Bayesian forecasting and inference in latent structure for the Brazilian Industrial Production Index

Gabriel Huerta
Centro de Investigación en Matemáticas
Apartado Postal 402. Guanajuato, Gto., 36000, México.

and

Hedibert Freitas Lopes
Institute of Mathematics, Federal University of Rio de Janeiro
Caixa Postal 68530. Rio de Janeiro, RJ 21945-970, Brazil

Summary. We consider the analysis of the Brazilian industrial production index (IPI) using statistical tools recently developed for time series. The main purpose is short-term forecasting and structural decomposition of the data through an autoregressive model that allows, but not imposes, nonstationary behavior. A very strong point of this model is that it incorporates all kinds of uncertainties by averaging forecasts across competing models, weighted by their posterior probabilities, in contrast with traditional analyses which assign probability one to a particular model. Additionally, the model considers innovation errors with heavy-tailed distributions and consequently acomodates for outlying observations. We interpret the results of the analysis in terms of its relation to the Brazilian economy.

Some key words: Bayesian time series; Autoregressive component models; Model uncertainty; Bayesian forecasting; Heavy-tailed errors.

1 Introduction

Several authors had explained and made short and long term predictions of the Brazilian Industrial Production Index (IPI) or some of its variants. For example, Lopes *et al.* (1999), Schmidt *et al.* (1999), Gamerman and Moreira (1998) use dynamic linear models (DLMs) with local linear trends, seasonality and cycles, to describe the behavior of the monthly observed IPI. Although the analyses based on these models are relevant for the IPI , they do not account for the uncertainty due to the specifications of the trend/seasonal/cycle terms. All of these papers produce their results by using a particular model selected with some optimal criteria.

Enormous amount of work has been devoted to develop useful models, but little to incorporate model uncertainty as another crucial aspect in statistical analysis. For instance, Draper (1995) states that model uncertainty should be taken very seriously to produce forecasts and obtain parameter estimates. In time series, model uncertainty was introduced by Harrison and Stevens (1976) that developed the multi-process approach to combine aspects of different DLMs under consideration. More recently Barnett *et al.* (1996), Barbieri and O'Hagan (1997), Troughton and Godsill (1997), Huerta and West (1999) developed Markov Chain Monte Carlo (MCMC) methods to incorporate model uncertainty within a linear autoregressive (AR) framework.

Specifically, in this paper we analyze the Brazilian IPI using the AR model with prior specifications on latent components and characteristic roots as in Huerta and West (1999). These specifications lead to a new class of prior distributions in autoregressive component structure which has the following properties:

- they permit arbitrary collections of real and complex conjugate pairs of characteristic roots;
- they allow for zero values among the characteristic roots, so taking care of prior uncertainty about model order;
- they allow unit roots, and so cater for persistent low frequency trends and sustained quasi-periodic components;
- they incorporate unobserved initial values of the data process as uncertain latent variables, so that all resulting inferences are formally based on incorporating full uncertainties about initial values.

In particular, such a class of priors naturally avoids any of the corrections in significance tests proposed for the unitary root problem through posterior probability statements on the number and type of unitary roots. For example, Cribari-Neto (1993) illustrates the complications of frequentist procedures when testing whether a root is unitary or not. Detailed discussion about unit root tests and their pitfalls is presented by Campbell and Perron (1991). Furthermore, the class of priors is identified by a small number of hyper-parameters, which may be chosen based on specific forms of quantitative prior information. Alternatively, these hyper-parameters can be assigned essentially uniform or "reference" prior distributions themselves, so inducing what may be viewed as a non-informative analysis.

Additionally, we extend the prior modeling of Huerta and West (1999) to allow heavy-tailed innovation errors which permits the accomodation of outlying observations. Such extensions and modeling issues are presented in Section 2, along with prior elicitation, posterior and predictive inference. We strongly believe that an AR model that fully recognizes uncertainty on the order, model parameters, number of unitary roots and considers heavy-tailed errors, is very helpful to describe economics as encompassed by the Brazilian IPI. Empirical arguments are provided in Section 3. Section 4 presents our final remarks and possible extensions.

2 Time series model and Methods

2.1 The Model and a Decomposition Result

Define $\{x_t\}$ as the realization of an AR process of order p,

$$x_t = \phi(B)\epsilon_t$$

where $Bx_t = x_{t-1}$ and $t \in \{0, 1, \dots n\}$. $\phi(u) = 1 - \phi_1 u - \dots - \phi_p u^p$ is the characteristic polynomial, $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)'$ is the vector of standard coefficients, and $\{\epsilon_t\}$ are zero-mean uncorrelated errors with $\epsilon_t \sim N(0, \sigma^2/\gamma_t)$. The quantities γ_t are assumed independent with a common distribution $p(\gamma_t)$. Note that the introduction of the parameters γ_t implies a scale mixture of Normals on the error terms which allows for heavy-tailed innovations. Some cases of a scale mixture of Normals include the Laplace, exponential power and Student t distributions. This is important in our application since it is well-known that several macroeconomic interventions took place in Brazil during the last two decades. By allowing heavy-tailed distributions for the innovations, such temporal interventions will have lower impact in the model specification and conclusions. We will return to this point later when we discuss the analysis of the IPI.

