Bayesian analysis of autoregressive panel data model: application in genetic evaluation of beef cattle

Fabyano Fonseca e Silva[1,3]*; Thelma Sáfadi[2]; Joel Augusto Muniz[2]; Guilherme Jordão Magalhães Rosa[3]; Luiz Henrique de Aquino[2]; Gerson Barreto Mourão[4]; Carlos Henrique Osório Silva[1]

[1]*UFV – Depto. de Estatística, 36570-000 – Viçosa, MG – Brasil.*
[2]*UFLA – Depto. de Ciências Exatas, 37200-000 – Lavras, MG – Brasil.*
[3]*University of Wisconsin – Animal Science, Madison, WI – USA.*
[4]*USP/ESALQ – Depto. de Zootecnia, C.P. 09 – 13418-900 – Piracicaba, SP – Brasil.*
**Corresponding author <fabyanofonseca@ufv.br>*

ABSTRACT: The animal breeding values forecasting at futures times is a relevant technological innovation in the field of Animal Science, since its enables a previous indication of animals that will be either kept by the producer for breeding purposes or discarded. This study discusses an MCMC Bayesian methodology applied to panel data in a time series context. We consider Bayesian analysis of an autoregressive, AR(*p*), panel data model of order *p*, using an exact likelihood function, comparative analysis of prior distributions and predictive distributions of future observations. The methodology was tested by a simulation study using three priors: hierarchical Multivariate Normal-Inverse Gamma (model 1), independent Multivariate Student's t – Inverse Gamma (model 2) and Jeffrey's (model 3). Comparisons by Pseudo-Bayes Factor favored model 2. The proposed methodology was applied to longitudinal data relative to Expected Progeny Difference (EPD) of beef cattle sires. The forecast efficiency was around 80%. Regarding the mean width of the EPD interval estimation (95%) in a future time, a great advantage was observed for the proposed Bayesian methodology over usual asymptotic frequentist method.
Key words: MCMC, time series forecasting, prior comparison, predictive distribution

Análise Bayesiana do modelo auto-regressivo para dados em painel: aplicação na avaliação genética de bovinos de corte

RESUMO: A previsão dos valores genéticos de animais em tempos futuros constitui importante inovação tecnológica para a área de Zootecnia, uma vez que possibilita planejar com antecedência o descarte ou a manutenção de animais no rebanho. No presente estudo considerou-se uma análise Bayesiana de modelos auto-regressivos de ordem *p*, AR(*p*), para dados em painel, de forma a utilizar a função de verossimilhança exata, a análise de comparação de distribuições a priori e a obtenção de distribuições preditivas de dados futuros. A metodologia utilizada foi testada mediante um estudo de simulação usando a priori hierárquica Normal multivariada-Gama inversa (modelo 1), a priori independente t-Student Gama inversa (modelo 2) e a priori de Jeffreys (modelo 3). As comparações entre os modelos, realizadas por meio do Pseudo-Fator de Bayes, indicaram uma superioridade do modelo 2 em relação aos demais. Realizou-se uma aplicação em resultados reais referentes as DEP de touros da raça Nelore, sendo que, em média, a eficiência de previsão dos valores de DEP para um ano futuro foi próxima de 80%. Constatou-se considerável vantagem da metodologia proposta em relação a metodologia frequentista usual, uma vez que a implitude dos intervalos de credibilidade de 95% foram muito menores que aquelas apresentadas pelos intervalos de confiança assintóticos.
Palavras-chave: MCMC, previsão em séries temporais, comparação de prioris, distribuição preditiva

Introduction

The advantage of simultaneously modeling several time series, also called panel data analysis, is the possibility of generating more accurate predictions for individual outcomes by pooling the data rather than generating predictions of individual outcomes using the data on the individual series only. The pooling takes place because the parameters of all time series are assumed to arise from the same distribution (Liu and Tiao, 1980). The convenience in the specification of this distribution indicates that the Bayesian procedure has a theoretical advantage over the frequentist approaches, since panel data analysis is directly related to prior information.

A commonly used subjective prior distribution for parameters of auto-regressive models of order p, AR(*p*), is a multivariate normal, but other distributions have also been suggested such as the multivariate Student's t (Barreto and Andrade, 2004) and independent rescaled beta distribution (Liu and Tiao, 1980), in addition to non-informative prior (Sun and Ni, 2003). Thus, the choice of a prior distribution is a relevant topic in the analysis of autoregressive panel data models.

Often, only an approximate likelihood function is attempted in the Bayesian analysis of AR(*p*) models, because the unconditional or exact function does not provide posterior conditional distributions with closed form. The condi-

tionality for p initial observations in AR(p) panel data model, defined by the order p of each series, represents a larger information loss, and although the exact likelihood function increases the analysis complexity, it may still be recommended (Liu and Tiao, 1980). The Bayesian analysis of time series panel data models generates forecasts by combining all the information and sources of uncertainty into a predictive distribution for future values. The interval estimation process inherent to this distribution is theoretically more precise than those obtained from frequentist methods, in which asymptotic approximations are indiscriminately used. To illustrate this fact, we will discuss an application in the field of animal breeding.

