmetros deste modelo. Primeiramente consideramos a priori não informativa de Jeffreys, onde a densidade a posteriori marginal, para os parâmetros do modelo, pode ser calculada analiticamente e mostra-se que o valor esperado dessa posteriori coincide com o estimador de máxima verossimilhança. No segundo caso, adotamos uma função densidade a priori conjugada normal-gama. Aqui, as densidades a posteriori também podem ser calculadas analiticamente, resultando em uma densidade t -Student p -dimensional para os parâmetros do modelo e uma densidade gama para o inverso da variância dos resíduos. A análise da expressão do estimador de Bayes, para os parâmetros do modelo, mostra que este estimador é uma média ponderada entre o estimador de máxima verossimilhança e o valor esperado da densidade a priori.

O uso de densidades a priori conjugadas em geral recebe críticas por serem estas funções pouco realista para tratamento de problemas reais. Para contornar estas críticas adotamos no terceiro caso uma densidade a priori informativa t -Student, p -dimensional, para os parâmetros e uma densidade a priori gama para o inverso da variância dos resíduos. Isto resulta em uma densidade a posteriori não padronizada. Neste caso a análise a posteriori só pode ser feita numericamente. Aqui optamos pelo uso de algoritmos de simulação de Monte Carlo em Cadeia de Markov (MCMC), do tipo Metropolis-Hastings [Chib e Greenberg - 1995] e Gibbs Sampling [Casella e George - 1992] os quais vêm se destacando com muita eficiência neste tipo de abordagem, porém, exigindo para isso considerações práticas.

Um critério de seleção de modelos baseado na densidade preditiva é aplicado. Esses resultados são comparados com a análise clássica onde a identificação do modelo é feita através da função de autocorrelação, FAC e função de autocorrelação parcial, FACP, a escolha do melhor modelo para um conjunto de dados é feita usando-se o Critério de Informação de Akaike AIC e o Critério de Informação Bayesiano BIC.

Este trabalho está organizado da seguinte forma: na seção 2 apresentamos o desenvolvimento da função de máxima verossimilhança para um Processo $AR(p)$. Na seção 3 apresentamos a abordagem bayesiana para as três priori e uma análise preditiva com as priori estudadas neste trabalho. Na seção 4 apresentamos as densidades condicionais deduzidas da abordagem bayesiana e o algoritmo MCMC para os modelos $AR(p)$. Na seção 5 é apresentado um estudo numérico para dois conjuntos de dados gerados e um conjunto de dados real, representando a componente estocástica de vazões médias mensais em (m^3/s) que chegam ao reservatório da hidroelétrica de Furnas, no sudeste do Brasil.

2. Processo Auto-regressivo, AR (p)

Considere a série temporal Z_t , $t = 1, \dots, N$ gerada por um processo auto-regressivo de ordem p , AR(p):

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t \quad (2.1)$$

onde $\phi \in R^p$ é um vetor $p \times 1$, $\{a_t, t = 1, 2, \dots\}$ é uma sequência de variáveis aleatórias i.i.d. $N(0, \tau^{-1})$ e $\tau = 1/\sigma^2 > 0$. Vamos supor agora um conjunto de N observações para essa série que denotamos por $Z = (Z_{p+1}, \dots, Z_N)'$, e vamos denotar por $Z_p = (Z_1, \dots, Z_p)'$ as p primeiras observações do processo. Escrevendo o modelo (2.1) como:

$$a_t = Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} \quad (2.2)$$

Podemos escrever que a densidade de probabilidade conjunta para $\{a_t, t = p, p+1, \dots, N\}$ como:

$$f(a_p, \dots, a_N | \phi, Z_p) \propto \tau^{\frac{N-p}{2}} \exp\left\{-\frac{\tau}{2} \sum_{t=p+1}^N a_t^2\right\}$$

A equação (2.2) relaciona os ruídos $(a_{p+1}, \dots, a_N)$ com os dados Z e Z_p . Levando em consideração que a relação (2.1) tem Jacobiano unitário, podemos escrever a densidade de probabilidade conjunta dos dados Z , condicionada aos p -primeiros dados Z_p , como:

$$f(Z | \phi, \tau, Z_p) \propto \tau^{\frac{N-p}{2}} \exp\left\{-\frac{\tau}{2} \sum_{t=p+1}^N (Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p})^2\right\} \quad (2.3)$$

Considerando o conjunto de dados Z e definindo-se a matriz X :

$$Z = \begin{pmatrix} Z_{p+1} \\ Z_{p+2} \\ \vdots \\ Z_N \end{pmatrix}; X = \begin{pmatrix} Z_1 & \dots & Z_p \\ Z_2 & \dots & Z_{p+1} \\ \dots & \dots & \dots \\ Z_{N-p} & \dots & Z_N \end{pmatrix}_{N \times p} \text{ e } \phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix}_{p \times 1}$$

A função de verossimilhança, associada com os parâmetros (ϕ, τ) , condicionada aos valores iniciais Z_p pode ser escrita a partir de (2.3) como:

$$\mathcal{L}(\phi, \tau | Z, Z_p) \propto \tau^{\frac{N-p}{2}} \exp\left\{-\frac{\tau}{2} (Z - X\phi)'(Z - X\phi)\right\} \quad (2.4)$$

A função de verossimilhança (2.4) é chamada de função de verossimilhança aproximada por ter sido construída a partir da densidade conjunta dos dados condicionada aos p -primeiros valores Z_p . Tomando-se o logaritmo de (2.4) podemos calcular os estimadores de máxima verossimilhança (EMV) para ϕ e τ , seja:

$$l(\phi, \tau | Z, Z_p) = \ln(\mathcal{L}(\phi, \tau | Z, Z_p)) = \frac{N-p}{2} \ln(\tau) - \frac{\tau}{2} (Z - X\phi)'(Z - X\phi) \quad (2.5)$$

Derivando (2.5) com relação a ϕ e τ , e igualando a zero, temos:

$$\frac{\partial \ln(\mathcal{L}(\phi, \tau | Z, Z_p))}{\partial \phi} = \tau X'(Z - X\phi) = 0$$

$$\frac{\partial \ln(\mathcal{L}(\phi, \tau | Z, Z_p))}{\partial \tau} = \frac{N-p}{2} \tau^{-1} - \frac{1}{2} (Z - X\phi)'(Z - X\phi)$$

Os EMV são dados por:

$$\hat{\phi} = (X'X)^{-1}(X'Z)$$

$$\hat{\tau}^{-1} = \frac{1}{N-p} (Z - X\hat{\phi})'(Z - X\hat{\phi})$$

Podemos mostrar que a variância observada dos estimadores dos parâmetros $\hat{\phi}$ são dadas por:

$$Var(\hat{\phi}) = \tau^{-1} (X'X)^{-1}$$

Podemos verificar que:

$$(Z - X\phi)'(Z - X\phi) = (Z - \hat{Z})'(Z - \hat{Z}) + (\phi - \hat{\phi})'X'X(\phi - \hat{\phi}) \quad (2.6)$$

onde $\hat{Z}$ é dado por:

$$\hat{Z} = X\hat{\phi} \quad (2.7)$$

Substituindo (2.7) em (2.5) essa função pode ser escrita como:

$$\mathcal{L}(\phi, \tau | Z, Z_p) \propto \tau^{\frac{N-p}{2}} \exp\left\{-\frac{\tau}{2}[(\phi - \hat{\phi})' X' X (\phi - \hat{\phi}) + (Z - \hat{Z})'(Z - \hat{Z})]\right\} \quad (2.8)$$

Notando que $\mathcal{L}(\phi, \tau | Z, Z_p)$ é uma Normal-Gama, ou seja, $\mathcal{L}(\phi | \tau, Z, Z_p)$ é Normal com média $\hat{\phi}$ e matriz precisão $\tau X' X$ e $\mathcal{L}(\tau | \phi, Z, Z_p)$ é uma Gama.

2.1. Identificação da Ordem do Modelo

Identificar a ordem de um modelo auto-regressivo consiste em determinar o valor de p , ou seja determinar o número adequado de termos que devemos considerar no modelo auto-regressivo. A escolha de um valor pequeno de p , por exemplo, $p = 1$ pode levar a uma representação inadequada da série, por outro lado a escolha de um valor alto de p , por exemplo, $p = 6$ pode levar a um modelo com um grau de complexidade desnecessário. O critério de escolha do modelo adequado com o menor número de parâmetro possível é denominado por Box et alii-1994, como critério de parcimoniosidade. Dois critérios de escolha do modelo, propostos na literatura que levam em conta o critério de parcimoniosidade são: o critério de Akaike AIC, [Akaike -1974] e o Critério de informação Bayesiano BIC, [Schwarz -1978]. De acordo com esses critérios escolhe-se o modelo AR(p) que minimize os valores dos critérios de AIC e BIC, onde $\hat{\tau}$ denota o EMV de τ .

$$AIC(p) \simeq \ln \hat{\tau}^{-1} + \frac{2(p+1)}{N} = \ln \sigma^2 + \frac{2(p+1)}{N}$$

$$BIC(p) \simeq \ln \hat{\tau}^{-1} + (p+1) \frac{\ln(N)}{N} = \ln \sigma^2 + (p+1) \frac{\ln(N)}{N}$$

3. Análise Bayesiana do Modelo AR(p)

A análise Bayesiana de modelos AR(p) começa pela escolha de uma função densidade a priori para os parâmetros ϕ e τ .