Denote by $\{\alpha_1, \dots, \alpha_p\}$ the reciprocals of the characteristic roots or solutions of the equation $\phi(u) = 0$. If $|\alpha_j| \leq 1$ for all j, the process is stationary with unitary roots if any of these moduli equal one. Assume there are C pairs of complex conjugate roots and $R = p - 2C$ real roots. Denote the complex pairs by $r_j \exp(\pm i\omega_j)$ for $j = 1, \dots, C$, and the real roots by r_j for $j = 2C+1, \ldots, p$. As presented in West (1997), it can be shown that

$$x_t = \sum_{j=1}^{C} z_{tj} + \sum_{j=2C+1}^{p} a_{tj}$$

where the z_{tj} and a_{tj} are latent processes related to the complex and real roots respectively. Corresponding to the real roots $j = 2C+1, \ldots, p$, we have

$$(1 - r_j B)a_{tj} = b_j \epsilon_t$$

for some real constants b_j; thus the a_{tj} are correlated AR processes of order one. Corresponding to the complex conjugate pairs of roots $j = 1, \ldots, C$, we have

$$(1 - 2r_j cos(\omega_j)B + r_j^2 B^2)z_{tj} = (d_j + e_j B)\epsilon_t$$

for further real constants d_j and e_j; thus the z_{tj} are AR, moving average processes of order $(2, 1)$. In other words, a series that follows an autoregressive process can be expressed as the sum of simpler processes, some of periodic behavior and some with low frequency variation. In fact, the decomposition implies that z_{tj} has a quasi-periodic behavior with frequency ω_j, or periodicity $\lambda_j = 2\pi/\omega_j$ where the damping of the component is determined by the modulus of the defining complex root. Computation of the latent components may be handled through the DLM representation of an AR model and has been exemplified in the context of oxygen-isotope series, electroencephalogram traces and other types of data, in both West (1997), West *et al.* (1999) and West and Harrison (1997).

2.2 Prior Specifications

Huerta and West (1999) introduced a class of hierarchical priors defined on the component structure of an AR time series. We briefly review these specifications here.

The prior assumes fixed but arbitrary upper bounds C_+ and R_+, on the number of complex pairs and real roots, hence an upper bound $p_+ = 2C_+ + R_+$ on model order. Independent priors are specified on the real roots, the complex roots and the error terms variance. Each real root r_j has a prior that

- gives probability $\pi_{r,0}$ to $r_j = 0$,
- gives probability $\pi_{r,-1}$ to $r_j = -1$,
- gives probability $\pi_{r,1}$ to $r_j = 1$, and
- otherwise has a continuous density $g_r(r_j)$ from -1 to 1.

Each complex conjugate pairs of roots $r_j \exp(\pm i\omega_j)$ has a prior that

- gives probability $\pi_{c,0}$ to $r_j = 0$,
- gives probability $\pi_{c,1}$ to $r_j = 1$, and
- otherwise has r_j independent of ω_j. The modulus r_j follows a continuous density $g_c(r_j)$ with support in $(0, 1)$. The wavelength $\lambda_j = 2\pi/\omega_j$ has a continuous density $h(\lambda_j)$ with support on $(2, \lambda_u)$ where λ_u is an upper bound in periods. By default, λ_u can be fixed at $n/2$, the maximum period observable for a time series of length n.

Notice that the prior is defined on the parameters that determine the time series decomposition of Section 2.1 and implicitly, quantifies prior knowledge on the latent structure of an AR model. In applications, particular forms for $g_r(\cdot)$, $g_c(\cdot)$ and $h(\cdot)$ had involved truncated Normals, Uniform densities or more general Beta distributions. A detailed exploration of how particular forms of these functions determine priors in other quantities of interest, like the standard AR coefficients, has been fully addressed in Huerta and West (1999). For the analysis of the Brazilian IPI and in a non-informative sense, we adopt the benchmark prior known as the *component reference prior* which establishes that $g_r(\cdot)$ is a Uniform on $(-1, 1)$, $g_c(\cdot)$ is a Beta(3,1) and $h(\lambda_j) \propto \sin(2\pi/\lambda_j)/\lambda_j^2$ with λ_j ranging form 2 to λ_u. The marginals of this prior, correspond to the standard reference prior obtained by treating the parameters of each component process z_{tj} and a_{tj} individually.

Furthermore, the constant scale factor that appears in each error term is assumed independent of the roots and has a specific marginal prior, usually a conditionally conjugate inverse gamma prior, i.e., $p(\sigma^2) \sim IG(a, b)$. Priors for the point-masses may be assigned as context dependent, but for simplification, we use independent uniform Dirichlet distributions, namely $Dir(\pi_{r,0}, \pi_{r,1}, \pi_{r,-1}|1, 1, 1)$ and $Dir(\pi_{c,0}, \pi_{c,1}|1, 1)$. Note that the prior point masses at zero for the numbers of roots, both complex and real, may fall below the fixed upper bounds C_+ and R_+. This implies that the model order can take any value from 0 to p_+. Also, the point masses $\pi_{r,-1}$ and $\pi_{c,1}$ permit direct inferences on the number of unitary roots distinguishing between real and complex cases, something that is known to be controversially important in macroeconomic time series analysis (Nelson and Plosser, 1982).

We must note that the roots are not identified. The model coefficients ϕ are unchanged with arbitrary permutations of the roots. Identification of real roots can be imposed simply by relabeling them in order of increasing value. Identification may be achieved for the complex roots by relabeling them in order of increasing moduli or of increasing period or wavelength.

In extension to Huerta and West (1999), we assume that each γ_t is independent of the reciprocal roots and σ^2 with a prior distribution $p(\gamma_t) \sim Ga(\alpha, \beta)$. This specification defines errors that follow a Student t distribution.

2.3 Posterior and Predictive Analysis

Posterior and predictive calculations are developed using Markov chain Monte Carlo (MCMC) methods based on the Gibbs sampler. For explanations on MCMC methods with theory and applications, we recommend Gamerman (1997) and Gilks *et al.* (1996). Chib and Greenberg (1996) and Gamerman (2000) are references of applications of MCMC techniques in econometric problems. For our MCMC, we briefly outline the form of the relevant conditional posterior distributions.