Expected Progeny Difference (EPD) is an estimate of the individual's genetic merit for producing future progeny. EPD values are usually reported in the same measurement unit as the trait, and represent the solutions for additive genetic random effects in the Mixed Model Equations proposed by Henderson (1984). These values are published annually by the Sire Summaries, and may change from year to year; thus, several sires with EPD published in several years make up a panel data structure. In practice, to obtain EPD values in a future year for a given sire, the method used is a frequentist approach given by a 95% confidence range (CR), also called the possible chance (PC) interval defined as (Bourdon, 2000): $CR_{95\%} = EPD_a \pm 1.96\sqrt{(1-ACC^2)\sigma_g^2}$, where: EPD_a is EPD calculated in the current year, ACC is the accuracy of EPD_a and σ_g^2 is the estimated genetic variance. The ACC value ranges from 0 to 1, and it informs the reliability of an EPD for a particular animal. In other words, ACC is a measure of the risk associated with the genetic estimate given by: $ACC = \sqrt{1-(PEV/\sigma_g^2)}$, where PEV is the prediction error variance.

In this manuscript we propose a full Bayesian analysis of an autoregressive, AR(p), panel data model, which can be used to provide narrow estimate intervals for EPD values in one future time. The methodology considers an exact likelihood function, comparative analysis of prior distributions and predictive distributions of future observations. This methodology was evaluated by a simulation study using three priors: hierarchical Multivariate Normal-Inverse Gamma (model 1), independent Multivariate Student's t – Inverse Gamma (model 2) and Jeffrey's (model 3). Model comparisons were performed using the Pseudo-Bayes Factor. An application was performed on real data from Nellore sires Expected Progeny Difference (EPD), observed during a six year period (2003-2008).

Material and Methods

Let y_{it} denote m time series realizations, $i = 1,2,...,m$, with time index given by $t = 1,2,...,n_i$. An autoregressive AR(p) panel data model can be described as (Liu and Tiao, 1980):

$y_{it} = \phi_{i1}y_{i(t-1)} + \phi_{i2}y_{i(t-2)} + ... + \phi_{mp}y_{m(t-p)} + e_{it}$, or

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

where: y_{it} is the actual value of a stochastic process; $y_{i(t-1)}$, $y_{i(t-2)}$, ... , $y_{i(t-p)}$ represent values assumed in the past, ϕ_{i1}, ϕ_{i2}, ..., ϕ_{ip} are autoregressive coefficients for each individual; and e_{it} is a non-observable error term, assumed as

$$e_{it} \overset{iid}{\sim} N(0, \sigma_e^2).$$

The exact likelihood function for model (1) is given by:

$$L(\boldsymbol{Y} \mid \boldsymbol{\Phi}, \sigma_e^2) \propto \Psi(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}_p)\sigma_e^{2^{-\left(\frac{m(n-p)}{2}\right)}}$$

$$\exp\left\{-\frac{1}{2\sigma_e^2}\sum_{i=1}^{m}\sum_{t=p+1}^{n}(y_{it} - \sum_{j=1}^{p}\phi_{ij}y_{i(t-j)})^2\right\}, \quad (2)$$

where:

$$\Psi(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}_p) = \sigma_e^{2^{-\left(\frac{mp}{2}\right)}}\left|\boldsymbol{V}_p\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma_e^2}\boldsymbol{Y}_p'\boldsymbol{V}_p\boldsymbol{Y}_p\right\}, \text{ and } (3)$$

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

The matrix $\boldsymbol{V}_p$ is obtained by the Yule-Walker equation (Box et al., 1994). Here, we generalized this matrix for panel data using the block diagonal structure, which is illustrated for the AR(1) and AR(2) autoregressive models, respectively:

$$\boldsymbol{V}_1 = \begin{bmatrix} 1-\phi_{11}^2 & 0 & 0 & 0 \\ 0 & 1-\phi_{21}^2 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1-\phi_{m1}^2 \end{bmatrix}_{mp \times mp}$$

$$\boldsymbol{V}_2 = \begin{bmatrix} 1-\phi_{12}^2 & -\phi_{11}(1+\phi_{12}) & 0 & 0 & 0 & 0 & 0 \\ -\phi_{11}(1+\phi_{12}) & 1-\phi_{12}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\phi_{22}^2 & -\phi_{21}(1+\phi_{22}) & 0 & 0 & 0 \\ 0 & 0 & -\phi_{21}(1+\phi_{22}) & 1-\phi_{22}^2 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1-\phi_{m2}^2 & -\phi_{m1}(1+\phi_{m2}) \\ 0 & 0 & 0 & 0 & 0 & -\phi_{m1}(1+\phi_{m2}) & 1-\phi_{m2}^2 \end{bmatrix}_{mp \times mp}$$

The likelihood function (2) can be written in matrix notation as:

$$L(\boldsymbol{Y} \mid \boldsymbol{\Phi}, \sigma_e^2) \propto \Psi(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}_p)\sigma_e^{2^{-\left(\frac{m(n-p)}{2}\right)}}$$

$$\exp\left\{-\frac{1}{2\sigma_e^2}(\boldsymbol{Y}_1 - \boldsymbol{X}\boldsymbol{\Phi})'(\boldsymbol{Y}_1 - \boldsymbol{X}\boldsymbol{\Phi})\right\} \quad (4)$$

where:

$Y_1 = [y_{1p+1}, y_{1p+2}, ..., y_{1n}, y_{2p+1}, y_{2p+2}, ..., y_{2n}, ..., y_{mp+1}, y_{mp+2}, ..., y_{mn}]'_{m(n-p) \times 1}$,

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 & 0 & 0 & 0 \\ 0 & \boldsymbol{X}_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \boldsymbol{X}_m \end{bmatrix}_{m(n-p) \times mp}, \text{ with } \boldsymbol{X}_i = \begin{bmatrix} y_{ip} & \cdots & y_{i1} \\ y_{ip+1} & \cdots & y_{i2} \\ \vdots & \vdots & \vdots \\ y_{in-1} & \cdots & y_{in-p} \end{bmatrix}_{(n-p) \times p}$$

and,

$\Phi = [\phi_{11}, \phi_{12}, ..., \phi_{1p}, \phi_{21}, \phi_{22}, ..., \phi_{2p}, ..., \phi_{m1}, \phi_{m2}, ..., \phi_{mp}]'_{mp \times 1}$, $\Phi \in \Re^{mp}$.

