A Bayesian Analysis of Autoregressive Models with Random Normal Coefficients

[1]Thelma Sáfadi and [2]Pedro A. Morettin

[1] Federal University of Lavras and [2] University of São Paulo, Brazil

Abstract

In this paper we consider a Bayesian analysis for an autoregressive model with random normal coefficients (RCA). For the proposed procedure we use conjugate priors for some parameters and improper vague priors for others. The inference for the parameters is made via Gibbs sampler and the convergence is assessed with multiple chains and Gelman and Rubin criterium. Forecasts are based on the predictive density of future observations. Some remarks are also made regarding order determination and stationarity. Applications to simulated and real series are given.

Key Words: Autoregressive models, Bayesian analysis, Gibbs sampler, Predictive distribution.

1 Introduction

In the last years non-linear models have been studied thoroughly and their analysis is facilitated due to increasing developments in computational methodologies. We mention specially the MCMC (Markov chain Monte Carlo) methods.

Several authors in various areas of applications have considered the use of linear models with random coefficients. The random coefficient regression models for example have been widely used in economics, sociology and industrial panel data in applications where it is reasonable to assume that the modelling conditions change over time.

The problem of classical parameter estimation for autoregressive models with random coefficients is presented by Nicholls and Quinn(1980) and Quinn and Nicholls(1981).

The use of Bayesian procedure has the advantage that any information available, prior of the collection of the data, can be easily incorporated. Also, the exact posterior distribution may be obtained for some cases, eliminating the use of asymptotic results.

A Bayesian analysis of the random coefficient regression models was done by Liu(1981). Liu and Tiao(1980) used a Bayesian approach to fit a random coefficient first order autoregressive model to panel data in a time series context. The coefficient of the model changed for the different series and its value was assumed to be generated by a rescaled beta density. In such a problem, they had a panel of short time series, e.g., growth curves of plants or animals, or observations on some economic indicator from various regions of a country.

A comparison of forecasts for ARMA models between a random coefficient approach and a Bayesian approach was presented by Liu(1995). He proved analytically that the point forecast and the forecast variation via the Bayesian approach could be represented by formulas similar to those obtained via the random coefficient approach. Specially when the improper prior density is applied, the point forecast for the Bayesian approach is exactly equal to the one obtained via the random coefficient approach.

A Bayesian forecasting for an AR(1) models with normal coefficients was entertained by Diaz(1990), where numerical methods were used since it was not possible to obtain the marginal posterior for some parameters in closed form.

In this work we consider a Bayesian analysis of the RCA(p) model. The estimation of the parameters is obtained using the Gibbs sampler, since it is possible in this case to derive the full conditional posterior distribution for each parameter. The convergence is assessed via the Gelman and Rubin criterium (Gelman and Rubin, 1992).

The paper is organized as follows. The RCA model is introduced in Section 2 and a posterior analysis is given in Section 3. The predictive distribution is entertained in Section 4. Two applications are given in Section 5 and some futher comments are made in Section 6.

2 The RCA Model

A time series $\{Y_t\}$ follows a random coefficient autoregressive model of order p, RCA(p), if Y_t satisfies

$$Y_t = \sum_{i=1}^{p} (\beta_i + B_i(t)) Y_{t-i} + e_t, \tag{1}$$

where

(i) $\{e_t\}$ is a sequence of independent, identically distributed (i.i.d) random variables, with zero means and common precision τ;

(ii) the β_i are unknown constant parameters, for $i = 1, ..., p$;

(iii) if $B(t) = (B_1(t), ..., B_k(t))^{'}$, then $\{B(t)\}$ is a sequence of independent $p \times 1$

random vectors with zero means and $E(B(t)B^{'}(t)) = C$;

(iv) $B(t)$ and e_t are independent.

It is fairly straightforward to extend (1) to a q-dimensional vector time series $\{\mathbf{Y}_t\}$ by replacing the β_i by $q \times q$ matrices and $\{e_t\}$ by a q-dimensional time series. Conditions (i)-(iv) are easily generalized. For details see Tong(1990).

In this paper we consider the RCA(p) model , where the autoregressive parameter θ_{jt}, follows the model $\theta_{jt} = \mu_j + \epsilon_{jt}$, with $e_t \sim N(0, \tau^{-1})$ and $\epsilon_{jt} \sim N(0, \xi^{-1})$, independent of each other. The RCA(1) model, where the autoregressive parameter $\theta_t = \beta_1 + B_1(t)$, follows the model $\theta_t = \mu + \epsilon_t$, was analyzed by Diaz(1990), with the purpose of determining the predictive density of future observations. Diaz(1990) noted that although it was not possible to obtain a closed form for the marginal densities of μ and Y_{n+1}, it was not difficult to use numerical integration to obtain their means and variances. The marginal posterior densities for μ and Y_{n+1} are more complicatd to obtain, because one has a perfom a numerical integration with respect to ξ for each one of the values of μ and Y_{n+1}, for the corresponding joint posterior densities.