3.1. Considerando Priori Não Informativa.

Quando temos "pouca" informação sobre os parâmetros podemos adotar a priori não informativa de Jeffreys; [Box e Tiao -1973]. Supondo ϕ e τ , independentes, temos a densidade conjunta a priori dada por:

$$\Pi_0(\phi, \tau) \propto 1/\tau, \tau > 0 \quad (3.1)$$

Combinando-se a função de verossimilhança (2.8) com a priori (3.1), temos a densidade a posteriori:

$$\Pi(\phi, \tau) \propto \mathcal{L}(\phi, \tau | Z, Z_p) \Pi_0(\phi, \tau) \quad (3.2)$$

A densidade conjunta a posteriori (3.2) é uma densidade Normal-Gama, dada por:

$$\Pi(\phi, \tau) \propto \tau^{\frac{N-p}{2}-1} \exp\left\{-\frac{\tau}{2}[(\phi - \hat{\phi})' X' X (\phi - \hat{\phi}) + (Z - \hat{Z})'(Z - \hat{Z})]\right\} \quad (3.3)$$

Neste caso em particular, onde estamos usando uma priori não informativa, podemos calcular as densidades para os parâmetros ϕ e τ , integrando diretamente (3.3). Assim temos:

$$\Pi(\tau) \propto \tau^{\frac{N-p}{2}-1} \exp\left\{-\frac{\tau}{2}(Z - \hat{Z})'(Z - \hat{Z})\right\} \int_{\phi} \exp\left\{-\frac{\tau}{2}[(\phi - \hat{\phi})' X' X (\phi - \hat{\phi})]\right\} d\phi \quad (3.4)$$

A integral (3.4) representa uma integral multipla para os parâmetros ϕ . Resolvendo essa integral temos:

$$\Pi(\tau) \propto \tau^{\frac{1}{2}(N-2p)-1} \exp\left\{-\frac{\tau}{2}(Z - \hat{Z})'(Z - \hat{Z})\right\} \quad (3.5)$$

Portanto τ tem uma densidade a posteriori Gama $((\frac{N-2p}{2}), (Z - \hat{Z})'(Z - \hat{Z})/2)$. temos então:

$$E(\tau | Z) = (N - 2p) [(Z - \hat{Z})'(Z - \hat{Z})]^{-1}$$

$$Var(\tau | Z) = 2(N - 2p) [(Z - \hat{Z})'(Z - \hat{Z})]^{-2}$$

Para encontrar agora a densidade a posteriori para os parâmetros ϕ , integramos (3.3) com relação a τ , temos:

$$\Pi(\phi) \propto \int_0^{\infty} \tau^{\frac{N-p}{2}-1} \exp\left\{-\frac{\tau}{2}[(\phi - \hat{\phi})' X' X (\phi - \hat{\phi}) + (Z - \hat{Z})'(Z - \hat{Z})]\right\} d\tau$$

$$\Pi(\phi) \propto [(\phi - \hat{\phi})' X' X (\phi - \hat{\phi}) + (Z - \hat{Z})'(Z - \hat{Z})]^{-\frac{N-p}{2}}$$

Após algumas simplificações algébricas, podemos escrever a densidade a posteriori dos parâmetros ϕ , como:

$$\Pi(\phi) \propto [1 + \frac{(\phi - \hat{\phi})' V (\phi - \hat{\phi})}{N-2p}]^{-\frac{N-p}{2}} = [1 + \frac{(\phi - \hat{\phi})' V (\phi - \hat{\phi})}{v}]^{-\frac{1}{2}(v+p)}$$

onde $V = (N - 2p)(X'X)[(Z - \hat{Z})'(Z - \hat{Z})]^{-1}$ e $v = N - 2p$

Portanto ϕ tem densidade a posteriori t -Student, com $v = (N - 2p)$ graus de liberdade, vetor de localização $\hat{\phi}$ e matriz de precisão V . O estimador bayesiano dos parâmetros, é o vetor esperado a posteriori de ϕ ou seja, $\phi_b = E(\phi|Z)$. sendo:

$$E(\phi|Z) = \hat{\phi} = (X'X)^{-1} X'Z$$

$$Var(\phi|Z) = \frac{v}{v-2} V^{-1} = \frac{(X'X)^{-1} [(Z - \hat{Z})'(Z - \hat{Z})]}{N-2p-2}$$

Notamos que neste caso, quando usa-se densidade a priori não-informativa, o estimador bayesiano coincide com o EMV, $\phi_b = \hat{\phi}$.

3.2. Considerando Priori Conjugada Normal-Gama

A função de verossimilhança (2.8) sugere o uso de uma densidade a priori conjugada Normal-Gama. Assim, temos:

$$\Pi_0(\phi, \tau) = \Pi_1(\phi|\tau) \Pi_2(\tau) \quad (3.6)$$

$$\Pi_1(\phi|\tau) \propto \tau^{\frac{p}{2}} \exp\left\{-\frac{\tau}{2}(\phi - \mu)' P(\phi - \mu)\right\}$$

$$\Pi_2(\tau) \propto \tau^{\alpha-1} \exp\{-\beta\tau\}$$

onde $\phi \in R^p$, $P \in R^{p \times p}$ e α, β escalares são hiperparâmetros. Combinando a função de verossimilhança (2.8) com a densidade a priori (3.6) temos a densidade a posteriori dada por:

$$\Pi(\phi, \tau) \propto \tau^{\frac{(N-p+2\alpha)}{2}-1} \exp\{-\tau D\} \tau^{\frac{p}{2}} \exp\{-\frac{\tau}{2}(\phi - \phi_b)'V(\phi - \phi_b)\} \quad (3.7)$$

onde,

$$\phi_b = (X'X + P)^{-1}(X'Z + P\mu)$$

$$V = X'X + P$$

$$D = \beta + \frac{1}{2}[(Z'Z + \mu'P\mu) - (X'Z + P\mu)'(X'X + P)^{-1}(X'Z + P\mu)]$$

Devido a conjugação, temos que a densidade conjunta a posteriori (3.7) é uma Normal-Gama. Assim, podemos determinar a densidade a posteriori de τ , integrando (3.7) com relação aos parâmetros ϕ . Procedendo desta forma, temos:

$$\begin{aligned} \Pi(\tau) &\propto \tau^{\frac{(N-p+2\alpha)}{2}-1} \exp\{-\tau D\} \int_{\phi} \tau^{\frac{p}{2}} \exp\{-\frac{\tau}{2}[(\phi - \phi_b)'V(\phi - \phi_b)]\} d\phi \\ \Pi(\tau) &\propto \tau^{\frac{(N-p+2\alpha)}{2}-1} \exp\{-\tau D\} \end{aligned} \quad (3.8)$$

Temos portanto que o valor esperado e a variância a posteriori de τ são dados por:

$$E(\tau|Z) = (N - p + 2\alpha)(2D)^{-1} \quad (3.9)$$

$$Var(\tau|Z) = E(\tau|Z)(D)^{-1} \quad (3.10)$$

Da mesma forma, integrando (3.7) com relação a τ , temos a densidade marginal a posteriori para os parâmetros ϕ , dada por:

$$\begin{aligned} \Pi(\phi) &\propto \int_0^{\infty} \tau^{\frac{(N+2\alpha)}{2}-1} \exp\{-\frac{\tau}{2}[(\phi - \phi_b)'V(\phi - \phi_b) + 2D]\} d\tau \\ \Pi(\phi) &\propto [2D + (\phi - \phi_b)'V(\phi - \phi_b)]^{-\frac{N+2\alpha}{2}} \end{aligned} \quad (3.11)$$

reagrupando os termos, podemos escrever (3.11) como:

$$\Pi(\phi) \propto \left[1 + \frac{(\phi - \phi_b)'W(\phi - \phi_b)}{N - p + 2\alpha}\right]^{-\frac{N+2\alpha}{2}} \quad (3.12)$$

onde,

$$W = (N - p + 2\alpha)V(2D)^{-1}$$

Portanto, vemos que ϕ tem densidade a posteriori t -Student, p -dimensional, com $v = (N - p + 2\alpha)$ graus de liberdade, vetor de localização ϕ_b e matriz de precisão W . Então podemos estimar os parâmetros ϕ , pelo valor esperado da densidade (3.12) e a matriz de covariância dos parâmetros pode ser obtida diretamente da matriz de precisão W . Assim temos:

$$E(\phi|Z) = (X'X + P)^{-1}(X'Z + P\mu) \quad (3.13)$$

$$Var(\phi|Z) = \frac{v}{v-2} W^{-1} = \frac{(X'X+P)^{-1}(2D)}{(N-p+2\alpha-2)} \quad (3.14)$$

Note de (3.13) que o valor esperado a posteriori, ϕ_b , é uma média ponderada do estimador de máxima verossimilhança $\hat{\phi}$ e da média da densidade a priori, μ , ou seja:

$$\phi_b = (X'X + P)^{-1}(X'Z + P\mu) = (X'X + P)^{-1}[(X'X)\hat{\phi} + P\mu] \quad (3.15)$$

3.3. Considerando Priori t -Student

Supondo agora uma densidade a priori para os parâmetros ϕ , como sendo uma densidade t -Student p -dimensional, com graus de liberdade κ , e vamos considerar uma densidade a priori Gama com parâmetros α e β para τ , assim temos:

$$\Pi_1(\phi) \propto \left(1 + \frac{(\phi-\mu)'P(\phi-\mu)}{\kappa}\right)^{-\frac{(\kappa+p)}{2}} \quad (3.16)$$

$$\Pi_2(\tau) \propto \tau^{\alpha-1} \exp\{-\beta\tau\} \quad (3.17)$$

com grau de liberdade κ , vetor de localização μ e matriz de precisão P .