An important issue with the Bayesian implementation of the AR(p) model refers to the prior specification, because the use of an inappropriate prior may lead to elevated fluctuations in the parameter estimates (Kass and Raftery, 1995). The hierarchical multivariate Normal - Inverse Gamma prior (model 1), independent multivariate Student's t - Inverse Gamma prior (model 2), and the Jeffrey's prior (model 3).

In the case of a hierarchical multivariate Normal- Inverse Gamma prior, we have:

$P(\boldsymbol{\Phi}, \sigma_e^2) = P(\boldsymbol{\Phi} \mid \sigma_e^2) P(\sigma_e^2)$, with $\boldsymbol{\Phi} \mid \sigma_e^2 \sim N(\boldsymbol{\mu}, \sigma_e^{-2}\boldsymbol{P})$ and $\sigma_e^2 \sim IG(\alpha, \beta)$, such that:

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

and $P(\sigma_e^2) \propto (\sigma_e^2)^{-(\alpha+1)} \exp\left\{-\frac{\beta}{\sigma_e^2}\right\}$, leading to the following joint prior distribution:

$$P(\boldsymbol{\Phi}, \sigma_e^2) \propto \sigma_e^{2^{-(\frac{mp+2\alpha}{2}+1)}} \exp\left\{-\frac{1}{2\sigma_e^2}\left[2\beta + (\boldsymbol{\Phi}-\boldsymbol{\mu})' \boldsymbol{P}^{-1}(\boldsymbol{\Phi}-\boldsymbol{\mu})\right]\right\}. \quad (5)$$

The independent multivariate Student t – Inverse Gamma prior is givem by: $P(\boldsymbol{\Phi}, \sigma_e^2) = P(\boldsymbol{\Phi}) P(\sigma_e^2)$, in which $\Phi \sim$*Mult. Student t* (μ, P), with v degrees of freedom, and $\sigma_e^2 \sim IG(\alpha, \beta)$, such that:

$P(\boldsymbol{\Phi}) \propto \left[1 + (\boldsymbol{\Phi}-\boldsymbol{\mu})' \boldsymbol{P}^{-1}(\boldsymbol{\Phi}-\boldsymbol{\mu})\right]^{-\frac{v+mp}{2}}$. The resulting joint prior distribution is then:

$$P(\boldsymbol{\Phi}, \sigma_e^2) \propto \left[1 + (\boldsymbol{\Phi}-\boldsymbol{\mu})' \boldsymbol{P}^{-1}(\boldsymbol{\Phi}-\boldsymbol{\mu})\right]^{-\frac{v+mp}{2}} (\sigma_e^2)^{-(\alpha+1)} \exp\left\{-\frac{\beta}{\sigma_e^2}\right\}. \quad (6)$$

The components $\boldsymbol{\mu}$, $\boldsymbol{P}$, α and β in the expressions 5 an 6 are hyperparameters, whose values must be specified. An alternative for prior specification is the use of a non-informative prior, which represents the lack of knowledge when a certain parametrical family is chosen. Jeffrey's prior approach to autoregressive model used here was presented by Broemiling and Cook (1993), which is given by: $P(\boldsymbol{\Phi}, \sigma_e^2) \propto \frac{1}{\sigma_e^2}$.

Under the Bayesian approach, the prior beliefs about parameters are updated with the information from the data to produce the posterior distribution of the parameters. Thus, it summarizes the current state of knowledge about all the uncertain quantities in a model (Gelman et al., 2003). In this study, we used the same sample information, that is, the same exact likelihood function, combined with each alternative prior distribution, in order to obtain three different joint posterior distributions. Using the Bayes Theorem, $P(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}) \propto L(\boldsymbol{Y} \mid \boldsymbol{\Phi}, \sigma_e^2) P(\boldsymbol{\Phi}, \sigma_e^2)$, such that the resulting joint posterior distribution for each prior specification,is given by:

Model 1:

$$P(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}) \propto \Psi(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}_p) \sigma_e^{2^{-\left(\frac{mn+2\alpha}{2}+1\right)}} \exp\left\{-\frac{1}{2\sigma_e^2}[2D + (\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}}_{\boldsymbol{B}})' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}}_{\boldsymbol{B}})]\right\}, \quad (7)$$

where:

$$D = \beta + \frac{(\boldsymbol{Y}_1' \boldsymbol{Y}_1 + \boldsymbol{\mu}' \boldsymbol{P}^{-1}\boldsymbol{\mu}) - (\boldsymbol{X}' \boldsymbol{Y}_1 + \boldsymbol{P}^{-1}\boldsymbol{\mu})' (\boldsymbol{X}' \boldsymbol{X} + \boldsymbol{P}^{-1})^{-1} (\boldsymbol{X}' \boldsymbol{Y}_1 + \boldsymbol{P}^{-1}\boldsymbol{\mu})}{2},$$

$\sum = X'X + P^{-1}$ and $\hat{\boldsymbol{\Phi}}_{\boldsymbol{B}} = (\boldsymbol{X}' \boldsymbol{X} + \boldsymbol{P}^{-1})^{-1} (\boldsymbol{X}' \boldsymbol{Y}_1 + \boldsymbol{P}^{-1}\boldsymbol{\mu})$.