Our aim is to make inferences on the parameters through the Gibbs sampler. We consider multiple sequences and the Gelman and Rubin criterium, R, to verify the convergence (Gelman and Rubin, 1992). One-step ahead forecasts also will be considered. Details on the Gibbs sampler can be seen in Casella and George(1992) and Gamerman(1997).

The model to be considered is then

$$\begin{aligned} Y_t &= \sum_{j=1}^{p} \theta_{jt} Y_{t-j} + e_t, \quad t = 1, \ldots, n, \\ \theta_{jt} &= \mu_j + \epsilon_{jt}, \;\; j = 1, ..., p, \end{aligned} \tag{2}$$

where $e_t \sim N(0, \tau^{-1})$ and $\epsilon_{jt} \sim N(0, \xi^{-1})$, independent of each other. The parameters μ_j are supposed to be real.

We assume that the order p is fixed. Otherwise it may be selected through some information criterion, such as AIC (the Akaike information criterion) or BIC (the Bayesian information criterion). Another possibility is to use a reversible jump MCMC computation (see Green, 1995 or Chen et al., 2000). This will not be further pursued in this work.

3 Bayesian Analysis

The likelihood function for model (2) is given by

$$\mathcal{L}(\theta,\tau,\xi,\mu|D) \;\propto\; \tau^{n/2}\exp\{-\frac{\tau}{2}[\sum_{t=1}^{n}(Y_t-\sum_{j=1}^{p}\theta_{jt}Y_{t-j})^2]\}(\xi)^{pn/2}\exp\{-\frac{\xi}{2}[\sum_{t=1}^{n}(\sum_{j=1}^{p}(\theta_{jt}-\mu_j)^2)]\}$$

$$\propto\; \tau^{n/2}\xi^{np/2}\exp\{-\frac{\tau}{2}[\sum_{t=1}^{n}(Y_t-\sum_{j=1}^{p}\theta_{jt}Y_{t-j})^2]-\frac{\xi}{2}[\sum_{t=1}^{n}(\sum_{j=1}^{p}(\theta_{jt}-\mu_j)^2)]\}, \quad (3)$$

where $\theta = (\theta_1, \theta_2, ..., \theta_n)'$, $\theta_j = (\theta_{j1}, \ldots, \theta_{jn})^{'}$, $\mu = (\mu_1, \ldots, \mu_p)^{'}$ and $D = (Y_1, Y_2, ..., Y_n)^{'}$ is the vector of observations, with $Y_0, Y_{-1}, ...Y_{1-p}$ assumed fixed.

For the parameters μ, τ and ξ we consider the independent priors $\tau \sim \text{Gamma}(\alpha_1, \beta_1)$, $\xi \sim \text{Gamma}(\alpha_2, \beta_2)$ and for $\mu_j \sim \text{Normal}(m_j, \tau_j), j = 1, ..., p$:

$$P(\tau) \;\propto\; \tau^{\alpha_1-1}\exp\{-\tau\beta_1\}, \quad P(\xi) \;\propto\; \xi^{\alpha_2-1}\exp\{-\xi\beta_2\},$$

$$P(\mu_j) \;\propto\; \tau_j^{\frac{1}{2}}\exp\{-\frac{\tau_j}{2}(\mu_j-m_j)^2\}.$$

It follows that the joint prior distribution is given by

$$P(\mu,\tau,\xi) \;\propto\; \tau^{\alpha_1-1}\xi^{\alpha_2-1}\exp\{-\frac{1}{2}[2\tau\beta_1+2\xi\beta_2+\sum_{j=1}^{p}\tau_j(\mu_j-m_j)^2]\}. \quad (4)$$

Theorem 1. *The joint posterior distribution, obtained by combining (3) and (4), is given by*

$$P(\theta,\tau,\xi,\mu|D) \;\propto\; \tau^{\frac{n+2\alpha_1}{2}-1}\xi^{\frac{pn+2\alpha_2}{2}-1}\times$$

$$\exp\{-\frac{\tau}{2}[2\beta_1+\sum_{t=1}^{n}(Y_t-\sum_{j=1}^{p}\theta_{jt}Y_{t-j})^2]-\frac{\xi}{2}[2\beta_2+\sum_{t=1}^{n}\sum_{j=1}^{p}(\theta_{jt}-\mu_j)^2]-\frac{1}{2}\sum_{j=1}^{p}\frac{\tau_j}{2}(\mu_j-m_j)^2\}. \quad (5)$$

This set of prior densities produces standard posterior densities that are easy to work with. To implement the Gibbs sampler, we need to derive the conditional posterior distribution of each parameter given all the others (full conditional distributions).