A análise a posteriori é obtida combinando-se a função de verossimilhança (2.8) com (3.16) e (3.17), desta forma temos:

$$\Pi(\phi, \tau) \propto \tau^{\frac{N-p+2\alpha}{2}-1} \exp\{-\tau[\beta + B(\phi)/2]\} \Pi_1(\phi) \quad (3.18)$$

onde:

$$B(\phi) = (\phi - \hat{\phi})'(X'X)(\phi - \hat{\phi}) + (Z - \hat{Z})'(Z - \hat{Z}).$$

Notamos aqui que a densidade conjunta a posteriori não é uma densidade conhecida, portanto só podemos avaliar as densidades a posteriori marginais por meio de métodos de aproximação, tais como o método de Laplace [Tierney, Kass e Kadane -1986] ou usando métodos de simulação de Monte Carlo em Cadeia de Markov, tais como, Gibbs Sampling [Casela e George -1992] e Metropolis-Hastings [Chib e Greenberg -1995]. Neste trabalho vamos optar pelos métodos de simulação por serem precisos e de fácil implementação computacional. Com a finalidade de fazer uma ilustração do uso de algoritmos de MCMC vamos fazer uma estimação dos parâmetros ϕ e τ para os casos em que temos prioris conjugada.

3.4 Análise Preditiva

O objetivo dos modelos de séries temporais é fazer previsão de valores futuros da série, $Z_{N+1}, Z_{N+2}, \dots$.

Previsão um passo a frente

Considerando uma observação futura Z_{N+1} , dos dados $Z = (Z_1, Z_2, \dots, Z_N)'$.

Seja,

$$\Pi(Z_{N+1}|Z, \phi, \tau) \propto \tau^{\frac{1}{2}} \exp\left\{-\frac{\tau}{2}(Z_{N+1} - \sum_{i=1}^p \phi_i Z_{N-(i-1)})^2\right\} \quad (3.19)$$

sendo:

$$(Z_{N+1} - \sum_{i=1}^p \phi_i Z_{N-(i-1)})^2 = Z_{N+1}^2 - 2 Z_{N+1} \sum_{i=1}^p \phi_i Z_{N-(i-1)} + (\sum_{i=1}^p \phi_i Z_{N-(i-1)})^2$$

fazendo:

$$Z_{N+1} \sum_{i=1}^p \phi_i Z_{N-(i-1)} = [\phi_1, \dots, \phi_p] Z_{N+1} \begin{bmatrix} Z_N \\ \vdots \\ Z_{N-(p-1)} \end{bmatrix} = \phi' F$$

e

$$(\sum_{i=1}^p \phi_i Z_{N-(i-1)})^2 = \phi' E \phi$$

onde E é uma matriz $p \times p$ (onde o ij^{th} elemento é $Z_{N-(i-1)} Z_{N-(j-1)}$) e F é uma matriz $p \times 1$ (onde o j^{th} elemento é $Z_{N+1} Z_{N-(j-1)}$). Assim temos a densidade preditiva condicional,

$$\Pi(Z_{N+1}|Z, \phi, \tau) \propto \tau^{\frac{1}{2}} \exp\left\{-\frac{\tau}{2}(Z_{N+1}^2 + \phi' E \phi - 2 \phi' F)\right\} \quad (3.20)$$

A preditiva posteriori condicional é dada por:

$$\Pi(Z_{N+1}, \phi, \tau | Z) \propto \Pi(Z_{N+1} | Z, \phi, \tau) \Pi(\phi, \tau | Z) \quad (3.21)$$

onde $\Pi(\phi, \tau | Z)$ é a posteriori condicional dos parâmetros dado as observações.

Modelo com Prioris não Informativa

Considerando uma priori não informativa, a posteriori é dada por (3.3), fazendo o produto temos:

$$\begin{aligned} \Pi(Z_{N+1}, \phi, \tau | Z) \propto & \tau^{\frac{N-p-1}{2}} \exp\left\{-\frac{\tau}{2}[(Z - \hat{Z})'(Z - \hat{Z}) + Z_{N+1}^2]\right\} \\ & \exp\left\{-\frac{\tau}{2}[(\phi - \phi_b)'V(\phi - \phi_b) + \phi'E\phi - 2\phi'F]\right\} \end{aligned} \quad (3.22)$$

Integrando (3.22) em ϕ e τ temos:

$$\Pi(Z_{N+1} | Z) \propto \left[1 + \frac{\mathbb{P}(Z_{N+1} - \eta)^2}{N-2p}\right]^{-\frac{N-2p+1}{2}} \quad (3.23)$$

com

$$\eta = (1 - F^{*'}(V + E)^{-1}F^*)^{-1}F^{*'}(V + E)^{-1}C$$

$$V = X'X$$

$$C = X'Z$$

$$\mathbb{P} = (N - 2p)(1 - F^{*'}(V + E)^{-1}F^*)^{-1}$$

$$[2D + C'V^{-1}C - C'(V + E)^{-1}(1 - F^{*'}(V + E)^{-1}F^*)^{-1}C]^{-1}$$

onde F^* é um vetor $p \times 1$ com o j^{th} componente dado por $Z_{N-(j-1)}$ e $j = 1, \dots, p$.

Portanto (3.23) é a função densidade preditiva um passo a frente, e é uma t -Student com $v = (N - 2p)$ graus de liberdade, localização η e precisão $\mathbb{P}$. Então podemos estimar o valor esperado da densidade (3.23) e a variância pode ser obtida diretamente da matriz de precisão $\mathbb{P}$. Assim temos:

$$E(Z_{N+1} | Z) = (1 - F^{*'}(V + E)^{-1}F^*)^{-1}F^{*'}(V + E)^{-1}C$$

$$Var(Z_{N+1} | Z) = \frac{v}{v-2} \mathbb{P}^{-1}$$

Modelo com Priori Informativa Normal-Gama

Considerando uma priori conjugada Normal-Gama, a posteriori é dada por (3.7), fazendo o produto temos:

$$\begin{aligned} \Pi(Z_{N+1}, \phi, \tau | Z) &\propto \tau^{\frac{(N-p+2\alpha)-1}{2}} \exp\left\{-\frac{\tau}{2}[(2D + Z_{N+1}^2] \right. \\ &\quad \left. \tau^{\frac{p}{2}} \exp\left\{-\frac{\tau}{2}[(\phi - \phi_b)'V(\phi - \phi_b) + \phi'E\phi - 2\phi'F]\right\}\right\} \end{aligned} \quad (3.24)$$

Integrando (3.24) em ϕ e τ temos:

$$\Pi(Z_{N+1} | Z) \propto \left[1 + \frac{\mathbb{P}(Z_{N+1} - \eta)^2}{N - p + 2\alpha}\right]^{-\frac{N - p + 2\alpha + 1}{2}} \quad (3.25)$$

com

$$\eta = (1 - F^{*'}(V + E)^{-1}F^*)^{-1}F^{*'}(V + E)^{-1}C$$

$$V = X'X + P$$

$$C = (X'Z + P\mu)$$

$$P = (N - p + 2\alpha)(1 - F^{*'}(V + E)^{-1}F^*)^{-1}$$

$$[2D + C'V^{-1}C - C'(V + E)^{-1}(1 - F^{*'}(V + E)^{-1}F^*)^{-1}C]^{-1}$$

Portanto (3.25) é a função densidade preditiva um passo a frente, e é uma t -Student, p dimensional, com $v = (N - p + 2\alpha)$ graus de liberdade, localização η e precisão P .

Então podemos estimar o valor esperado da densidade (3.25) e a variância pode ser obtida diretamente da matriz de precisão P . Assim temos:

$$E(Z_{N+1} | Z) = (1 - F^{*'}(V + E)^{-1}F^*)^{-1}F^{*'}(V + E)^{-1}C$$

$$Var(Z_{N+1} | Z) = \frac{v}{v-2} P^{-1}$$

Modelo com Priori Informativa *t*-Student

Considerando uma priori *t*-Student, a posteriori é dada por (3.18), fazendo o produto temos:

$$\begin{aligned} \Pi(Z_{N+1}, \phi, \tau | Z) &\propto \tau^{\frac{(N-p+2\alpha)-1}{2}} \exp\left\{-\frac{\tau}{2}[\beta + Z_{N+1}^2]\right\} \\ &\exp\left\{-\frac{\tau}{2}[B(\phi) + \phi' E \phi - 2\phi' F]\right\} \Pi_1(\phi) \end{aligned} \quad (3.26)$$

Neste caso não é possível integrar a equação (3.26), a previsão de valores futuros só podem ser calculados por simulação ou aproximação numérica.

3.5 Previsão *k* passos a frente

Denotando por Z_f os *k* valores da série a serem previstos e por Z_p as últimas *p* observações, ou seja:

$$Z_f = (Z_{(N+1)}, Z_{(N+2)}, \dots, Z_{(N+k)})'$$

$$Z_p = (Z_{N-(p-1)}, Z_{N-(p-2)}, \dots, Z_N)'$$

Seja $\Pi(\phi, \tau | Z)$ a densidade a posteriori para os parâmetros ϕ e τ . A densidade preditiva pode ser calculada por:

$$\Pi(Z_f | Z) = \int_{\phi} \int_{\tau} \Pi(Z_f | \phi, \tau, Z_p) \Pi(\phi, \tau | Z_p) d\phi d\tau$$

onde $\int_{\phi}$ denota as *p*-integrais $\int_1, \dots, \int_p$.