Model 2:

$$P(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}) \propto \Psi(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}_p) \sigma_e^{2^{-\left[\frac{m(n-p)+2\alpha}{2}+1\right]}} \times \left[1 + (\boldsymbol{\Phi}-\boldsymbol{\mu})' \boldsymbol{P}^{-1}(\boldsymbol{\Phi}-\boldsymbol{\mu})\right]^{-\left(\frac{v+p}{2}\right)}$$
$$\times \exp\left\{-\frac{1}{\sigma_e^2}\left[\left(\frac{(\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}})' (\boldsymbol{X}' \boldsymbol{X})(\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}}) + (\boldsymbol{Y}_1 - \hat{\boldsymbol{Y}}_1)' (\boldsymbol{Y}_1 - \hat{\boldsymbol{Y}}_1)}{2}\right) + \beta\right]\right\}, \quad (8)$$

where:

$\hat{\boldsymbol{\Phi}} = (\boldsymbol{X}' \boldsymbol{X})^{-1} (\boldsymbol{X}' \boldsymbol{Y}_1)$ and $\hat{\boldsymbol{Y}}_1 = \boldsymbol{X}\hat{\boldsymbol{\Phi}} = \boldsymbol{X}(\boldsymbol{X}' \boldsymbol{X})^{-1} (\boldsymbol{X}' \boldsymbol{Y}_1)$.

Model 3:

$$P(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}) \propto \sigma_e^{2^{-\left(\frac{mn}{2}+1\right)}} \left|\boldsymbol{V}_p\right|^{-1/2} \exp\left\{-\frac{1}{2\sigma_e^2}\left[\boldsymbol{Y}_p' \boldsymbol{V}_p^{-1} \boldsymbol{Y}_p + \right.\right.$$
$$+ (\boldsymbol{\Phi} - (\boldsymbol{X}' \boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{Y}_1)' (\boldsymbol{X}' \boldsymbol{X})(\boldsymbol{\Phi} - (\boldsymbol{X}' \boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{Y}_1) +$$
$$\left.\left. + (\boldsymbol{X}' \boldsymbol{Y}_1)' (\boldsymbol{X}' \boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{Y}_1 + \boldsymbol{Y}_1' \boldsymbol{Y}_1\right]\right\}. \quad (9)$$

In a Bayesian analysis, the marginal posterior distributions, which contain all relevant information about the unknown parameters, are obtained by the multidimensional integration of the joint posterior distribution. These integrals are usually impossible to evaluate analytically, but Markov chain Monte Carlo (MCMC) simulation, such as the Gibbs sampler and Metropolis-Hastings, can be used instead (Gelman et al, 2003). The full conditional posterior distributions for $\boldsymbol{\Phi}$ and σ_e^2, are necessary in order to apply MCMC. These distributions can be represented as follows:

Model 1:

$$\boldsymbol{\Phi} \mid \sigma_e^2, \boldsymbol{Y} \sim \Psi(\boldsymbol{\Phi}, \sigma_e^2 \mid \boldsymbol{Y}_p) N(\hat{\boldsymbol{\Phi}}_{\boldsymbol{B}}, \sigma_e^{-2} \boldsymbol{\Sigma}^{-1}) \quad (10)$$

$$\sigma_e^2 \mid \boldsymbol{\Phi}, \boldsymbol{Y} \sim Inv.\ Gamma\left(\frac{mp + mn + 2\alpha}{2}, \frac{1}{2}(\boldsymbol{Y}_p' \boldsymbol{V}_p \boldsymbol{Y}_p) + \right.$$

$$+D+\frac{1}{2}(\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}}_{\boldsymbol{B}})'\boldsymbol{\Sigma}(\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}}_{\boldsymbol{B}})) \quad (11)$$

Model 2:

$$\boldsymbol{\Phi}\,|\,\sigma_e^2,\boldsymbol{Y}\sim\Psi(\boldsymbol{\Phi},\sigma_e^2\,|\,\boldsymbol{Y}_p)\times N(\hat{\boldsymbol{\Phi}},(\boldsymbol{X}'\boldsymbol{X})^{-1})$$

$$\times mult.\ t-Student(\boldsymbol{\mu},\boldsymbol{P^{-1}}) \quad (12)$$

$$\sigma_e^2\,|\,\boldsymbol{\Phi},\boldsymbol{Y}\sim Inv.\ Gamma\left(\frac{mn+2\alpha}{2},\frac{1}{2}(\boldsymbol{Y}_p{}'\boldsymbol{V}_p\boldsymbol{Y}_p)+\beta+\cdot\right.$$

$$\left.+\frac{(\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}})'(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{\Phi}-\hat{\boldsymbol{\Phi}})+(\boldsymbol{Y}_1-\hat{\boldsymbol{Y}}_1)'(\boldsymbol{Y}_1-\hat{\boldsymbol{Y}}_1)}{2}\right) \quad (13)$$