Theorem 2. *The full conditional distributions derived from (5) are given by:*

$$\tau|\theta,\xi,\mu,D \;\sim\; \text{Gamma}(\frac{n+2\alpha_1}{2},\frac{1}{2}(2\beta_1+\sum_{t=1}^{n}(Y_t-\sum_{j=1}^{p}\theta_{jt}Y_{t-j})^2, \tag{6}$$

$$\xi|\tau,\theta,\mu,D \;\sim\; \text{Gamma}(\frac{np+2\alpha_2}{2},\frac{1}{2}(2\beta_2+\sum_{t=1}^{n}\sum_{j=i}^{p}(\theta_{jt}-\mu_j)^2), \tag{7}$$

$$\mu_j|\theta,\tau,\xi,D \;\sim\; N(\frac{1}{n\xi+\tau_j}(\xi\sum_{t=1}^{n}\theta_{jt}+\tau_j m_j),(n\xi+\tau_j)^{-1}), j=1,...,p \tag{8}$$

and for $k=1,...,p$, and $t=1,...,n$, we have

$$\theta_{kt}|\theta_{(-k)},\mu,\tau,\xi,D \;\sim\; N(\frac{\xi\mu_k+\tau Y_{t-k}(Y_t-\sum_{j\neq k}\theta_{jt}Y_{t-j})}{(\tau Y_{t-k}^2+\xi)},(\tau Y_{t-k}^2+\xi)^{-1}). \tag{9}$$

Proof. The distributions of $\tau|\theta,\mu,\xi,D$ and $\xi|\theta,\mu,\tau,D$ are easily obtained, since

$$P(\tau|\theta,\mu,\xi,D) \;\propto\; \tau^{\frac{n+2\alpha_1}{2}-1}\exp\{-\frac{\tau}{2}[\sum_{t-1}^{n}[(Y_t-\sum_{j=1}^{p}\theta_{jt}Y_{t-j})^2]+2\beta_1]\}$$

and

$$P(\xi|\theta,\mu,\tau,D) \;\propto\; \xi^{\frac{pn+2\alpha_2}{2}-1}\exp\{-\frac{\xi}{2}[2\beta_2+\sum_{t=1}^{n}(\sum_{j=1}^{p}(\theta_{jt}-\mu_j)^2)]\}.$$

For $\mu|\theta,\tau,\xi,D$ let us notice that

$$P(\mu|\theta,\tau,\xi,D) \;\propto\; \exp\{-\frac{\xi}{2}[\sum_{t=1}^{n}\sum_{j=1}^{p}(\theta_{jt}-\mu_j)^2]-\frac{1}{2}\sum_{j=1}^{p}\tau_j(\mu_j-m_j)^2\}$$

$$\propto \prod_{j=1}^{p} \exp\{-\frac{1}{2}[\xi \sum_{t=1}^{n} (\mu_j - \theta_{jt})^2 + \tau_j(\mu_j - m_j)^2]\}$$

$$\propto \prod_{j=1}^{p} \exp\{-\frac{1}{2}[(n\xi\mu_j^2 - 2\xi\mu_j \sum_{t=1}^{n} \theta_{jt} + \xi \sum_{t=1}^{n} \theta_{jt}^2 + \tau_j\mu_j^2 - 2\tau_j m_j \mu_j + \tau_j m_j^2]\}$$

$$\propto \prod_{j=1}^{p} \exp\{-\frac{n\xi + \tau_j}{2}[\mu_j - (n\xi + \tau_j)^{-1}(\xi \sum_{t=1}^{n} \theta_{jt} + \tau_j m_j)]^2\}.$$

Then for $j = 1, ..., p$ we have

$$P(\mu_j|\tau, \xi, \theta, D) \propto \exp\{-\frac{n\xi + \tau_j}{2}[\mu_j - (n\xi + \tau_j)^{-1}(\xi \sum_{t=1}^{n} \theta_{jt} + \tau_j m_j)]^2\}.$$

For the parameter θ we have that

$$P(\theta|\mu, \tau, \xi, D) \propto \exp\{-\frac{\tau}{2} \sum_{t=1}^{n} (Y_t - \sum_{j=1}^{p} \theta_{jt} Y_{t-j})^2 - \frac{\xi}{2} \sum_{t=1}^{n} \sum_{j=1}^{p} (\theta_{jt} - \mu_j)^2\}$$

$$\propto \prod_{t=1}^{n} \exp\{-\frac{1}{2}[\tau(Y_t - \sum_{j=1}^{p} \theta_{jt} Y_{t-j})^2 + \xi \sum_{j=1}^{p} (\theta_{jt} - \mu_j)^2]\}$$