A previsão dos valores futuros Z_f pode ser obtida resolvendo-se a integral múltipla,

$$E(Z_f | Z) = \int Z_f \Pi(Z_f | Z) dZ_f$$

$$E(Z_f | Z) = \int Z_f \left[\int_{\phi} \int_{\tau} \Pi(Z_f | \phi, \tau, Z_p) \Pi(\phi, \tau | Z_p) d\phi d\tau \right] dZ_f$$

Alterando a ordem das integrações temos,

$$E(Z_f | Z) = \int_{\phi} \int_{\tau} E(Z_f | \phi, \tau, Z_p) \Pi(\phi, \tau | Z_p) d\phi d\tau \quad (3.27)$$

Usando o modelo $AR(p)$, temos:

$$Z_{(N+k)} = \phi_1 Z_{(N+k-1)} + \dots + \phi_p Z_{(N-(p-k))} + a_{(N+k)}$$

Colocando na forma matricial

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi_1 & 1 & 0 & \dots & 0 \\ -\phi_2 & -\phi_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\phi_{k-1} & -\phi_{k-2} & -\phi_{k-3} & \dots & 1 \end{pmatrix}_{k \times k}$$

$$B = \begin{pmatrix} -\phi_p & -\phi_{p-1} & -\phi_{p-2} & \dots & -\phi_1 \\ 0 & -\phi_p & -\phi_{p-1} & \dots & -\phi_2 \\ 0 & 0 & -\phi_p & \dots & -\phi_3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\phi_k \end{pmatrix}_{k \times p}$$

Podemos escrever:

$$AZ_f + BZ_p = a$$

$$Z_f = -A^{-1}BZ_p + A^{-1}a$$

onde a é um vetor de variáveis aleatória i.i.d. $N(0, \tau^{-1})$. Assim temos que Z_f também é um vetor de variáveis aleatória normalmente distribuída $N[E(Z_f|\phi, \tau, Z), Var(Z_f|\phi, \tau, Z)]$ onde:

$$E(Z_f|\phi, \tau, Z) = -A^{-1}BZ_p \quad (3.28)$$

$$Var(Z_f|\phi, \tau, Z) = \tau^{-1}(A^{-1})(A^{-1})' \quad (3.29)$$

Substituindo (3.28) em (3.27) podemos calcular a previsão dos valores futuros Z_f , no entanto, para isso é necessário resolver a integral

$$E(Z_f|Z) = \int_{\phi} \int_{\tau} (-A^{-1}BZ_p) \Pi(\phi, \tau|Z) d\phi d\tau \quad (3.30)$$

A variância é calculada como:

$$\text{Var}(Z_f|Z) = E_{\phi, \tau}[\text{Var}(Z_f|\phi, \tau, Z)] + \text{Var}_{\phi, \tau}[E(Z_f|\phi, \tau, Z)] \quad (3.31)$$

onde:

$$\begin{aligned} E_{\phi, \tau}[\text{Var}(Z_f|\phi, \tau, Z)] &= \int_{\phi} \int_{\tau} \text{Var}(Z_f|\phi, \tau, Z) \Pi(\phi, \tau|Z) d\phi d\tau \\ &= \int_{\phi} \int_{\tau} \tau^{-1} (\mathbf{A}^{-1})(\mathbf{A}^{-1})' \Pi(\phi, \tau|Z) d\phi d\tau \end{aligned} \quad (3.32)$$

Devido a complexidade das expressões (3.30) e (3.31) quando $p \geq 2$, geralmente essa integral só pode ser calculada de forma aproximada. No entanto, mostraremos na seção 4 que $E(Z_f|Z)$ e $\text{Var}(Z_f|Z)$ podem ser calculadas facilmente usando métodos de simulação de Monte-Carlo.

3.6. Considerações sobre Seleção de Modelo

Na análise Bayesiana existem vários critérios para seleção de modelo. Aqui vamos considerar o critério baseado na densidade preditiva. A densidade preditiva para Z_{N+k} dado $Z_k = (Z_1, Z_2, \dots, Z_N, \dots, Z_{N+k})'$ e os parâmetros ϕ e τ é dada por:

$$\Pi(Z_{N+k}|Z_k) \propto \int_{\phi} \int_{\tau} \tau^{\frac{1}{2}} \exp\left\{-\frac{\tau}{2}(Z_{N+k} - \sum_{i=1}^p \phi_i Z_{N-(i-1)})^2\right\} \Pi(\phi, \tau|Z) d\phi d\tau \quad (3.33)$$

A equação (3.33) pode também ser calculada usando simulação de Monte Carlo. Neste caso a densidade preditiva é estimada por $c_k = \hat{\Pi}(Z_{N+k}|Z_k)$. Então fazemos o gráfico de c_k versus k ($k = 1, 2, \dots, n$) para diferentes modelos, o maior valor de c_k (em média) indica o melhor modelo. Escolhemos também o modelo tal que $c(l) = \prod_{i=1}^n c_i(l)$ é máximo (l é o índice do modelo).

4. Algoritmos de MCMC

4.1 Algoritmo Gibbs Sampling:

A técnica do Amostrador de Gibbs é usada para gerar variáveis aleatórias de uma distribuição sem utilizar sua densidade. O Amostrador de Gibbs é essencialmente um esquema iterativo de amostragem de uma cadeia de Markov cujo núcleo de transição é formado pelas distribuições condicionais.

Modelo com Prioris não Informativa

Neste caso, temos uma densidade conjunta a posteriori, Normal-Gama, (3.3), dada por:

$$\Pi(\phi, \tau) \propto \tau^{\frac{N-p}{2}-1} \exp\left\{-\frac{\tau}{2}[(\phi - \hat{\phi})'(X'X)(\phi - \hat{\phi}) + (Z - \hat{Z})'(Z - \hat{Z})]\right\}$$

As densidades a posteriori condicionais, são:

$$\Pi(\phi|\tau) \sim N(\hat{\phi}, (\tau V)^{-1}) \quad (4.1)$$

$$\Pi(\tau|\phi) \sim \Gamma\left(\frac{N-p}{2}, B\right) \quad (4.2)$$

onde, $B = \frac{1}{2}[(\phi - \hat{\phi})'(X'X)(\phi - \hat{\phi}) + (Z - \hat{Z})'(Z - \hat{Z})]$, $V = (X'X)^{-1}$

Modelo com Priori Informativa Normal-Gama

A posteriori conjugada para este caso dada por (3.7) é:

$$\Pi(\phi, \tau) \propto \tau^{\frac{(N-p+2\alpha)}{2}-1} \exp\{-\tau D\} \tau^{\frac{p}{2}} \exp\left\{-\frac{\tau}{2}(\phi - \phi_b)'V(\phi - \phi_b)\right\}$$

As densidades a posteriori condicionais, são:

$$\Pi(\phi|\tau) \sim N(\phi_b, (\tau V)^{-1}) \quad (4.3)$$

$$\Pi(\tau|\phi) \sim \Gamma\left(\frac{N+2\alpha}{2}, D + \frac{1}{2}(\phi - \phi_b)'V(\phi - \phi_b)\right) \quad (4.4)$$

onde,

$$V = (X'X + P)$$

$$D = \beta + \frac{1}{2}[(Z'Z + \mu'P\mu) - (X'Z + P\mu)'(X'X + P)^{-1}(X'Z + P\mu)]$$

passo 1: Inicialize o contador de iterações $j = 1$ e arbitre valores iniciais:

$$\phi^{(0)} = (\phi_1^{(0)}, \dots, \phi_p^{(0)}), \tau^{(0)}$$

passo 2: Obtenha os novos valores gerando-se a partir das densidades condicionais como se segue:

$$\phi^{(j+1)} \sim \Pi(\phi|\tau^{(j)})$$

$$\tau^{(j+1)} \sim \Pi(\tau|\phi^{(j+1)})$$

passo 3: Atualize o contador de iterações para $j \leftarrow j + 1$ e volte ao passo 2, repetir este procedimento até a convergência.

Após um número grande de iterações as sequências geradas (ϕ^j e τ^j) convergirão para uma amostra aleatória das variáveis $\phi \sim \Pi(\phi)$ e $\tau \sim \Pi(\tau)$. Para verificar se a convergência já foi atingida existem vários métodos [Cowles e Carlin -1996]. Neste trabalho vamos usar o critério de Gelman e Rubin (1992).

4.2 Algoritmo Metropolis-Hastings & Gibbs Sampling:

Quando as distribuições condicionais não são facilmente identificadas nós usamos o algoritmo de Metropolis-Hastings.