Model 3:

$$\boldsymbol{\Phi}\,|\,\sigma_e^2,\boldsymbol{Y}\sim N(0,\boldsymbol{V}_p^{-1})\times\sigma_e^{2^{-\left(\frac{mn}{2}+1\right)}}\exp\left\{-\frac{1}{2\sigma_e^2}\left[(\boldsymbol{\Phi}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\right.\right.$$

$$\left.\left.\boldsymbol{X}'\boldsymbol{Y}_1)'(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{\Phi}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_1)\right]\right\} \quad (14)$$

$$\sigma_e^2\,|\,\boldsymbol{\Phi},\boldsymbol{Y}\sim Inv.\ Gamma\left(\frac{mn}{2},\frac{1}{2}\left[\boldsymbol{Y}_p{}'\boldsymbol{V}_p\boldsymbol{Y}_p+\right.\right.$$

$$+(\boldsymbol{\Phi}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_1)'(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{\Phi}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_1)\Big)+$$

$$\left.\left.+(\boldsymbol{X}'\boldsymbol{Y}_1)'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_1+\boldsymbol{Y}_1{}'\boldsymbol{Y}_1\right]\right). \quad (15)$$

The conditional posterior distributions of σ_e^2 have a closed form for all considered priors, which is an Inverse Gamma probability density distribution. Therefore, the Gibbs Sampler algorithm can be used to generate samples of the σ_e^2 marginal posterior distribution. For the parameter *F*, however, a closed form can not be obtained, and so we used the Metroplis-Hastings algorithm.

For each prior, single chains with starting values obtained by maximum likelihood estimation were run. Although to use multiple chains is more reliable, avoiding that a single sequence is stuck in some unrepresentative small region of the parameter space, this method is not always recommended. According to Kass et al. (1998), if the convergence is very slow, as in the present study, sometimes runing just one chain can be better than several chains for correspondingly less time.

After several trials, by visual inspection of the graph, the chain length was set to 50,000 and the burn-in period 20,000 iterations, higher than the minimum burn-in required according to the Raftery and Lewis (1992) criteria. Sampling interval (thinning) was set to 3 iterations, so that a total of 10,000 samples were kept for posterior analyses. Convergence was tested for each chain separately using the Geweke (1992) criteria.

The Gibbs Sampler and Metropolis-Hastings algorithms were implemented in the R software (R Development Core Team, 2008). The *mvtnorm* package was used to generate random numbers of multivariate Normal and Student's t distributions, and the *rinvgamma* function (MCMCpack package) for the inverse Gamma distribution. The cited convergence criteria were implemented via the *BOA* (Bayesian Output Analysis) package (Smith, 2007).

In Bayesian forecasting, the prediction of future observations are obtained by a posterior predictive distribution, which is given by a conditional distribution of future data, conditioned on past data. In the present panel data analysis, for a specific individual *i*, the distribution for a future observation is given by:

$$P(\boldsymbol{Y}_{i(n+1)}\,|\,\boldsymbol{Y}_i)\propto\int_{\sigma_e^2}\int_{\boldsymbol{\phi}_i}P(\boldsymbol{Y}_{i(n+1)},\boldsymbol{\phi}_i,\sigma_e^2\,|\,\boldsymbol{Y}_i)\ d_{\boldsymbol{\phi}_i}d_{\sigma_e^2}\text{, where:} \quad (16)$$

$$P(\boldsymbol{Y}_{i(n+1)},\boldsymbol{\phi}_i,\sigma_e^2\,|\,\boldsymbol{Y}_i)\propto L(\boldsymbol{Y}_{i(n+1)}\,|\,\boldsymbol{\phi}_i,\sigma_e^2,\boldsymbol{Y}_i)P(\boldsymbol{\phi}_i,\sigma_e^2\,|\,\boldsymbol{Y}_i), \quad (17)$$

The term is obtained by the following model:

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

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

$e_{i(n+1)}\sim N(0,\sigma_e^2)$, such that:

$$L(\boldsymbol{Y}_{i(n+1)}\,|\,\boldsymbol{\phi}_i,\sigma_e^2,\boldsymbol{Y}_i)\propto\left(\sigma_e^2\right)^{-\frac{1}{2}}exp\left\{-\frac{1}{2\sigma_e^2}\left[Y_{i(n+1)}-\sum_{j=1}^{p}\phi_{ij}Y_{i(n+1-j)}\right]^2\right\}. \quad (18)$$

Generalizing the expression (17) for *m* individuals, we have:

$$L(\boldsymbol{Y}_{(n+1)}\,|\,\boldsymbol{\Phi},\sigma_e^2,\boldsymbol{Y})\propto\prod_{i=1}^{m}L(\boldsymbol{Y}_{i(n+1)}\,|\,\boldsymbol{\phi}_i,\sigma_e^2,\boldsymbol{Y}_i)$$

$$\propto\left(\sigma_e^2\right)^{-\frac{m}{2}}exp\left\{-\frac{1}{2\sigma_e^2}\sum_{i=1}^{m}\left[Y_{i(n+1)}-\sum_{j=1}^{p}\phi_{ij}Y_{i(n+1-j)}\right]^2\right\}. \quad (19)$$

Using matrix notation, the expression (19) can be expressed as:

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

$$\left.\left.'\left(\boldsymbol{Y}_{(n+1)}-\boldsymbol{X}\boldsymbol{\Phi}\right)\right]\right\}\text{ where:} \quad (20)$$

$$\boldsymbol{Y}_{(n+1)}=\begin{bmatrix}y_{1(n+1)}\\y_{2(n+1)}\\\vdots\\y_{m(n+1)}\end{bmatrix}_{\boldsymbol{m}\times 1},\quad \boldsymbol{X}=\begin{bmatrix}\boldsymbol{X}_1&0&0&0\\0&\boldsymbol{X}_2&0&0\\0&0&\ddots&0\\0&0&0&\boldsymbol{X}_m\end{bmatrix}_{\boldsymbol{m}\times\boldsymbol{mp}}\text{ and}$$

$$\boldsymbol{X}_i=\left[y_{in}\,y_{i(n-1)}\ldots y_{i(n+1-p)}\right]_{1\times\boldsymbol{p}}.$$