Then, for $j = k$, we have

$$P(\theta_k|\theta_{(-k)}, \mu, \tau, \xi, D) \propto \prod_{t=1}^{n} \exp\{-\frac{1}{2}[(\tau(Y_t - \sum_{j \neq k} \theta_{jt} Y_{t-j}) - \theta_{kt} Y_{t-k})^2 +$$

$$\xi(\theta_{kt} - \mu_k)^2 + \xi \sum_{j \neq k} (\theta_{jt} - \mu_j)^2]\}$$

$$\propto \prod_{t=1}^{n} \exp\{-\frac{1}{2}\{\tau[(\theta_{kt} Y_{t-k})^2 - 2Y_{t-k}(Y_t - \sum_{j \neq k} \theta_{jt} Y_{t-j})\theta_{kt} +$$

$$(Y_t - \sum_{j \neq k} \theta_{jt} Y_{t-j})^2] + \xi[\theta_{kt}^2 - 2\mu_k\theta_{kt} + \mu_k^2]\}$$

$$\propto \prod_{t=1}^{n} \exp\{-\frac{1}{2}[\theta_{kt}^2(\tau Y_{t-k}^2 + \xi) - 2[\tau Y_{j-k}(Y_t - \sum_{j\neq k}\theta_{jt}Y_{t-j}) + \xi\mu_k]\theta_{kt}]\}$$

$$\propto \prod_{t=1}^{n} \exp\{-\frac{\tau Y_{t-k}^2 + \xi}{2}[\theta_{kt} - (\tau Y_{t-k}^2 + \xi)^{-1}(\xi\mu_k + \tau Y_{t-k}(Y_t - \sum_{j\neq k}\theta_{jt}Y_{t-j}))]^2\}$$

Noting that $\theta_{kt}|\mu_k$ are conditionally independent, we have for $t = 1, ..., n$,

$$P(\theta_{kt}|\theta_{(-k)}, \mu, \tau, \xi, D) \propto \exp\{-\frac{\tau Y_{t-k}^2 + \xi}{2}[\theta_{kt} - (\tau Y_{t-k}^2 + \xi)^{-1}(\xi\mu_k + \tau Y_{t-k}(Y_t - \sum_{j\neq k}\theta_{jt}Y_{t-j}))]^2\}$$

4 Predictive Density

In this section we derive the density of Y_{n+1} conditioned on all parameters and past observations, in order to use the Gibbs sampling procedure to get one-step ahead forecasts.

The joint density of Y_{n+1} and $\theta_{n+1} = \{\theta_{1,n+1}, ..., \theta_{p,n+1}\}$, given θ, τ, ξ and μ is given by

$$\begin{aligned} P(Y_{n+1}, \theta_{n+1}|\theta, \tau, \xi, \mu) \quad \propto \quad & \tau^{\frac{1}{2}}\xi^{\frac{p}{2}} \exp\{\frac{-\tau}{2}(Y_{n+1} - \sum_{j=1}^{p}\theta_{j,n+1}Y_{n+1-j})^2 - \\ & -\frac{\xi}{2}\sum_{j=1}^{p}(\theta_{j,n+1} - \mu_j)^2\}. \end{aligned} \tag{10}$$

Combining (5) and (10) we obtain the joint density

$$\begin{aligned} P(Y_{n+1}, \theta_{n+1}, \theta, \tau, \xi, \mu|D) \quad \propto \quad & \tau^{\frac{n+2\alpha_1+1}{2}-1}\xi^{\frac{p(n+1)+2\alpha_2}{2}-1} \exp\{\frac{-\tau}{2}[2\beta_1 + \\ & \sum_{t=1}^{n+1}(Y_t - \sum_{j=1}^{p}\theta_{jt}Y_{t-j})^2] - \frac{\xi}{2}[2\beta_2 + \sum_{t=1}^{n+1}\sum_{j=1}^{p}(\theta_{jt} - \mu_j)^2]\}. \end{aligned}$$

The posterior of Y_{n+1}, given $\theta_{n+1}, \theta, \tau, \xi, \mu$ and D is given by

$$P(Y_{n+1}|\theta, \theta_{n+1}, \tau, \xi, \mu, D) \quad \propto \quad \exp\{\frac{-\tau}{2}[\sum_{t=1}^{n+1}(Y_t - \sum_{j=1}^{p}\theta_{jt}Y_{t-j})^2]\}$$

$$\propto \quad \exp\{\frac{-\tau}{2}[(Y_{n+1} - \sum_{j=1}^{p} \theta_{j,n+1} Y_{n+1-j})^2 + \sum_{t=1}^{n} (Y_t - \sum_{j=1}^{p} \theta_{jt} Y_{t-j})^2]\}$$

$$\propto \quad \exp\{\frac{-\tau}{2}(Y_{n+1} - \sum_{j=1}^{p} \theta_{j,n+1} Y_{n+1-j})^2\}.$$

Therefore, we have

$$Y_{n+1}|\theta, \theta_{n+1}, \tau, \xi, \mu, D \quad \sim \quad N(\sum_{j=1}^{p} \theta_{j,n+1} Y_{n+1-j}, \tau^{-1}). \tag{11}$$

5 Applications

In this section we consider a simulated example and an application to the analysis of the well known lynx data.

Example 1. For this application we simulated observations for a RCA(2) model and using the Gibbs sampler estimated the parameters τ, μ_1, μ_2, ξ and produced one-step ahead forecast. The distributions involved are those of Theorem 2 and (11).