First some notation. Write $\mathbf{X} = \{x_1, \ldots, x_n\}$ for the observed time series and, given the maximum model order p_+ write $\mathbf{Y} = \{x_0, x_{-1}, \ldots, x_{-(p_+-1)}\}$ for the latent initial values. The MCMC includes formal inference on these initial values. The model parameters are denoted by

$$\boldsymbol{\psi} = \{\alpha_j,\ j = 1, \ldots, p_+;\ (\pi_{r,-1}, \pi_{r0}, \pi_{r1});\ (\pi_{c0}, \pi_{c1});\ \sigma^2;\ \gamma_t, t = 1, \ldots, n\}.$$

Posterior inferences are based on summarizing the full posterior $p(\boldsymbol{\psi}, \mathbf{Y}|\mathbf{X})$. For any subset $\boldsymbol{\xi}$ of elements of $\boldsymbol{\psi}$, let $\boldsymbol{\psi}\backslash\boldsymbol{\xi}$ denote the complementary elements, i.e., $\boldsymbol{\psi}$ with $\boldsymbol{\xi}$ removed. The MCMC method iteratively simulate elements of $\boldsymbol{\psi}$ and $\mathbf{Y}$ from their conditional posteriors with all conditioning parameters fixed at their latest sampled values. Specifically,

- for each $j = 2C_+ + 1, \ldots, p_+$, the real roots are sampled individually from

  $$p(r_j|\boldsymbol{\psi}\backslash r_j,\ \mathbf{X}, \mathbf{Y}).$$

  Assuming $g_r(r_j)$ is Uniform from -1 to 1, this conditional posterior is a mixture of a truncated Normal at $(-1, 1)$ with three points masses at 0, -1 and 1 respectively. This mixture posterior is easily sampled via CDF inversion of a truncated Normal.

- For each $j = 1, \ldots, C_+$, the complex roots are sampled individually from

  $$p(r_j, \lambda_j|\boldsymbol{\psi}\backslash(r_j, \lambda_j),\ \mathbf{X}, \mathbf{Y}).$$

  Even with simple models for $g_c(r_j)$ and $h(\lambda_j)$, this conditional posterior is difficult. A MCMC reversible jump step is used to sample from this conditional distribution which is a mixture of a continuous component with two point masses.

- The hyperparameters are sampled from conditionally independent posteriors

  $$p(\pi_{r,-1}, \pi_{r0}, \pi_{r1}|\boldsymbol{\psi}\backslash(\pi_{r,-1}, \pi_{r0}, \pi_{r1}),\ \mathbf{X}, \mathbf{Y})$$

  and

  $$p(\pi_{c0}, \pi_{c1}|\boldsymbol{\psi}\backslash(\pi_{c0}, \pi_{c1}),\ \mathbf{X}, \mathbf{Y}).$$

Assuming the Dirichlet priors introduced in Subsection 2.2, this conditional distributions are respectively $Dir(\cdot|r_{-1}+1, r_0+1, r_1+1)$ and $Dir(\cdot|c_0+1, c_1+1)$ where (r_{-1}, r_0, r_1) denote the number of real roots equal to -1,0 and 1 respectively. (c_0, c_1) are the number of complex roots with modulus 0 and 1 respectively.

- The error variance is sampled from

$$p(\sigma^2|\psi\backslash\sigma^2, \mathbf{X}, \mathbf{Y}).$$

Assuming that $p(\sigma^2) \sim IG(a, b)$, the conditional posterior follows an $IG(a^{'}, b^{'})$ where

$$a^{'} = a + n/2; \quad b^{'} = b + \sum_{t=1}^{n} \gamma_t \epsilon_t^2/2;$$

$\epsilon_t = x_t - \sum_{j=1}^{p} \phi_j x_{t-j}$ are the error innovations computed with the implied AR parameter vector ϕ obtained with the current values for the roots α_j.

- For $t = 1, \ldots, n$ each scale parameter γ_t is sampled individually from

$$p(\gamma_t|\psi\backslash\gamma_t, \mathbf{X}, \mathbf{Y}).$$

If $p(\gamma_t) \sim Ga(\alpha, \beta)$, then the conditional posterior follows a $Ga(\alpha^{'}, \beta^{'})$ where

$$\alpha^{'} = \alpha + 1; \quad \beta^{'} = \beta + \epsilon_t^2/2\sigma^2.$$

Once again, $\epsilon_t = x_t - \sum_{j=1}^{p} \phi_j x_{t-j}$ are the error innovations computed with the implied AR parameter vector ϕ obtained with the current values for the roots α_j.

- The initial values are sampled from

$$p(\mathbf{Y}|\psi\backslash\mathbf{Y}, \mathbf{X}).$$

Under the prior specifications of the previous section, the AR process is not strictly stationary but it turns out that the reverse time model produces samples from the correct conditional distribution for the initial latent values. This important result is shown in Huerta and West (1999). The simulation consists in sequentially sampling $x_0, x_{-1}, \ldots, x_{-(p_+-1)}$ in turn, conditioning on the most recent sampled values in the reverse time model $x_t = \sum_{j=1}^{p} \phi_j x_{t+j} + \epsilon_t; t = 0, \ldots, -(p-1)$ sampling ϵ_t at each step. As for previous conditional distributions, the current roots α_j imply current values for ϕ_j. Notice that this operation is essentially the same as used in sampling future values x_{n+k} for $k > 0$, in forecasting ahead from the end of the data using the forward-time model $x_t = \sum_{t=1}^{p} \phi_j x_{t-j} + \epsilon_t$.

In the next section, we implement this machinery to explore the time series behavior and to forecast some of the levels of the Brazilian IPI.

3 Analyzing the Brazilian Industrial Production Index

The data we analyze correspond to 215 monthly observations of the *Brazilian industrial production index* (IPI), from February 1980 to December 1997. The data, displayed in Figure 1, presents a strong seasonal pattern and the "ups" and "downs" characteristic of a trend. Any econometric analysis of such macroeconomic series must be performed with extra care, since the brazilian economy has suffered several macroeconomic shocks in the past two decades; some of them with temporary effects, others with permanent effects. Allowing the innovations to follow heavy-tailed distributions is a conservative way of weighing the information in the data as it becomes more or less important.

— **Figure 1 about here** —

To study different aspects of the series, a component structured AR model was considered with $C_+ = 20$ and $R_+ = 20$ which implies a maximal model order $p_+ = 60$. The MCMC described in Subsection 2.3, was iterated 10000 times with a burn-in of 5000 iterations the following 5000 samples used for posterior inference. First, we present the marginal posterior distribution for model order in Figure 2. The Figure shows that the posterior distribution for p mostly favors values from 16 to 35 and has a mode at $p = 24$. This posterior distribution reflects large uncertainty upon the lag of the AR model.