Modelo com Priori Informativa t -Student

Neste caso a posteriori conjugada é dada em (3.18).

$$\Pi(\phi, \tau) \propto \tau^{\frac{N-p+2\alpha}{2}-1} \exp\left\{-\tau[\beta + B(\phi)/2]\right\} \Pi_1(\phi)$$

As posteriors condicionais são:

$$\Pi(\phi|\tau) \propto \tau^{\frac{N-p+2\alpha}{2}-1} \exp\left\{-\tau\left[\beta + \frac{1}{2}(\phi - \hat{\phi})'(X'X)(\phi - \hat{\phi})\right]\right\} \Pi_1(\phi) \quad (4.5)$$

$$\Pi(\tau|\phi) \sim \Gamma\left(\frac{N-p-2\alpha}{2}, \beta + \frac{1}{2}B(\phi)\right) \quad (4.6)$$

Aqui é conveniente escrever:

$$\Pi(\phi|\tau) \propto \Psi(\phi) \Pi_1(\phi)$$

onde:
$$\Psi(\phi) = \tau^{\frac{N-p+2\alpha}{2}-1} \exp\{-\tau[\beta + B(\phi)/2]\}$$

e usamos $\Pi_1(\phi)$ como núcleo para gerar os valores do parâmetro ϕ como se segue:

1) Inicialize o contador de iterações com $j = 1$ e arbitre valores iniciais com

$$\phi^{(0)} = (\phi_1^{(0)}, \dots, \phi_p^{(0)}), \tau^{(0)}$$

2) Obtenha um novo valor a partir da função transição $\Pi_1(\phi, \phi^*)$

3) Calcule a probabilidade de aceitação do novo valor na iteração j por:

$$\alpha(\phi^j, \phi^*) = \begin{cases} \min\{1, \frac{\psi(\phi^*)}{\psi(\phi^j)}\} & \text{se } \psi(\phi^j) > 0 \\ 1 & \text{caso contrário.} \end{cases}$$

4) Gerar u de Uniforme[0, 1] e faça:

$$\phi^{j+1} = \begin{cases} \phi^* & \text{se } u \leq \alpha(\phi, \phi^*) \\ \phi^j & \text{caso contrário} \end{cases}$$

O valor gerado ϕ^* é aceito com probabilidade α ou rejeitado com probabilidade $(1 - \alpha)$.

5) Gerar $\tau^{(j)} \sim \text{Gama}(\tau|\phi^{j+1})$

6) Atualize o contador de iteração j para $j + 1$ e volte para o passo 2 até a convergência.

4.3 Previsão usando MCMC

Retomando a equação (3.30) dada por:

$$E(Z_f|Z) = \int_{\phi} \int_{\tau} (-A^{-1}BZ_p) \Pi(\phi, \tau|Z) d\phi d\tau$$

O valor esperado pode ser calculado usando algoritmos de simulação de Monte-Carlo. Para isso seja a amostra $\phi^{(j)} = (\phi_1^{(j)}, \phi_2^{(j)})$; $j = 1, \dots, m$ gerados por simulação (Amostrador de Gibbs ou Metropolis-Hastings). Podemos calcular $E(Z_f|Z)$ usando essa amostra, por:

$$E(Z_f|Z) = \frac{1}{m} \sum_{j=1}^m -A^{-1}(\phi^{(j)})B(\phi^{(j)})Z_p$$

$$Var(Z_f|Z) = \int \phi Z_f \Pi(Z_f|Z) dZ_f$$

$$Var(Z_f|Z) = \frac{1}{N} \sum_{i=1}^N \phi(Z_i)$$

Retomando as equações (3.31) e (3.32) temos:

$$Var(Z_f|Z) = E_{\phi, \tau}[Var(Z_f|\phi, \tau, Z)] + Var_{\phi, \tau}[E(Z_f|\phi, \tau, Z)]$$

Usando simulação de Monte Carlo podemos calcular cada termo desta equação como:

$$E_{\phi, \tau}[Var(Z_f|\phi, \tau, Z)] = \frac{1}{m} \sum_{j=1}^m (\tau^{(j)})^{-1} [A^{-1}(\phi^{(j)})][A^{-1}(\phi^{(j)})]' \quad (4.7)$$

sendo:

$$E(Z_f|\phi^{(j)}, \tau^{(j)}, Z) = -A^{-1}(\phi^{(j)})B(\phi^{(j)})Z_p, \quad j = 1, \dots, m$$

Podemos calcular a $Var_{\phi, \tau}$ como:

$$Var_{\phi, \tau}[E(Z_f|\phi, \tau, Z)] = \frac{1}{m} \sum_{j=1}^m (A^{-1}(\phi^{(j)})B(\phi^{(j)})Z_p)^2 - \left(\frac{1}{m} \sum_{j=1}^m A^{-1}(\phi^{(j)})B(\phi^{(j)})Z_p \right)^2 \quad (4.8)$$

Usando as equações (4.7) e (4.8) podemos calcular $Var(Z_f|Z)$.

O critério de seleção de modelo utilizando simulação de Monte Carlo é feito através do cálculo da preditiva $c_k = f(Z_{N+k})$ o qual pode ser estimado usando os resultado da simulação MCMC, como:

$$\hat{c}_k = \frac{1}{m} \sum_{j=1}^m (\tau^{(j)})^{-1} \exp\left\{ -\frac{\tau^{(j)}}{2} \left[Z_{N+k} - \sum_{i=1}^p \phi_i^{(j)} Z_{N-(i-1)} \right]^2 \right\}$$

Fazendo o gráfico de c_k versus k ($k = 1, 2, \dots$) e calculando $\hat{c}(l) = \prod_{k=1}^n c_k(l)$ para cada modelo l , escolhemos o modelo que apresenta o maior $\hat{c}(l)$.

5. Estudo de Casos

A comparação entre as abordagens clássica e bayesiana é feita neste trabalho considerando-se duas séries de dados gerados e uma série real correspondente a componente estocástica dos dados correspondente as vazões médias mensais que chegam ao reservatório de Furnas no sudeste do Brasil. Uma descrição dos três componentes de dados é apresentado na Tabela 1.

Tabela 1- Dados usados para ajuste dos modclos

	Descrição	ϕ_1	ϕ_2	τ	N
série 1	AR(2)	0.8	- 0.5	0.5	200
série 2	AR(2)	- 0.5	0.3	1.0	200
série 3	Dados reais	-	-	-	540

No caso da série 3 a componente estocástica é calculada por:

$$Z_{t(r,m)} = \frac{X_{t(r,m)} - \mu_m}{\sigma_m}$$

onde m representa o número de meses $m = 1, 2, \dots, 12$, r o número de anos $r = 1, 2, \dots, 45$ e $t(r, m) = (r - 1)12 + m$, μ_m e σ_m são as médias e desvio-padrão estimado dos dados para cada mês do ano. Para as três séries o modelo ajustado é um AR(2), ou seja:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t$$

Um gráfico com as séries 1, 2 e com os últimos 200 pontos da série 3 é mostrado nas figuras 1a, 1b e 1c.

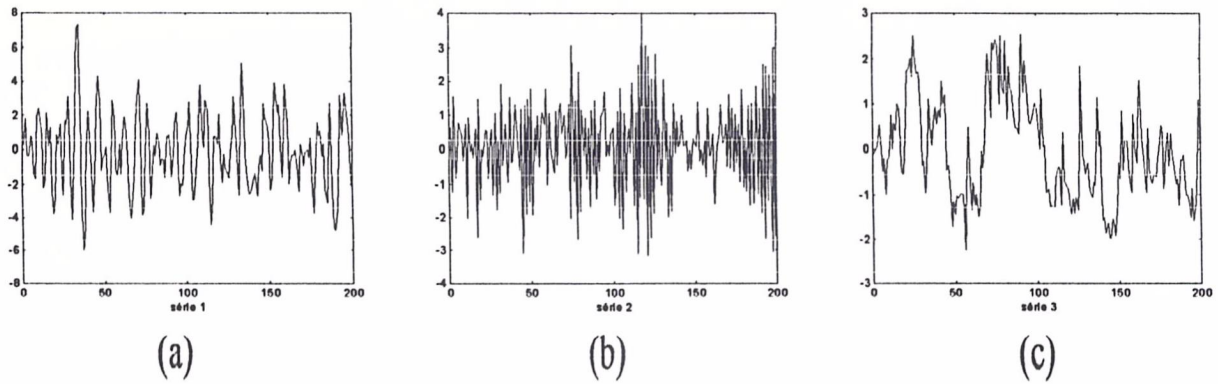


Figura 1 (a) série 1, (b) série 2, (c) últimos 200 pontos da série 3.

A Figura 2 a seguir mostra a função de autocorrelação (FAC) e a função de autocorrelação parcial (FACP), para os dados da série 3 (Furnas).

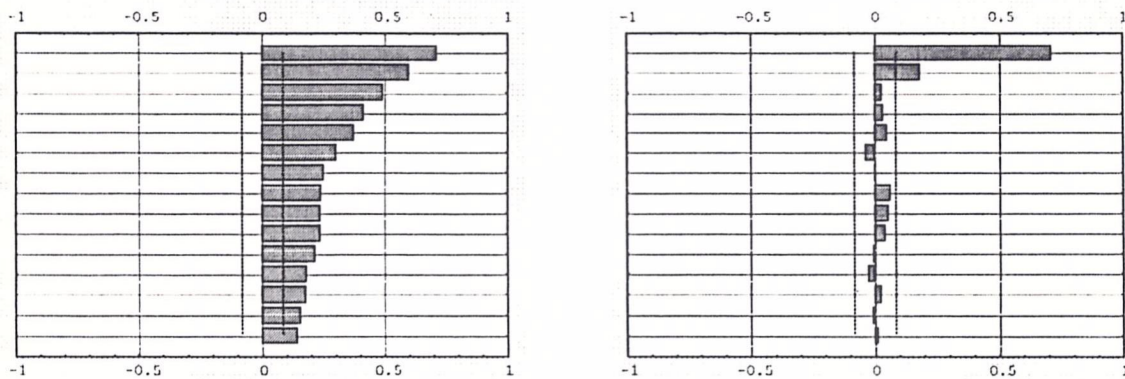


Figura 2: FAC e FACP, para os dados da série 3.