We generated 201 observations for the model RCA(2), choosing for the parameters the values $\tau = 5, \mu_1 = -0.4$, $\mu_2 = 0.3$ and $\xi = 2$ and considered samples with 200 observations.We consider two parallel chains with different initial values. Table 1 gives the initial values for the parameters τ, ξ, μ_1, μ_2. For the θ_1 and θ_2 values we considered, for each chain, an arithmetic sequence, with step 0.1, starting on -0.62 and -0.42 for θ_1, and -0.92 and -1.22 for θ_2. For the hiperparameters values we considered $\alpha_1 = 1, \alpha_2 = 2, \beta_1 = 2, \beta_2 = 1$ and for μ_1, μ_2 a Normal(0,1).

For each chain we had 10000 iterations, skept the first 4000 and for the remaining observations, took one in 25, in order to have approximate independent values. The mean values obtained are shown in Table 2, together with the respective standard deviations and R factor.

In Figure 1 we show the plot of the series and its histogram. We can see the evident variability due to the random coefficients. The random nature of the coefficients θ_{1t} and θ_{2t} is the reason for the temporary explosive behavior in the series. This occurs when large values of the observations Y_t coincide with values of θ_{1t} and θ_{2t} which lie outside of the stationary region. The series becomes temporarily non stationary and, if later the values of the coefficients fall within the stationary region, the series stabilizes again. This behavior can be seen in Figure 1. See section 6 for further discussion on stationarity.

Table 1 - Initial values for the parameters

Parameters	Initial values	
	Chain 1	Chain 2
τ	2.0	4.0
ξ	0.4	0.6
μ_1	0.2	-0.5
μ_2	-0.4	0.1

The estimated value, standard deviation and R factor for the one-step ahead forecast are shown in Table 3. Figure 2 shows the convergence for the parameters as well as the for the one-step ahead forecast, together with the respective histograms.

Tabel 2 - Mean, standard deviation and Gelman and Rubin factor, R.

Parameters	True value	Estimated value	s.d.	R
τ	5.0	4.46	0.88	1.002
ξ	2.0	2.65	0.39	1.007
μ_1	-0.4	-0.45	0.09	1.004
μ_2	0.3	0.27	0.08	0.999

Table 3 - One-step ahead forecast : mean,standard deviation and Gelman and Rubin factor

Forecasts	True value	Estimated value	s.d.	R
Y_{201}	0.967	0.588	0.796	1.003

Example 2. In this example we consider the log lynx data (1821-1920), after removing the sample mean. Nicholls and Quinn (1982) have fitted a RCA(2) model of the form (1) to the log lynx using the maximum likelihood and the conditional least-squares methods.

We consider two parallel chains with different initial values. For each chain we had 10000 iterations, skept the first 4000 and for the remaining observations, took one in 45. The mean values obtained are shown in Table 4, together with the respective standard deviations and R factor.

Tabel 4 - Mean, standard deviation and Gelman and Rubin factor, R.

Parameters	Estimated value	s.d.	R
τ	9.84	1.67	1.003
ξ	6.96	1.43	0.997
μ_1	1.09	0.10	1.000
μ_2	-0.57	0.11	0.996

The estimated value, standard deviation and R factor for the one-step ahead forecast are -0.372, 0.59 and 1.008, respectvely. These results are not directly comparable with those of Nicholls and Quinn, but the estimated coefficients have the same signs and are reasonably close.

In Figure 3 we show the plot of the series and its histogram and also the convergence for the parameters and for the one-step ahead forecast, together with the respective histograms.

6 Some Remarks

The purpose of this work was to develop a Bayesian analysis for the RCA(p) model through the Gibbs sampling procedure. It was observed in the simulated series the presence of possible outliers, or a nonstationary behavior during a certain period.

Closed-form expressions were obtained for the complete distributions, so neither the Metropolis-Hastings algorithm nor numerical methods had to be employed, as those in Diaz(1990).

For the simulated example and sample sizes of 200 good estimates for the parameters were obtained, convergence being attained for chains with 5,000 iterations. Forecasts for one-step ahead have shown to be very imprecise, with high standard deviations, eventhough the R factor of Gelman and Rubin indicated convergence of the iterations.

For the real series example a better forecast for the next observation (Y_{101}) was obtained. A similar model was entertained by Nichols and Quinn (1982). A thorough analysis of the lynx series was done by Tong (1990), fitting to it several classes of models.

As in Prado et al. (2000), no explicit stationarity conditions are imposed to the autoregressive parameters at each time t. However, in many applications such parameters lie in the stationary region, that is, the roots of $1-\theta_{1t}z-\theta_{2t}z^2-\ldots-\theta_{pt}z^p=0$, for each t, lie outside the unit circle. For further discussion on stationarity for RCA models see Nichols and Quinn (1982, ch.2).

West (1997) assumes that the coefficients in (2.2) evolve according to a random walk. A review of this class of models, called TVAR (for "time varying autoregressive") is given by Prado et al. (2000). Further references on autoregressive and RCA models are Yang and Chen (1995), Barnett et al. (1996) and Barbieri and O'Hagan (1997).