— **Figure 2 about here** —

To exhibit the model structure in terms of complex and real roots, in Figure 3 we present the marginal posterior distribution for the number of complex pairs of roots and the number of real roots. The model prefers 6, 7 or 8 complex pairs with large probability. The posterior distribution for the real roots favors a wide range of values, which is typical when components of very low frequency exist in the data.

— **Figure 3 about here** —

To summarize some of the posterior samples of the real roots, in Figure 4 we show histograms of samples for the 2 smallest and two largest real roots when the roots are ordered from lower to higher. Also, the positive probabilities of a point mass at -1,0 and 1 are reported in the Figure. It is interesting to note that the largest root (labeled $r(20)$) has 0.67 probability of being unitary and the smallest root (labeled $r(1)$) has 0.41 probability of being equal to -1. Indeed, this shows evidence that the data is non-stationary with random walks of order one driving the trend of the series.

— Figure 4 about here —

Posterior summaries for some of the complex pairs of roots appear in both Figures 5 and 6. The figures show boxplots of samples corresponding to the modulus and wavelength of 5 complex pair roots respectively. For identification, the roots were ordered by wavelength with the label "1" denoting the root with the larger period and the label "5", the root with the smallest period. The boxplots for moduli do not consider samples where the modulus is equal to one. Instead, the posterior probability of a unitary modulus is reported in the left side of Figure 5. We observe that the root that has the larger period or wavelength, has a posterior probability of being unitary equal to 0.96 and a period of about 12 time units. This complex root defines a quasi-cyclical non-stationary component that correspond to the seasonality in the data. Also, the other four roots have a positive probability of being unitary with periods at about 6, 4, 3, and 2.4 units of time. These periodicities are basically harmonics of the fundamental periodicity of 12. Notice that the fifth harmonic is more likely to correspond to a non-stationary latent component in comparison to the third and fourth harmonics.

— Figures 5 and 6 about here —

Posterior samples of the roots directly lead to samples for the components associated to the complex and real roots simply because these components are functions of the parameters in the AR model. In consequence, posterior summaries of the decomposition can be displayed as with other quantities of interest. In fact, Figure 7 presents the data with posterior means for two components corresponding to the complex roots and two components corresponding to the real roots. The quasi-cyclical component labeled by (C1) is associated to the complex pair that has a periodicity of 12 months and is essentially the underlying seasonality in the data. This component has a time-varying amplitude comparable to the amplitude presented by the series. Furthermore, the component has two high peaks between 1990-1992, a period where the brazilian economy was experiencing major macroeconomic interventions, such as the Summer Plan in February 1989, the first and second Collor's Plan in March 1990 and February 1991, respectively. Thus, the component captures the higher level of uncertainty presented in the data during the early 90's.

— Figure 7 about here —

The component labeled by (C2) corresponds to the root that has a harmonic periodicity of 6 months. It shows a very low amplitude compared to the data and all other complex components have similar low amplitudes. The component labeled by (R3) is associated to the maximal real root (r(20)) and has an amplitude comparable to the amplitude of the data. This component is basically the underlying trend of the series. The last component displayed (R4) corresponds to the smallest real root (r(1)); its amplitude is very low compared to the series and has switches characteristic of an AR(1) process with a root equal or close to -1. Mostly, components that have very low amplitude represent complicated noise structure in the information.

Posterior summaries for samples of γ_t are indicative of possible outlying observations. For instance, in Figure 8 we present 95% posterior intervals and posterior means based on 5000 posterior samples for γ_t, t corresponding to the months of February 89 to January 92. We assigned a prior for $\gamma_t \sim Ga(1,1)$ which puts .95 prior probability to values between 0.29 and 3.69. Other $Ga(\alpha,\alpha)$ prior distributions were used essentially leading to the same results. Most of the posterior intervals reported in Figure 8 are consistent with the hypothesis that γ_t may be close to one, except for the intervals corresponding to April 1990 and April 1991. Both periods are close to Collor's plans, so reassuring the strength of our modeling strategy in describing more important movements and trends of the data and giving lower weight to specific idiosyncratics. In these two cases, the small values for γ_t favor a larger error variance σ^2/γ_t. The introduction of γ_t helped the AR model to accomodate these anomalous observations.

— **Figure 8 about here** —

To consider model validation and forecasting, we implemented again the MCMC but only with the observations previous to and including January 1997. As explained in Subsection 2.3, samples of multiple step-ahead forecasts can be generated conditional on all other parameters using the autoregressive equation that defines the model. Based on 5000 of these posterior samples, Figure 9 presents the 95% predictive probability intervals and posterior means for forecasts corresponding to February 1997 until December 1997, compared against the actual observed values. In general, we notice that the posterior means are lower but close to the observed values. The predictive intevals contain the observed data, except for December 97, where the observation is below the lower limit of the corresponding interval.

— **Figure 9 about here** —

For comparison with other possible approaches, we computed forecasts with AR models that have a constant variance and the model order selected using the AIC criteria. For the Brazilian IPI, AIC leads to a model order of 13 which has zero posterior probability in Figure 2. Under AIC, we obtained the maximum likelihood estimator (MLE) of ϕ and σ^2 and generated samples of "future" values for February 1997-December 1997. Additionally, assuming the standard reference prior for the AR model, $p(\phi, \sigma^2) \propto 1/\sigma^2$, we generated samples of "future" values for the same time period with the corresponding Normal-Gamma reference posterior. In fact, using the posterior mean of the samples as point estimators of the future values, we computed the mean square error (MSE) of the AR-AIC models and our AR model that incoporates model order uncertainty, unitary roots and heavy-tailed errors. The MSE for the AR with the standard reference posterior is 67.64; with the MLE treated as the "real" parameter, the MSE is 86.06. For our AR model, the MSE is 30.14, showing that is worth the effort of recogninzing different levels of uncertainty in forecasting time series under AR models.