Therefore,

$P(\boldsymbol{Y}_{(n+1)},\boldsymbol{\Phi},\sigma_e^2 \mid \boldsymbol{Y}) \propto L(\boldsymbol{Y}_{(n+1)} \mid \boldsymbol{\Phi},\sigma_e^2,\boldsymbol{Y})P(\boldsymbol{\Phi},\sigma_e^2 \mid \boldsymbol{Y})$, where $P(\boldsymbol{\Phi},\sigma_e^2 \mid \boldsymbol{Y})$ is the joint posterior distribution previously presented. Then, by the generalization of expression (16), we have that:

$$P(\boldsymbol{Y}_{(n+1)} \mid \boldsymbol{Y}) \propto \int_{\sigma_e^2}\int_{\boldsymbol{\Phi}} \left(\sigma_e^2\right)^{-\frac{m}{2}} exp\left\{-\frac{1}{2\sigma_e^2}\left[\left(\boldsymbol{Y}_{(n+1)}-\boldsymbol{X\Phi}\right)'\left(\boldsymbol{Y}_{(n+1)}-\boldsymbol{X\Phi}\right)\right]\right\} P(\boldsymbol{\Phi},\sigma_e^2 \mid \boldsymbol{Y}) d_{\boldsymbol{\Phi}} d_{\sigma_e^2}. \quad (21)$$

This integral does not present an analytical solution, but according to Heckman and Leamer (2001) it is possible to obtain an approximation via MCMC algorithm with the distribution: $\boldsymbol{Y}_{(n+1)}^{(q)} \mid \boldsymbol{Y} \sim N\left(\boldsymbol{X\Phi}^{(q)},\sigma_e^{2(q)}\boldsymbol{I}\right)$, $\boldsymbol{I}_{mp\ x\ mp}$, where $\boldsymbol{I}$ is an identity matrix. The set of values generated by this multivariate normal distribution, at each q MCMC iteration step, constitutes a sample of future observations from the posterior predictive distribution. Then, the point estimate of this value for future observations, is given by the mean of this sample $\hat{P}(\boldsymbol{Y}_{(n+1)} \mid \boldsymbol{Y})$.

The comparison between priors used the Pseudo-Bayes Factor (PBF) (Gelfand, 1996), which is defined as the ratio of $\hat{P}(\boldsymbol{Y}_{(n+k)} \mid \boldsymbol{Y})$ produced by models M_z and M_w:

$$PBF_{zw} = \ln\left(\frac{\prod_{t=n}^{k}\hat{P}(\boldsymbol{Y}_{(n+k)} \mid \boldsymbol{Y},M_z)}{\prod_{t=n}^{k}\hat{P}(\boldsymbol{Y}_{(n+k)} \mid \boldsymbol{Y},M_w)}\right),$$

where k is the number of future time points to be predicted. If $PBF_{zw} > 0$ then M_z is selected, otherwise M_w is selected. In the present study it was assumed that k = 1, however, as the panel data structure demands a generalization for each individual series i, we have:

$$PBF_{zw} = \ln\left(\frac{\prod_{i=1}^{m}\hat{P}(Y_{i(n+1)} \mid \boldsymbol{Y},M_z)}{\prod_{i=1}^{m}\hat{P}(Y_{i(n+1)} \mid \boldsymbol{Y},M_w)}\right). \quad (22)$$

A simulation study was conducted to evaluate the proposed methodology. The AR(2) model was used because it is the most simple multi-parametric autoregressive approach. It is given by:

$$Y_{it} = \phi_{i1}Y_{i(t-1)} + \phi_{i2}Y_{i(t-2)} + e_{it};\ \text{with } i = 1,2,...,10,\ \text{and } t = 1,2,...,12,$$

where:

$$\boldsymbol{\phi}_i = [\phi_{i1},\phi_{i2}]'\ \text{with}\ \phi_{i1}+\phi_{i2}<1,\ \phi_{i2}-\phi_{i1}<1\ \text{and}\ -1<\phi_{i2}<1,$$

such that the series is stationary.

The parameter values, ϕ_{i1} and ϕ_{i2}, were generated by multivariate normal (model 1) and multivariate Student's t distributions (model 2), so it is possible to assume these distributions as the true prior distributions in the efficiency analysis of the Pseudo-Bayes Factor. The distributions used to generate the parameter values for models 1 and 2 are given, respectively, by:

$$\boldsymbol{\phi}_i \sim N\left(\begin{bmatrix}0.5\\-0.5\end{bmatrix},\begin{bmatrix}0.025 & 0\\0 & 0.010\end{bmatrix}\right)\text{and}$$

$$\boldsymbol{\phi}_i \sim Mult.\ Student\ t\left(\begin{bmatrix}0.5\\-0.5\end{bmatrix},\begin{bmatrix}0.025 & 0\\0 & 0.010\end{bmatrix},gl = n-1\right).$$

The residual distribution was Normal, $e_{it} \sim N(0,\sigma_e^2)$, were σ_e^2 was drawn from an Inverse Gamma distribution, $\sigma_e^2 \sim IG(3,2)$.

This simulation study also provides an alternative way to evaluate the autoregressive panel data model predictive ability, which is assumed by predicting the last observation, $\hat{Y}_{il2}$, which was excluded from the analysis for parameter estimation.