A escolha da ordem do modelo da série 3 é feita através da análise da FAC e FACP e usando-se os critérios AIC e BIC. O valor desses critérios para os modelos ajustados é apresentado na Tabela 2.

Tabela 2 identificação do modelo

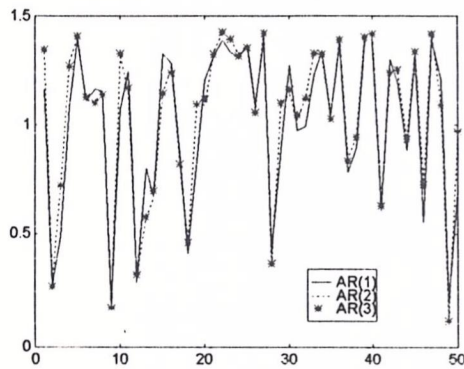
	AR(1)	AR(2)	AR(3)
AIC	- 0.783516	- 0.811755	- 0.807023
BIC	- 0.767621	- 0.787913	- 0.775234

Usando o procedimento bayesiano na verificação da adequabilidade do modelo através da densidade preditiva $\Pi(Z_{N+1:k}|Z_k)$, o valor médio de c_k e o produto $c(l)$ para as séries estudadas neste trabalho é apresentado na tabela 3.

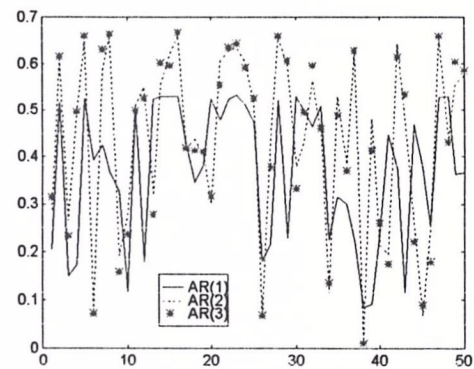
Tabela 3 seleção de modelos

Furnas		AR(1)	AR(2)	AR(3)
	$\bar{c}_k$	0.9887	1.0107	1.0114
	$c(l)$	4.7841e-005	5.6579e-005	5.4942e-005
série 1		AR(1)	AR(2)	AR(3)
	$\bar{c}_k$	0.3726	0.4379	0.4345
	$c(l)$	1.8114e-024	1.7825e-022	6.8961e-023
série 2		AR(1)	AR(2)	AR(3)
	$\bar{c}_k$	0.7450	0.7824	0.7819
	$c(l)$	2.0413e-008	1.1434e-007	1.0365e-007

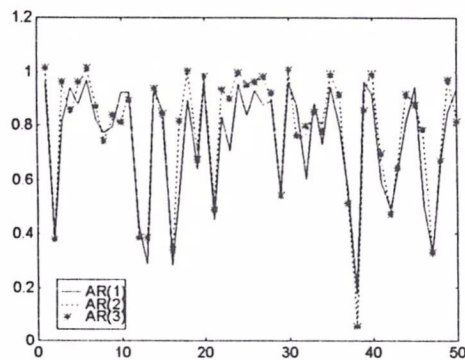
Nas tabelas 2 e 3 observamos que o modelo de ordem 2 é o mais adequado. As figuras 3-a, 3-b e 3-c mostra o gráfico de c_k versus k ($k = 1, 2, \dots, n$) para diferentes modelos, das séries Furnas, série 1 e série 2 respectivamente.



(a)



(b)



(c)

Figura 3: c_k versus k para diferentes modelos ajustados aos dados, (a) Furnas, (b) Série 1, (c) Série 2

Na Figura 4 apresentamos os histogramas estimado por simulação para os parâmetros ϕ_1 , ϕ_2 e τ , quando consideramos priori Normal-Gama, para a série 3 estudada neste trabalho.

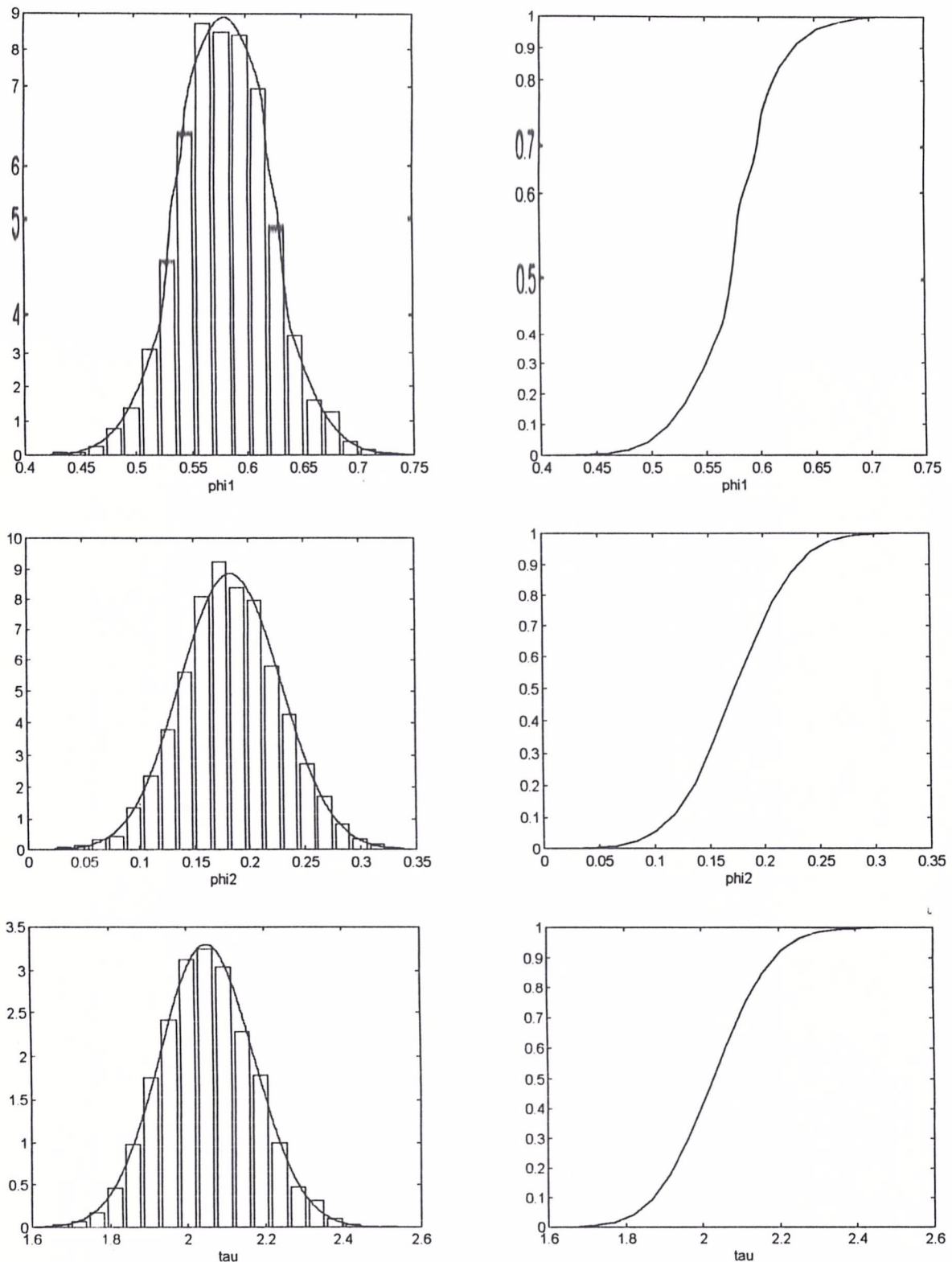


Figura 4: histograma estimado a posteriori para a série 3 (Furnas) considerando priori Normal-Gama para os parâmetros ϕ_1 , ϕ_2 e τ respectivamente.

Na Figura 4-a apresentamos os histogramas estimado por simulação para os parâmetros ϕ_1 , ϕ_2 e τ , quando consideramos priori t -Student, para a série 3 (Furnas) estudada neste trabalho.