Furthermore, we consider the forecasts produced for the IPI with dynamic linear models that have trend/seasonal components reported in Schmidt *et al.* (1999) and Gamerman and Moreira (1998). The predictions for March 1997-August 1997 obtained with these dynamic models and the predictions we obtained with AR models are plotted in Figure 10. We observed that all the five models underestimate the actual values and the AR-AIC models have very poor predictive performance. The AR with structure prior and heavy-tailed errors outperforms both dynamic models for May and July 1997. At March and April of 1997, our AR model is only superior to the DLM of Gamerman-Moreira. For this time period, the MSE of our AR model is 25.64, for the Gamerman-Moreira DLM is 25.34 and for the model of Schmidt *et al.* (1999) is 21.03.

— **Figure 10 about here** —

In terms of predictive intervals, Figure 11 compares the 95% probability intervals for the AR models with priors on component structure and the dynamic models of Schmidt *et al.* (1999) and Gamerman and Moreira (1998) for the period that covers March 1997-August 1997. The intervals for the AR model show a general tendency for higher predictive values during this time period. Surprisingly, the length of the intervals for the AR model are smaller with respect to those lengths obtained with DLMs.

— **Figure 11 about here** —

We summarize this application section with a list of points we find interesting and a few thought-provoking issues:

- One of the main aspects of our methodology is to allow the lag-length in an AR model to be uncertain, an aspect of great importance for economic time series. Figures 2 and 3 suggest that arbitrarily choosing an specific value for the lag-length may ignore a great amount of uncertainty and perhaps lead to over-optimistic inference and conclusions.

- Related to the last point is the fact that accounting for model uncertainty does not necessarily mean to increase uncertainty when forecasting a time series (see Figure 11).
- Another important issue is that the model can accomodate outlying observations that seem to have only marginal impact in the modeling by structuring the innovations with heavy-tailed distributions (see Figure 8).
- The real/complex unitary roots found represent long-term dependency in the economy. Figure 7, for instance, reveals a stochastically changing seasonality component (C1) and a stochastic trend component (R3) in agreement with previous analyses of the IPI.

4 Summary Remarks

This paper analyzes the Brazilian industrial production index using a Bayesian methodology based on a new class of prior distributions for AR models that is extended to allow for heavy-tailed errors. The analyses show how a unified approach is able to deal with model uncertainty, inference on latent structure, inference on unitary roots, forecasting and outliers, all at once. This type of modeling avoids the imposition of trends and polynomial seasonal components to capture structure and multiple significant tests to show the presence of an underlying stochastic trend. It also avoids the use of "ad-hoc" diagnostic tools to detect outlying observations by including scale-mixtures of Normals.

On the other hand, a current limitation of the model is that the generation of samples of futures is based on drawing the respective error terms by drawing γ_t from its prior distribution or assuming them equal to one, i.e., no outlying observations are expected in the future. In this direction, we recognize the need of a full study that measures the impact of such prior specifictions or others for forecasting and accomodating outlying information in time series. This is part of future research.

Acknowledgements

The authors are grateful to Ajax Moreira for providing the Industrial Production Index series and the results obtained with DLMs. The first author was partially supported by CONACYT and the Sistema Nacional de Investigadores of México. The second author would like to thank CAPES for partial financial support during his period at the Institute of Statistics and Decision Sciences, Duke University where part of the research was developed.

BIBLIOGRAPHY

Barbieri, M.M. and O'Hagan, A. (1997) A reversible jump MCMC sampler for Bayesian analysis of ARMA time series. Technical Report. Department of Statistics Universita "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.

Campbell, J. Y. and Perron, P. (1991) Pitfalls and opportunities: What macroeconomists should know about unit roots. NBER-Working paper series 100.

Chib, S. and Greenberg, E. (1996) Markov Chain Monte Carlo simulation methods in econometrics. *Econometric Theory*, **12,** 409–431.

Cribari-Neto, F. (1993) The cyclical component in Brazilian GDP. *Revista de Econometria*, **13,** 1–22.

Draper, D. (1995) Assessment and propagation of model uncertainty (with discussion). *Journal of the Royal Statistical Society-Series* B, **57,** 45–97.

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

Gamerman, D. (2000) MCMC in econometrics. *Economia*, 7–37.

Gamerman, D. and Moreira, A. R. B. (1998) Bayesian analysis of econometric time series models using hybrid integration rules. Technical Report. Department of Statistical Methods, Federal University of Rio de Janeiro.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996) *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

Harrison, P.J. and Stevens, C.F. (1976) Bayesian forecasting (with discussion). *Journal of the Royal Statistical Society-Series* B, **38,** 205–247.

Huerta, G. and West, M. (1999) Priors and component structures in autoregressive time series models. *Journal of the Royal Statistical Society-Series* B, **61,** 881–899.

Lopes, H. F., Moreira, A. R. B. and Schmidt, A. M. (1999) Hyperparameter estimation in forecast models. *Computation Statistics and Data Analysis*, **29,** 387–410.

Nelson, C.R. and Plosser, C.I. (1982) Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics*, **10,** 139–162.

Schmidt, A. M., Gamerman, D. and Moreira, A. R. B. (1999) An adaptive resampling scheme for cycle estimation. *Journal of Applied Statistics*, **26,** 619–641.

Troughton, P.T. and Godsill, S.J. (1997) Bayesian model selection for time series using Markov Chain Monte Carlo. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*.

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

West, M. and Harrison, J. (1997) *Bayesian Forecasting and Dynamic Linear Models (2nd Edn.)*. New York: Springer-Verlag.

West, M., Prado, R. and Krystal, A. (1999) Evaluation and comparison of EEG traces: latent structure in non-stationary time series. *Journal of the American Statistical Association*, **94,** 1083–1095.