References

Barbieri, M.M. and O' Hagan, A. (1997). A reversible jump MCMC sampler for Bayesian analysis of ARMA time series. TR, Department of Statistics, Universitá La Sapienza.

Barnett, G., Kohn, R. and Sheather, S. (1996). Bayesian estimation of an autoregressive model using Markov Chain Monte Carlo. *Journal of Econometrics*, **74**, 237–254.

Casella, G. and George, E.I.(1992).Explaining the Gibbs sampler. *The American Statistician*, **46** ,167–174.

Chen, M.-H., Shao, Q.-M. and Ibrahim, J.G. (2000). *Monte Carlo Methods in Bayesian Computation*. New York: Springer.

Cowles, M.K. and Carlin, B.P.(1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, **91**, 883–904.

Diaz, J.(1990). Bayesian forecasting for AR(1) models with normal coefficients. *Communications in Statistics- Theory and Methods*, **19**, 2229–2246.

Gamerman, D.(1997). *Markov Chain Monte Carlo*. Chapman and Hall.

Gelman, A. and Rubin, D.B.(1992). Inference from iterative simulation using multiple sequence. *Statistical Science*, **7**,457–511.

Green, P. (1995). Reversible jump MCMC computation and Bayesian model determination. *Biometrika*, **82**, 711–731.

Liu, L.M.(1981). Estimation of random coefficient regression models. *Journal of the Statistical Computation Simulation*, **13**, 27–39

Liu, L.M. and Tiao, G.C.(1980). Random coefficient first order autoregressive models. *Journal of Econometrics*, **13**, 305–325.

Liu, S.I.(1995), Comparison of forcasts for ARMA models between a random coefficient approach and a Bayesian approach. *Communications in Statistics-*

Theory and Methods, **24**, 319–333.

Nicholls, D.F. and Quinn, B.G.(1980). The estimation of random coefficient autoregressive models.I. *Journal of the Time Series Analysis*, **1**, 37–46.

Nicholls, D.F. and Quinn, B.G.(1982). *Random Coefficient Autoregressive Models: An Introduction.* Lecture Notes in Statistics. Vol. 11. New York: Springer.

Prado, R. Huerta, G. and West, M. (2000). Bayesian time-varying autoregressions: Theory, methods and applications. *Resenhas Journal*, **4**, 405–422.

Quinn, B.G. and Nicholls, D.F.(1981). The estimation of random coefficient autoregressive models.II. *Journal of the Time Series Analysis*, **2**, 185–203.

Tong, H.(1990).*Non-linear Time Series: A Dynamical Approach*. Oxford: Oxford University Press.

Yang, R. and Chen, M.-H. (1995). Bayesian analysis for random coefficient regression models using noninformative priors. *Journal of Multivariate Analysis*, **55**, 283–311.

West, M. (1997). Time series decomposition. *Biometrika,* **84**, 489–494.