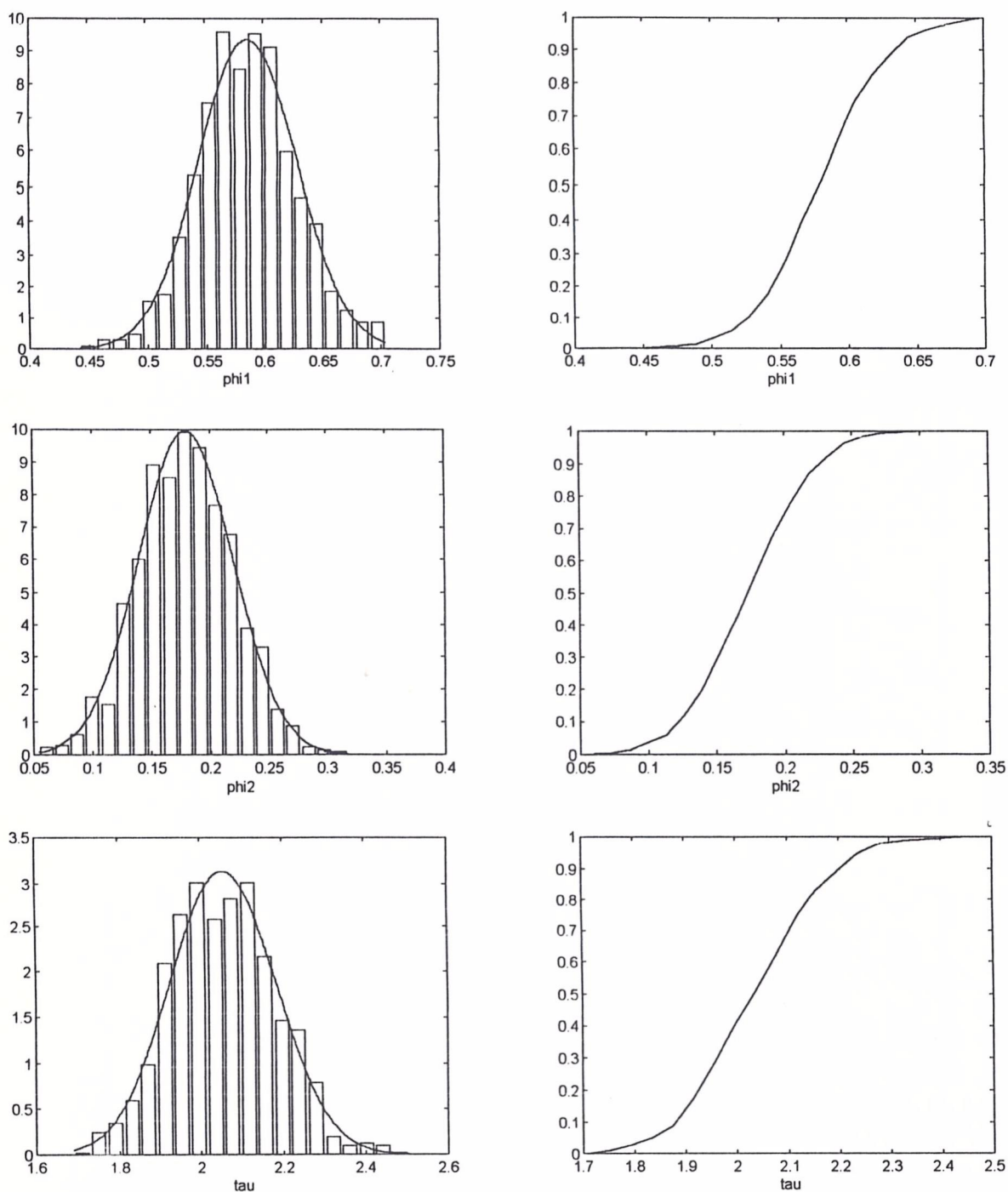


Figura 4-a: histograma estimado a posteriori para a série 3 (Furnas) considerando priori t -Student para os parâmetros ϕ_1 , ϕ_2 e τ respectivamente.

Observamos na Figura 4, que o comportamento da posteriori é aquele esperado na seção 3.2, e que com o uso da priori t -Student, apesar de não conhecermos a distribuição a posteriori, os histogramas se apresentam de uma forma simétrica.

No caso particular de priori t -Student, onde utilizamos o algoritmo metropolis-hastings foram geradas duas amostras de tamanho $N = 6.000$, desprezamos 3.000 e selecionamos para estimação uma amostra de tamanho 200 tomando-se um valor a cada 15 gerado na última sequência de 3.000 valores. Esse procedimento mostrou-se necessário para evitar a dependência entre os valores gerados devido a rejeição de novos valores. A Figura 5 a seguir mostra a correlação entre os valores gerados quando não se adota a seleção espaçada, para a série Furnas.

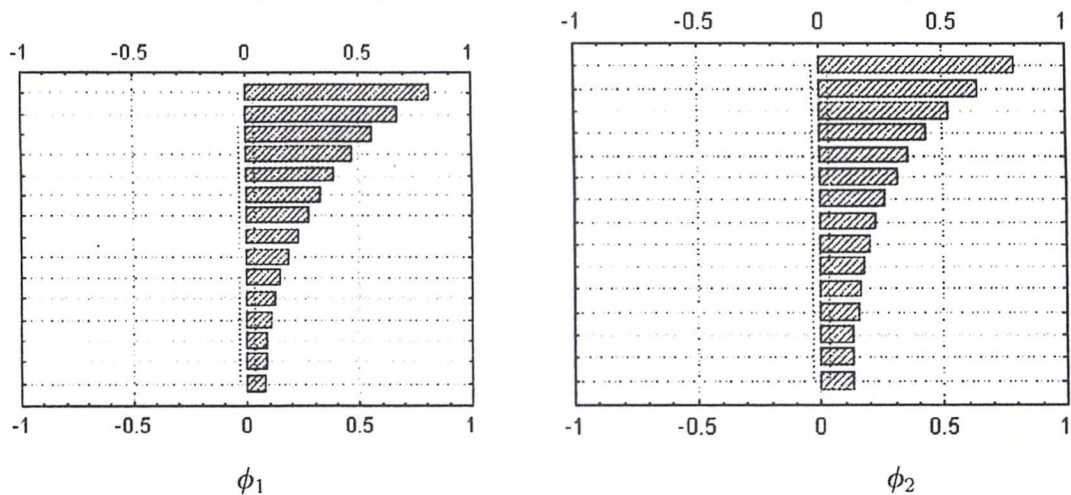


Figura 5 função de autocorrelação entre os valores gerados para ϕ_1 e ϕ_2 ; $N=15.000$.

A Figura 6 mostra a correlação entre os valores gerados quando se adota a seleção espaçada, procedendo assim diminui-se a correlação entre os dados gerados, e a variância do estimador.

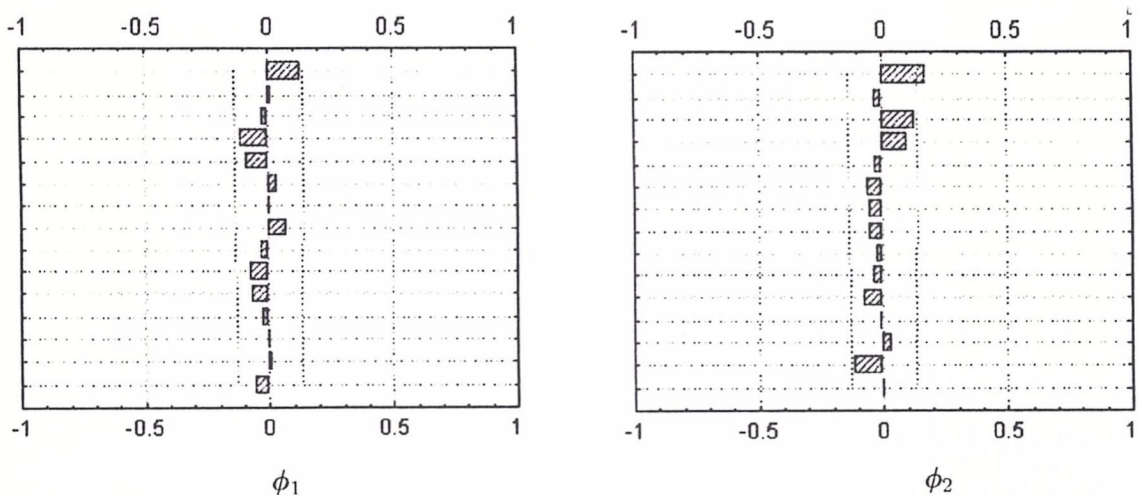


Figura 6- função de autocorrelação entre os valores gerados para ϕ_1 e ϕ_2 ; $N=1.000$.

A Figura 7 a seguir mostra a convergência, para os dados da série 3 (Furnas), partindo-se de uma condição inicial arbitrária.

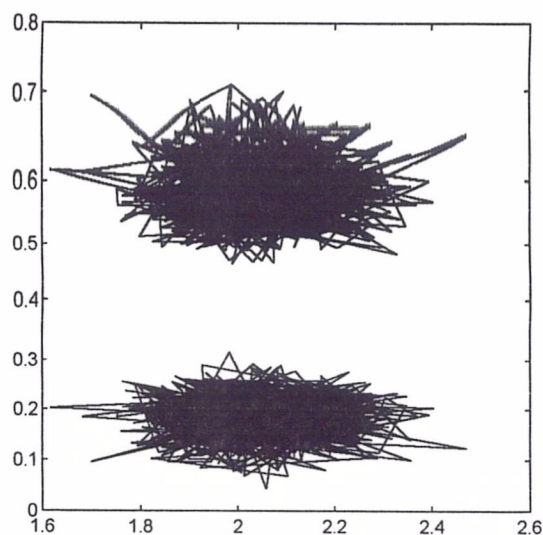


Figura 7 gráfico da convergência para ϕ_1, ϕ_2 e τ com os dados da série 3 (Furnas), 30.000 iterações.

O comportamento apresentado para a série de Furnas é o mesmo para as séries 1 e 2.

As Tabelas 3, 4 e 5 apresentam os estimadores dos parâmetros ϕ_1, ϕ_2 e τ para cada uma das três séries analisadas neste trabalho. EMV corresponde ao estimador de máxima verossimilhança; PNI(·) corresponde ao estimador quando adotamos prioris não-informativa, EX significa que esse é o estimador exato obtido pelas expressões do valor esperado da densidade a posteriori e GS significa o estimador obtido por simulação usando-se o algoritmo Gibbs Sampling; PNG(·) são os estimadores obtidos usando prioris conjugada Normal-Gama; PTS são os estimadores obtidos usando-se prioris t -Student com o algoritmo Metropolis-Hastings. Nestas tabelas apresentamos também a variância do estimador e o critério de convergência proposto por Gelman e Rubin (1992), quando os estimadores são obtidos por simulação. Este critério assegura a convergência quando $R \approx 1,0$.