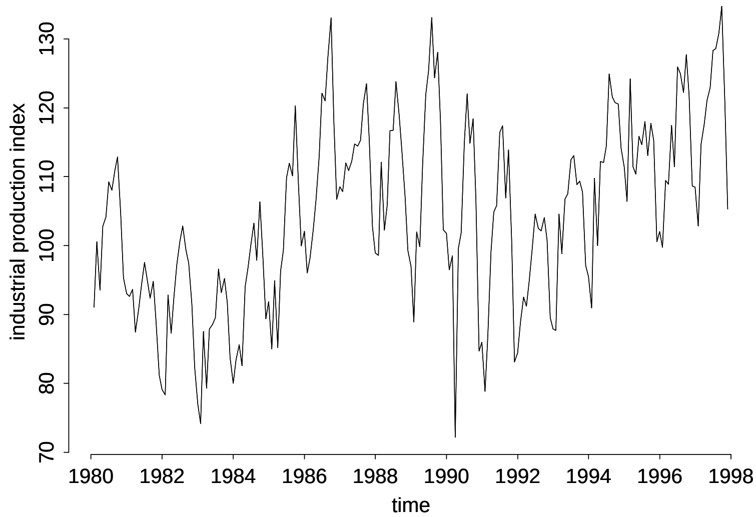

Figure 1: *Brazilian Industrial Production Index.* 215 monthly observations taken since February 1980.

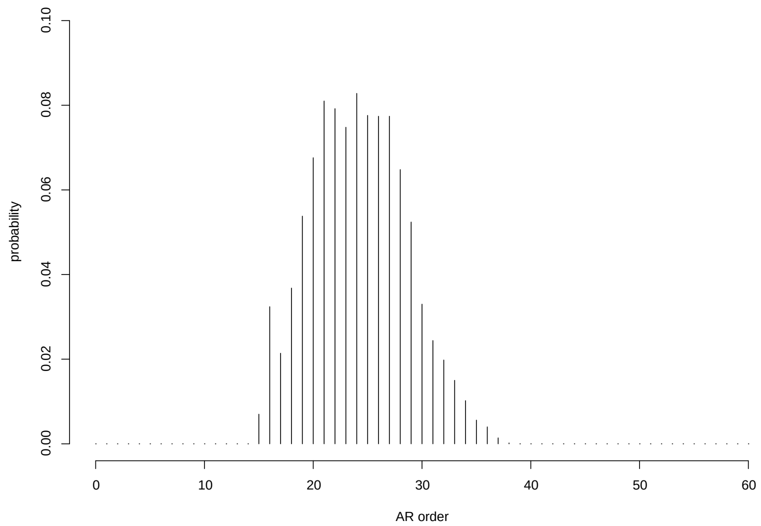

Figure 2: Posterior distribution for model order p based on 5000 posterior samples; $C_+ = 20$ and $R_+ = 20$.

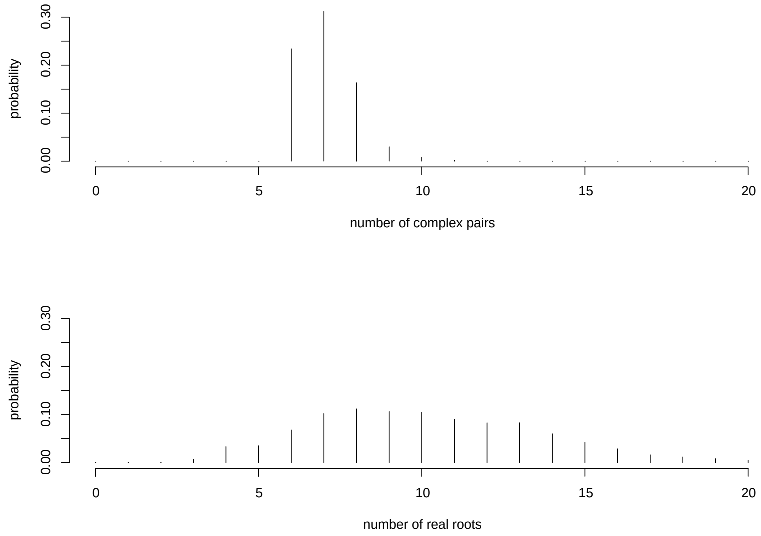

Figure 3: Posterior distributions for number of complex pairs and number of real roots based on 5000 posterior samples; $C_+ = 20$ and $R_+ = 20$

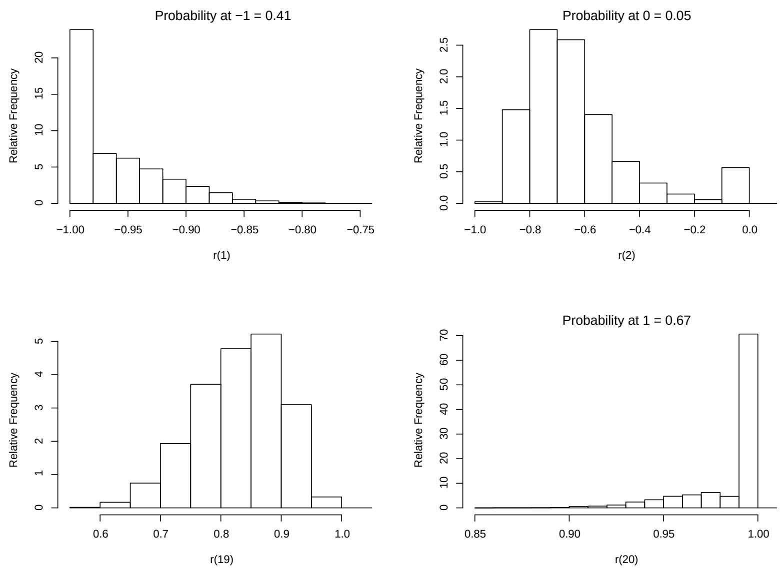

Figure 4: Histograms of samples for the two smallest and the two largest real roots with reported (positive) posterior probabilities of point masses.