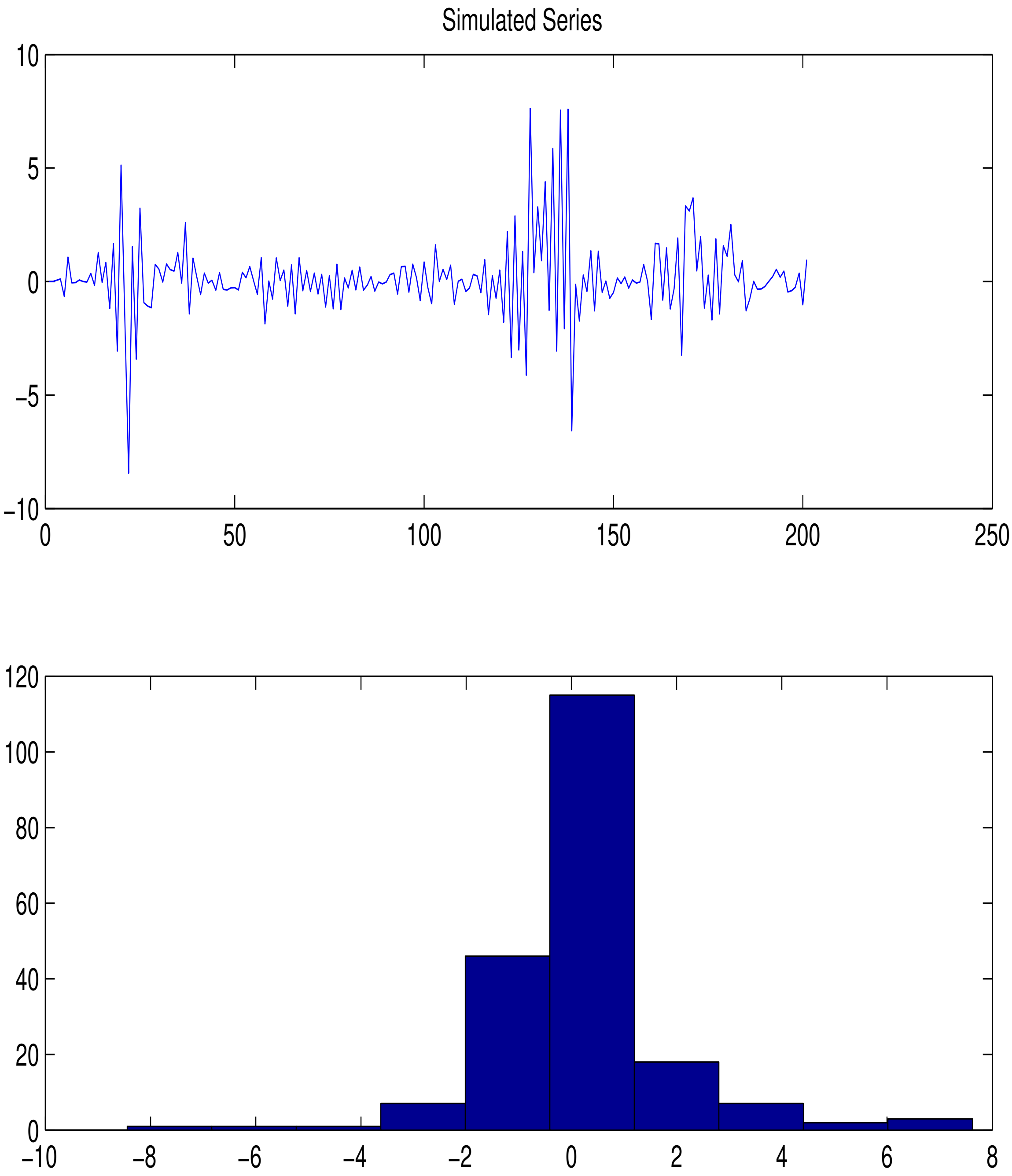

Figure 1: Plot for observations (n=200) and its histogram. Simulated data.