Tabela 4 Valores estimados dos parâmetros ϕ_1, ϕ_2 e τ para a série 1

	ϕ_1	SD	R	ϕ_2	SD	R	τ	SD	R
EMV	0.8156	0.0580	—	— 0.5773	0.0581	—	0.4119	—	—
PNI(EX)	0.8156	0.0586	—	— 0.5773	0.0587	—	0.4077	0.0412	—
PNI(GS)	0.8158	0.0581	1.0006	— 0.5773	0.0578	1.0001	0.4085	0.0411	1.0001
PNG(EX)	0.8141	0.0578	—	— 0.5759	0.0579	—	0.4178	0.0400	—
PNG(GS)	0.8130	0.0588	1.0002	— 0.5756	0.0580	1.0003	0.4186	0.0398	1.0001
PTS(MH)	0.8181	0.0550	1.0003	— 0.5806	0.0557	1.0078	0.4154	0.0385	1.0035

Tabela 5 Valores estimados dos parâmetros ϕ_1, ϕ_2 e τ para a série 2

	ϕ_1	SD	R	ϕ_2	SD	R	τ	SD	R
EMV	-0.5158	0.0679	—	0.2946	0.0684	—	1.2550	—	—
PNI(EX)	-0.5158	0.0686	—	0.2946	0.0691	—	1.2423	0.1255	—
PNI(GS)	-0.5163	0.0690	1.0000	0.2937	0.0685	1.0001	1.2443	0.1257	1.0001
PNG(EX)	-0.5141	0.0725	—	0.2951	0.0731	—	1.1003	0.1054	—
PNG(GS)	-0.5149	0.0737	1.0005	0.2947	0.0741	1.0001	1.1025	0.1058	1.0001
PTS(MH)	-0.5127	0.0620	1.0010	0.2965	0.0633	1.0053	1.0961	0.0987	1.0024

Tabela 6 Valores estimados dos parâmetros ϕ_1, ϕ_2 e τ para a série 3

	ϕ_1	SD	R	ϕ_2	SD	R	τ	SD	R
EMV	0.5834	0.0420	—	0.1815	0.0421	—	2.3244	—	—
PNI(EX)	0.5834	0.0421	—	0.1815	0.0421	—	2.3244	0.1417	—
PNI(GS)	0.5835	0.0421	1.0002	0.1809	0.0419	1.0000	2.3161	0.1409	1.0004
PNG(EX)	0.5815	0.0447	—	0.1825	0.0447	—	2.0528	0.1229	—
PNG(GS)	0.5821	0.0447	1.0000	0.1820	0.0446	1.0000	2.0531	0.1229	1.0002
PTS(MH)	0.5811	0.0406	1.0032	0.1824	0.0413	1.0003	2.0505	0.1223	1.0051

Observando os resultados obtidos concluímos que o uso dos algoritmos Gibbs Sampling e Metropolis-Hastings tem um bom desempenho na estimativa dos parâmetros.

No caso particular de priori não informativa ou priori conjugada Normal-Gama as posteriores dos parâmetros ϕ é uma t -Student então podemos usar a variância estimada $Var(\hat{\phi})$ para construir intervalos de confiança $\phi = \hat{\phi} \pm t_{\alpha/2} \sqrt{Var(\hat{\phi})}$, ou testar hipóteses $H_0 : \phi_i = 0$, $i = 1, 2$. Porém no caso de prioris t -Student a posteriori dos parâmetros não tem uma distribuição conhecida, portanto o intervalo de credibilidade é calculado empiricamente dos dados simulados, e os resultados obtidos são apresentados na Tabela 7.

Tabela 7 Intervalo de Credibilidade (0.95) para ϕ e τ

	ϕ_1	IC		ϕ_2	IC		τ	IC	
série 1	0.8181	0.6906;	0.9026	-0.5783	-0.6834;	-0.4617	0.4147	0.3583;	0.4967
série 2	-0.5127	-0.6390;	-0.3959	0.2965	0.1681;	0.4226	1.0961	0.9017;	1.2823
série 3	0.5811	0.4966;	0.6672	0.1824	0.0922;	0.2629	2.0505	1.8187;	2.2747

As Tabelas 8, 9 e 10 apresentam as previsões 5 passos a frente, através das abordagens clássica e bayesiana com o uso da priori Normal-Gama, para cada uma das três séries analisadas neste trabalho.

Tabela 8 - previsão 5 passos a frente para série 1

série real	prev.clássica	prev.bayesiana	int.credibilidade	
-0.9958	-0.3539	-0.3530	-0.4049	-0.1985
0.1286	-0.6941	-0.6917	-0.7641	-0.4844
-1.0206	-0.3618	-0.3579	-0.4001	-0.2382
1.0503	0.1056	0.1098	0.0526	0.2892
0.4561	0.2950	0.2966	0.2276	0.5167
EMQ	0.4883	0.4871	---	

Tabela 9 - previsão 5 passos a frente para série 2

série real	prev.clássica	prev.bayesiana	int.credibilidade	
-2.2663	-2.4515	-2.4487	-2.5581	-2.1280
2.3289	2.1505	2.1509	1.9932	2.6233
-1.5616	-1.8317	-1.8372	-2.0252	-1.2832
0.8875	1.5785	1.5939	1.3621	2.3098
0.4439	-1.3540	-1.3780	-1.5928	-0.7598
EMQ	0.7698	0.7918	---	

Tabela 10 - previsão 5 passos a frente para série 3 (Furnas)

série real	prev.clássica	prev.bayesiana	int.credibilidade	
-0.9881	-0.4809	-0.4809	-0.5077	-0.4021
-0.7129	-0.3643	-0.3631	-0.3872	-0.2916
-0.3989	-0.2998	-0.3004	-0.3287	-0.2190
-0.1546	-0.2410	-0.2420	-0.2695	-0.1625
0.1625	-0.1950	-0.1970	-0.2236	-0.1209
EMQ	0.1048	0.1052	---	

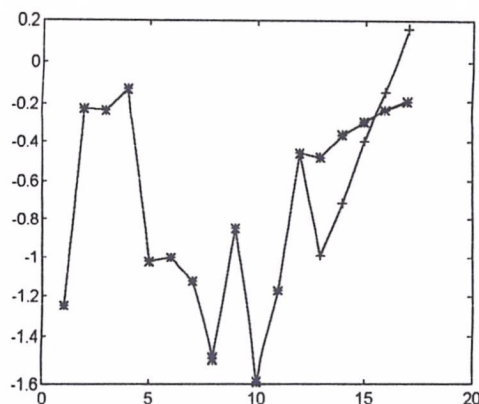


Figura 8 - gráfico da previsão 5 passos a frente para a série 3(Furnas)
(+)=real; (*)=previsto.

Nas Tabelas 8, 9 e 10, observamos um bom desempenho do método bayesiano. A Figura 8 que segue mostra a previsão, para a série de Furnas para 5 valores futuros. Nesta figura apresetamos também os valores da série observados após as previsões terem sido feitas. Notamos que as previsões acompanham a tendência da série. Um comportamento semelhante se observou com as demais séries analisadas neste trabalho.

7. Referências

Akaike, H., (1974). A new look at the statistical identification model. *IEEE trans. Auto. Control*, 19, 716-723.

Box, G.E.P. e Jenkins, G.M., Reinsel, (1994). Time Séries Analysis: Forecasting and Control, Holden-Day, 3th edition.

Box, G.E. e Tiao, G.C., (1973). Bayesian Inference in Statistical Analysis, NewYork: Addison Wesley.

Broemeling, L.D., e Shaarawy, S., (1988), Time Series: A Bayesian Analysis in Time Domain, in Bayesian Analysis of TimeSeries an Dynamic Models, *Marcel Dekker*.

Casella, G. e George, E.J., (1992), Explaining Gibbs Sampler, *Amer. Statis. Assoc.*, vol. 46, No.3.

Chib, S. e Greenberg, E., (1995), Understanding the Metropolis-Hastings Algorithm, *Amer. Statis. Assoc.*, vol.49, No.4.

Cowles, M.K. e Carlin, B.P., (1996), Markov Chain Monte Carlo Convergence Diagnostics: A comparative Review, *Journal of the American Statistical Association*, 91, 883-904.

de Alba, Enrique, (1993), Constrained forecasting in autoregressive time serie models: A Bayesian analysis, *Int. Journal of forecasting* 9, 95 – 108.

Gelman, A.E., Rubin, D., (1992). Inference from iterative simulation using multiple sequences, *Statistical Science*, 7, 457-472.

McCulloch, R.E. e Tsay, R., (1994), Bayesian Analysis of Autoregressive Time Series via The Gibbs Sampler, *Journ. of Time Series Analysis*, vol.15, No.2, 235-250.

Schwarz, G., (1978). Estimating the Dimension of a Model, *The Annals of Statistics*, vol.6, No.2, 461-464.

Tierney, L., Kass, R.E. e Kadane, J.B., (1986). Fully exponential Laplace approximations for expectations and variances of nonpositive functions. *Journal of the American Statistical Association*, 84, 710-716.