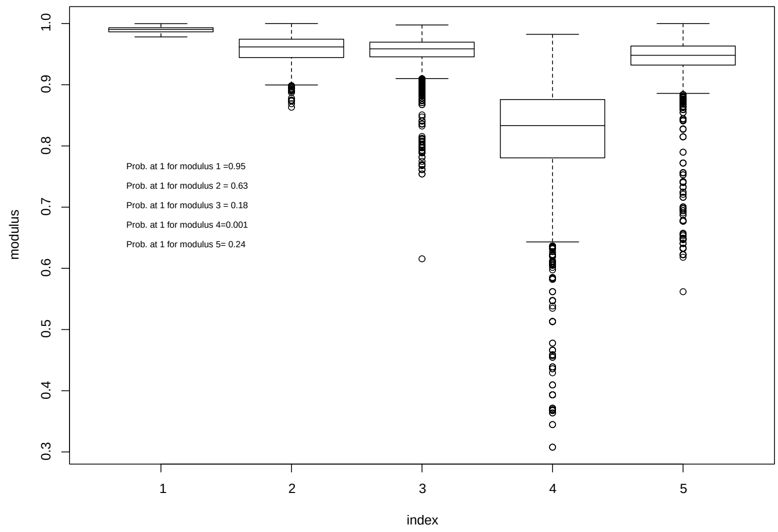

Figure 5: Boxplots of samples for moduli corresponding to the 5 largest roots ordered by wavelength with reported posterior probability of a point mass at one.

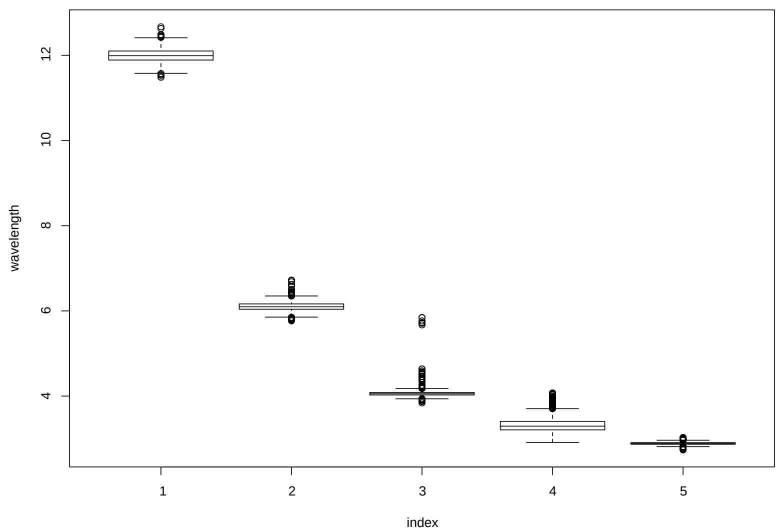

Figure 6: Boxplots of samples for wavelengths corresponding to the 5 largest roots ordered by wavelength.

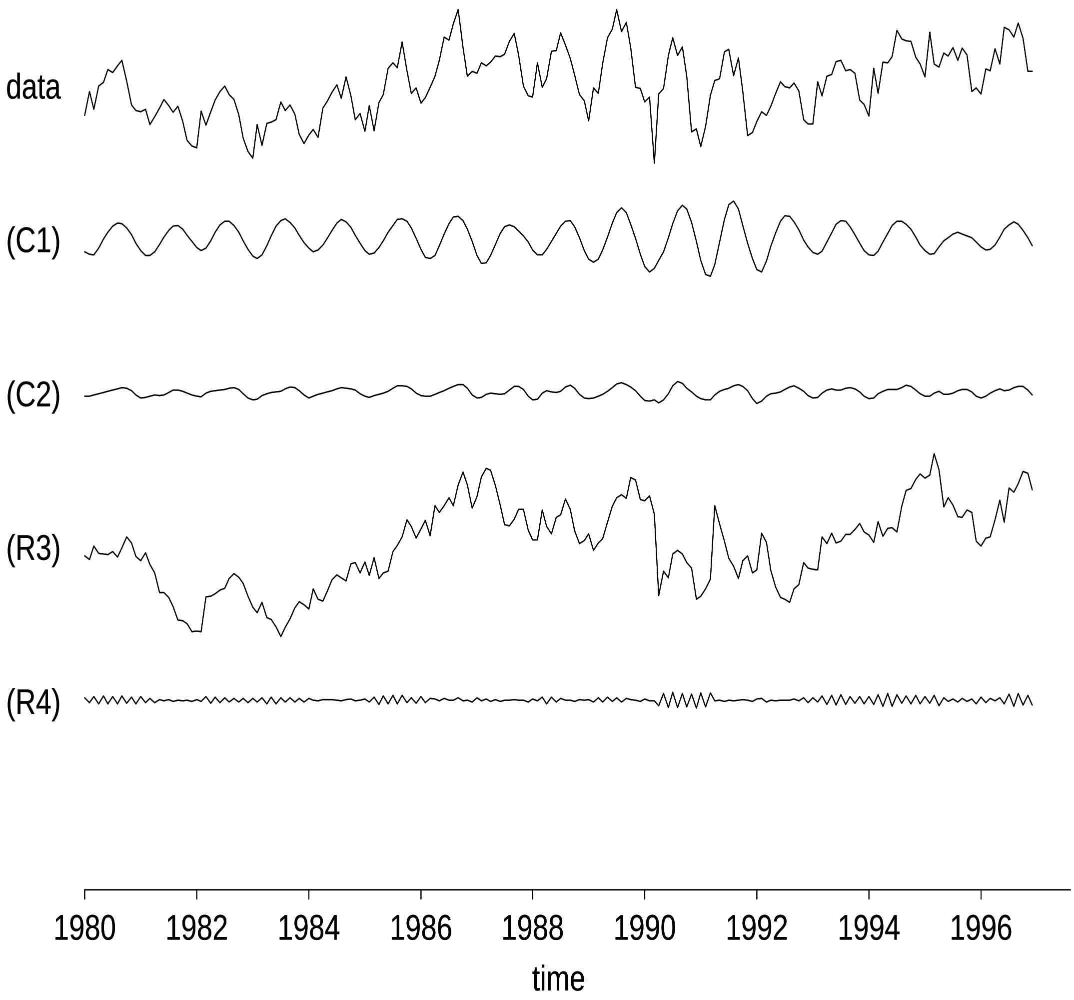

Figure 7: Data and posterior means for two latent components corresponding to complex roots and two latent components corresponding to real roots. Complex components (labeled C1 and C2) are the two largest when ordered by wavelength and real components (labeled R3 and R4) are the corresponding to the maximal and minimal roots.