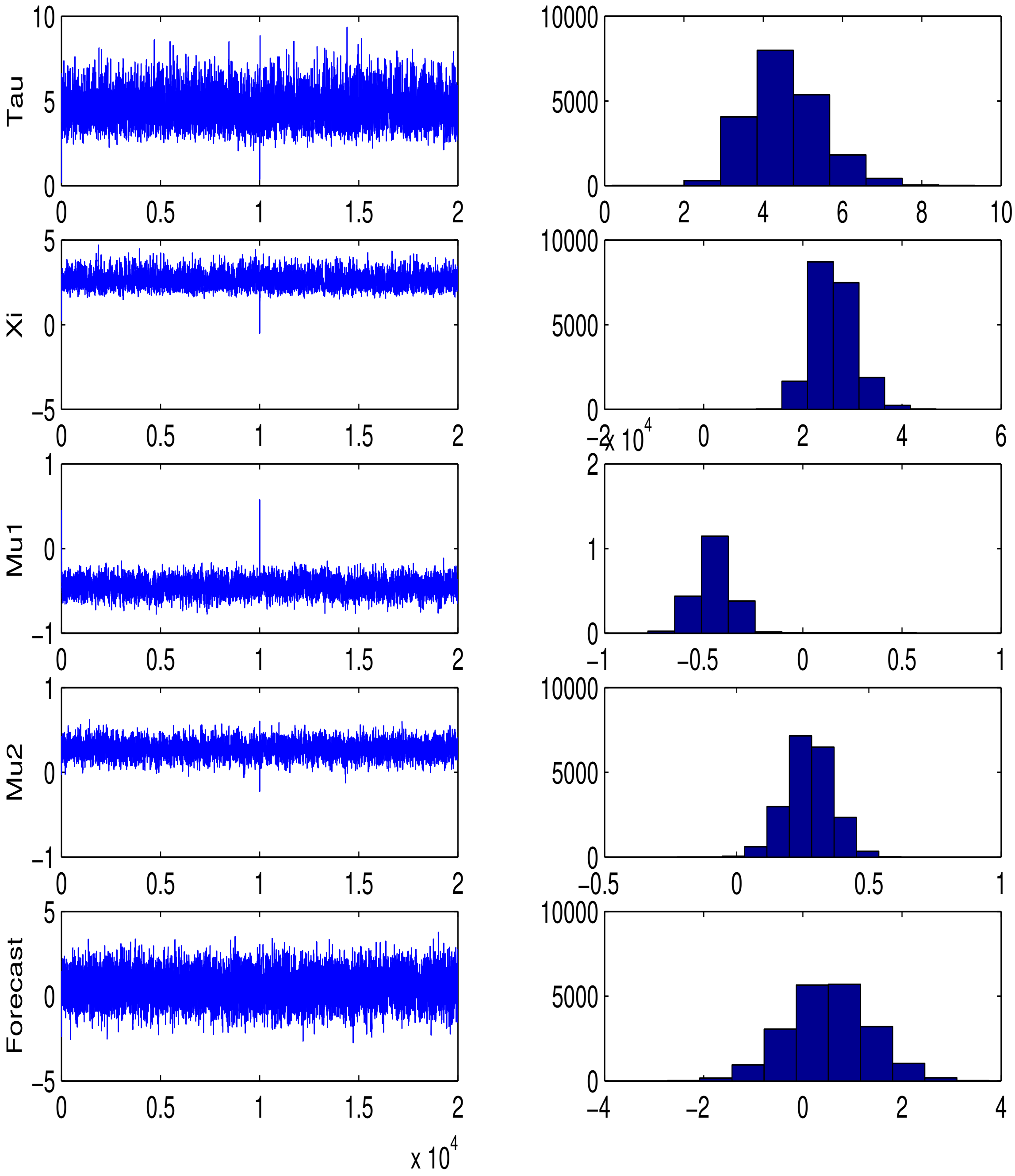

Figure 2: Convergence for the parameters τ, ξ, μ_1, μ_2, for the one-step ahead forecasts and histograms. Simulated data.

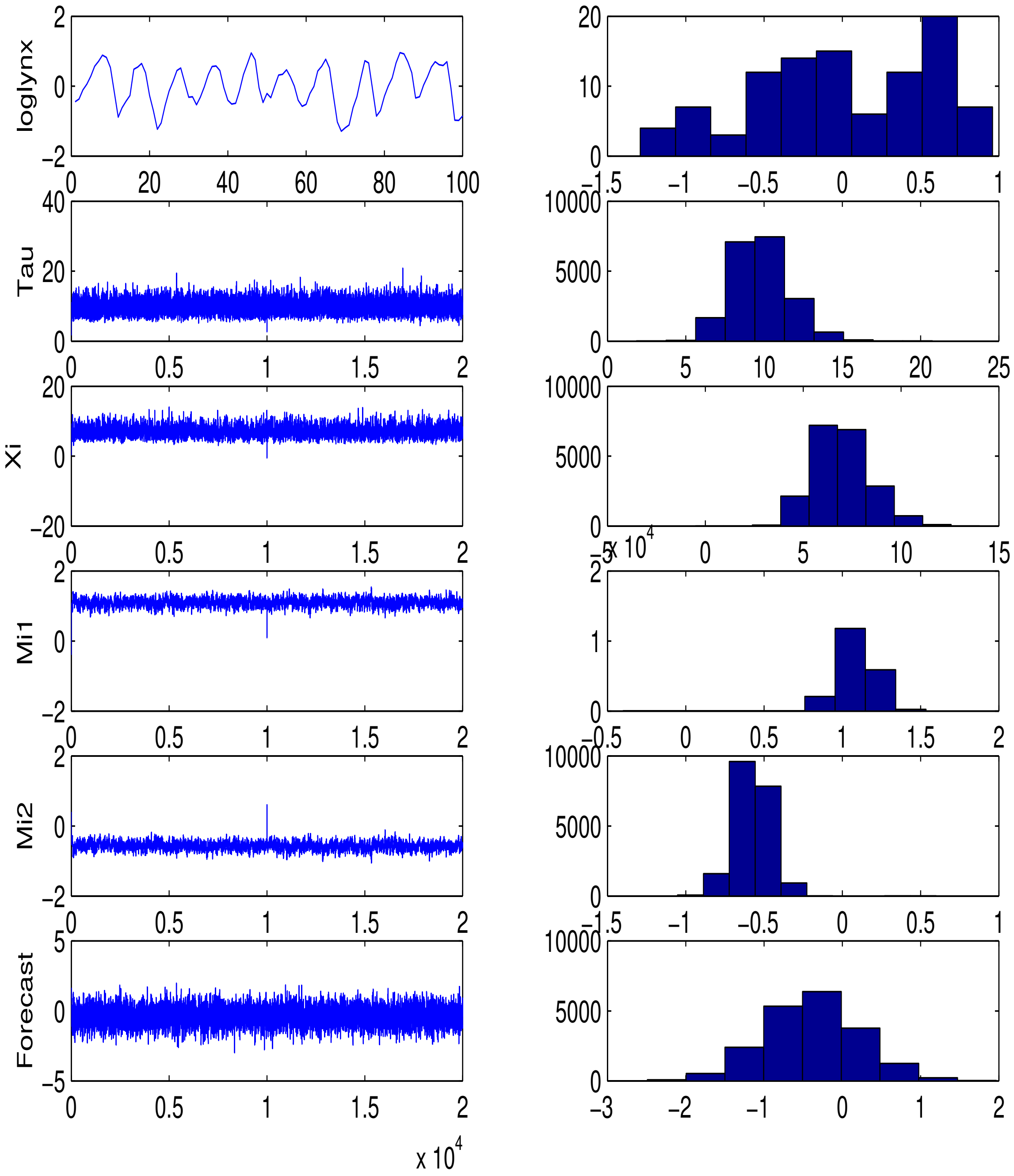

Figure 3: Plot for log lynx data (1821-1920) and its histogram. Convergence for the parameters τ, ξ, μ_1, μ_2, for the one-step ahead forecasts and histograms