The proposed methodology was applied to a real data set obtained from the animal breeding field, which refers to Expected Progeny Difference (EPD) for Nellore cattle. The data are from an annual genetic analysis of 117 sires and were provided by the Animal Breeding Group of the Universidade de São Paulo, in Pirassununga, state of São Paulo, Brazil. The EPD values for the weight gain from weaning (205 days of age) to 550 days of age were recorded between 2003 to 2008.

These EPD values were calculated with different accuracy levels, since these values are indicators of the reliability of the published genetic estimates for each animal. If all animals had the same accuracy level, characteristics such mean weight, age and herd, among others, could be used in order to classify latent sire groups. We split the animals into three groups according to their 2007 accuracy levels. These groups were classified as low (0 to 40%), median (41 to 60%) and high (above 60%) accuracy, which contained 31, 63 and 23 sires, respectively. These groups were constructed to ensure the homogeneity within each classification widely required for panel data analysis.

The order of the autoregressive process was identified with autocorrelogram plots, which showed that only first and second order models make sense. The AR(2) model was chosen, since this model includes the AR(1) structure.

Predictive ability was evaluated using 2008 EPD values, which were omitted from the estimated sample. To compare the forecasting precision of EPD values in one period ahead with the conventional method, we use a 95% confidence range (CR), and 95% Bayesian credible interval, which is based on 2.5 an 97.5 percentiles of the posterior predictive distribution, the mean width intervals were calculated. The point estimates from different methods were compared by through the root mean square error (RMSE).

Results and Discussion

The Bayesian methodology was efficient (Tables 1 and 2), with 95% credibility intervals containing 80% and 90% of the ϕ_1 parametric values, and 90% and 100% of the ϕ_2

parametric values, respectively for models 1 and 2. Furthermore, the estimates from model 2 also presented best agreement with the true values, since the RMSE for this model, regarding respectively ϕ_1 (0.0441) and ϕ_2 (0.0392), were smaller those obtained from model 1, respectively 0.0678 and 0.1346 for ϕ_1 and ϕ_2. These same considerations apply to residual variance estimates (Table 3).

The last three columns in Tables 1, 2 and 3 indicate that all chains give very similar results. The convergence was seemingly achieved, with the z-scores of the Geweke tests never higher than 1.96. The burn-in period values used were much higher than the minimum recommended by the procedure of Raftery and Lewis. For model 1, 60% of the credibility intervals contained the true values, and 80% for model 2 (Table 4). Thus, the joint evaluation of the two models produced an efficiency of 70%, which is similar to other studies that performed also model predictive ability evaluation. De Alba (1993) simulated independent time series with the AR(4) model and noted an efficiency of 75%, while Hay and Pettitt (2001) obtained 58% in the analysis of twelve pneumonia incidence time series using generalized the AR(1) model for count data. In relation to efficiency of point estimates, the RMSE calculated were very similar for the two models, 0.3541 and 0.3469, respectively for models 1 and 2.

Although model 2 was presented as the best for sample fit (Tables 1, 2 and 3), and the best in predictive ability (Table 4), the models were also compared by the Pseudo-Bayes Factor. The results are presented in Table 5, which show the model 2 superiority in relation to models 1 and 3, even when data were simulated using model 1. This interesting result can be in part explained by Student's t distribution with higher degrees of freedom, which is similar to the normal distribution. In general, the literature has related the Student's t prior quality for parameters of time series autoregressive models, and among these we refer the reader to Barreto and Andrade (2004).

The Pseudo-Bayes Factor values calculated to compare the three models fitted to the EPD data are presented in Table 6. The results are similar to those obtained with the simulated data, indicating a better performance for the independent multivariate Student's t - Inverse Gamma prior. The Pseudo-Bayes Factor magnitude in relation to the hierarchical multivariate Normal-Inverse Gamma is small, and the non-informative Jeffreys' prior had the worse results. This fact is very well discussed by Lambert et al. (2005), who indicate that, even if we do not have well determined information about the parameters, it is very important to use and compare the informative prior with the non-informative, be-

Table 1 – Individual series, parametric value (ϕ_1), posterior mean estimate ($\hat{\phi}_1$), 95% credibility intervals (LL and UL), Geweke Z score and burn-in values.

Hierarchical Multivariate Normal - Inverse Gamma prior (Model 1)						
Series	ϕ_1	$\hat{\phi}_1$	LL*	UL	Z	Burn-in
1	0.24	0.28	0.22	0.33	0.95	2
2	0.31	0.36	0.30	0.42	0.84	2
3	0.59	0.69	0.55	0.82	0.55	6
4	0.36	0.33	0.27	0.38	0.39	2
5	0.34	0.40	0.28	0.53	-0.57	3
6	0.41	0.48	0.39	0.58	-0.58	2
7	0.39	0.36	0.27	0.44	-0.90	2
8	0.60	0.70	0.61**	0.82	1.15	196
9	0.65	0.61	0.54	0.67	0.33	3
10	0.61	0.71	0.63	0.87	1.38	208
Independent multivariate Student's t - Inverse Gamma prior (Model 2)						
Series	ϕ_1	$\hat{\phi}_1$	LL*	UL	Z	Burn-in
1	0.47	0.51	0.32	0.68	0.42	3
2	0.65	0.67	0.59	0.74	0.63	3
3	0.39	0.45	0.41	0.55	1.42	250
4	0.47	0.49	0.35	0.65	0.33	3
5	0.84	0.89	0.78	1.11	1.31	105
6	0.60	0.69	0.59	0.78	0.63	2
7	0.33	0.36	0.24	0.50	0.58	2
8	0.49	0.51	0.33	0.67	0.47	3
9	0.30	0.34	0.24	0.45	0.97	2
10	0.41	0.41	0.34	0.49	0.45	5

*Lower Limit (LL) and Upper Limit (UL); **Miss-left intervals are highlighted in gray.