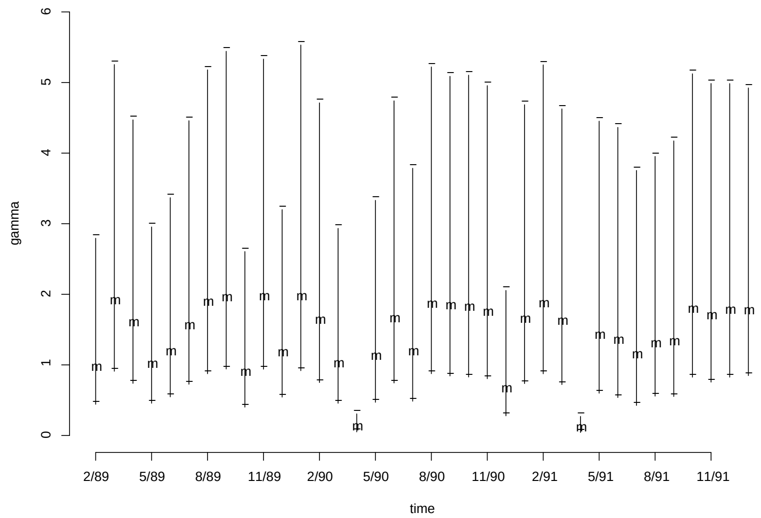

Figure 8: 95% posterior intervals for some of the parameters γ_t. "m" represents the posterior mean.

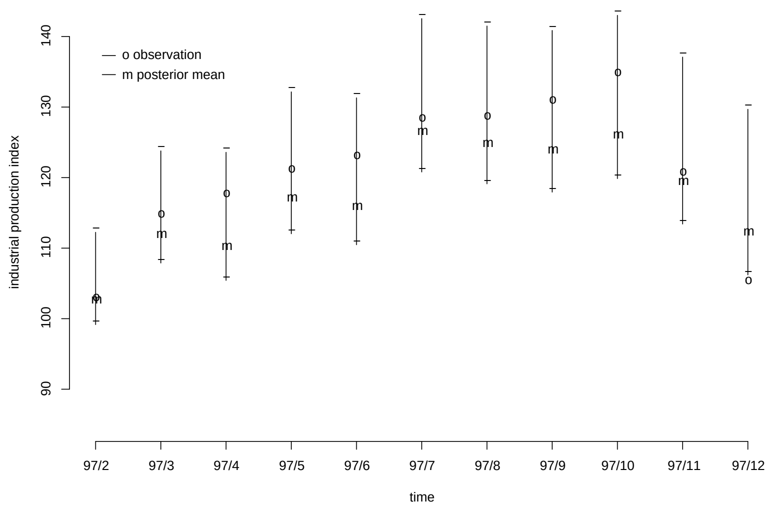

Figure 9: 95% predictive intervals based on AR model with priors on structure components including posterior means and actual observations for 97/3-97/12.

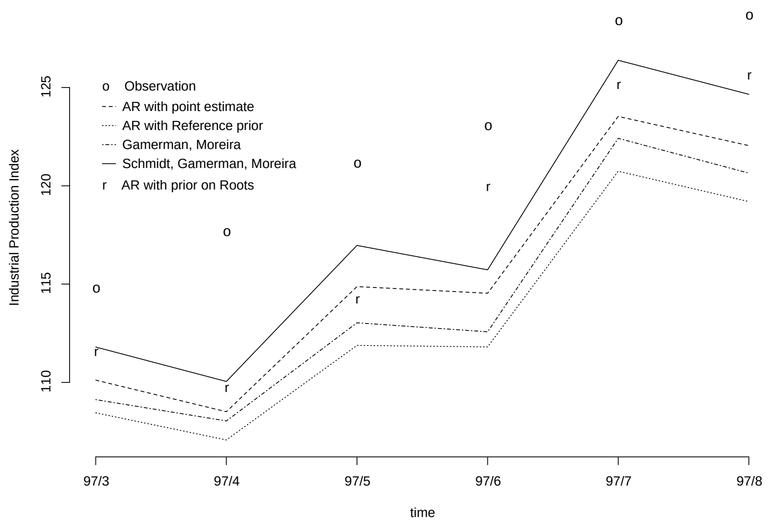

Figure 10: Observations and forecasts for 6 months, 97/3-97/8, corresponding to five time series models.

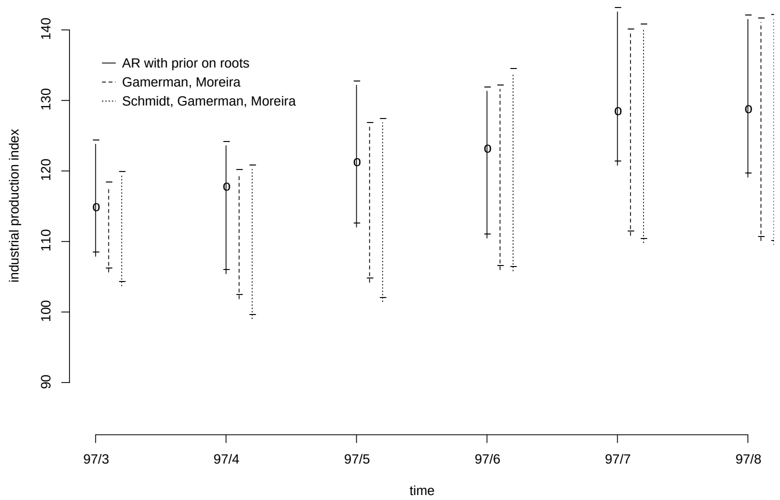

Figure 11: 95% predictive intervals for the period 97/3-97/8 based on AR models and dynamic linear models.