Table 2 – Individual series, parametric value (ϕ_2), posterior mean estimate ($\hat{\phi}_2$), 95% credibility intervals (LL and UL), Geweke Z score and burn-in values.

Hierarchical Multivariate Normal - Inverse Gamma prior (Model 1)						
Series	ϕ_2	$\hat{\phi}_2$	LL*	UL	Z	Burn-in
1	-0.60	-0.52	-0.62	-0.44	0.85	4
2	-0.57	-0.49	-0.60	-0.39	-0.64	3
3	-0.34	-0.39	-0.46	-0.33	0.35	3
4	-0.68	-0.58	-0.72	-0.43	0.54	3
5	-0.6	-0.51	-0.65	-0.36	1.17	205
6	-0.5	-0.43	-0.56	-0.32	0.14	3
7	-0.52	-0.45	-0.58	-0.33	0.78	2
8	-0.36	-0.71	-0.66	-0.49	-1.05	505
9	-0.64	-0.75	-0.91	-0.57	0.37	5
10	-0.81	-0.75	-0.92	-0.59	0.56	320
Independent multivariate Student's t - Inverse Gamma prior (Model 2)						
Series	ϕ_2	$\hat{\phi}_2$	LL*	UL	Z	Burn-in
1	-0.78	-0.82	-0.88	-0.75	0.14	2
2	-0.58	-0.66	-0.79	-0.53	0.25	2
3	-0.61	-0.61	-0.78	-0.53	0.33	1
4	-0.56	-0.61	-0.72	-0.49	-0.18	2
5	-0.57	-0.6	-0.66	-0.54	0.32	3
6	-0.4	-0.38	-0.55	-0.22	0.08	3
7	-0.6	-0.65	-0.71	-0.59	-0.96	3
8	-0.59	-0.62	-0.69	-0.54	1.25	102
9	-0.31	-0.32	-0.39	-0.24	-0.23	2
10	-0.38	-0.37	-0.42	-0.33	0.38	2

*Lower Limit (LL) and Upper Limit (UL); **Miss-left intervals are highlighted in gray.

Table 3 – Parametric value (σ_e^2), posterior mean estimate ($\hat{\sigma}_e^2$), 95% credibility intervals (LL and UL), Geweke Z score and burn-in values.

Hierarchical Multivariate Normal - Inverse Gamma prior (Model 1)					
σ_e^2	$\hat{\sigma}_e^2$	LL*	UL	Z	Burn-in
1.9905	2.7023	1.9414	4.3202	1.25	306
Independent multivariate Student's t - Inverse Gamma prior (Model 2)					
σ_e^2	$\hat{\sigma}_e^2$	LL*	UL	Z	Burn-in
1.0541	1.6061	0.9599	3.0250	1.33	450

*Lower Limit (LL) and Upper Limit (UL).

Table 4 – Last observation omitted values (Y_{12}), posterior mean estimate ($\hat{Y}_{12}$), 95% credibility intervals (LL and UL).

	Model 1				Model 2			
Series	Y_{12}	$\hat{Y}$	LL*	UL	Y_{12}		LL*	UL
1	0.69	0.43	0.12	0.74	0.60	0.46	0.24	0.68
2	0.28	0.12	-0.13	0.46	-1.58	-1.17	-1.63	-0.91
3	0.42	0.25	0.06	0.44	-0.56	0.02	-0.44	0.48
4	0.66	1.04	0.69**	1.39	1.49	1.92	1.35	2.49
5	1.18	1.36	1.13	1.59	-0.70	-0.92	-1.24	-0.60
6	1.18	1.41	0.89	1.93	0.03	0.14	-0.19	0.47
7	1.43	1.02	0.75	1.29	-0.54	0.06	-0.46	0.51
8	-0.55	-0.04	-0.42	0.26	-0.94	-0.69	-1.00	-0.38
9	1.27	0.97	0.65	1.30	-0.04	0.13	-0.06	0.33
10	0.63	1.25	0.95	1.55	-0.96	-1.19	-1.54	-0.84

*Lower Limit (LL) and Upper Limit (UL); **Miss-left intervals are highlighted in gray.

Table 5 – Model comparisons by Pseudo-Bayes Factor (PBF_{zw}^{*}).

True Model	Criteria
Hierarchical Multivariate Normal - Inverse Gamma prior (Model 1)	$PBF_{12} = \ln\left(\frac{-0.0025}{-0.4321}\right) = -5.1523$
	$PBF_{13} = \ln\left(\frac{-0.0025}{-0.0013}\right) = 0.6539$
	$PBF_{23} = \ln\left(\frac{-0.4321}{-0.0013}\right) = 5.8062$
Independent multivariate Student's t - Inverse Gamma prior (Model 2)	$PBF_{12} = \ln\left(\frac{-0.0000917}{-0.00032}\right) = -1.2497$
	$PBF_{13} = \ln\left(\frac{-0.0000917}{-0.0000222}\right) = 1.4184$
	$PBF_{23} = \ln\left(\frac{-0.00032}{-0.0000222}\right) = 2.6682$

*Indexes z and w, respectively indicate, the numerator and denominator models, model 3 is the non-informative Jeffreys' prior.

Table 6 – Comparison criteria given by Pseudo-Bayes Factor (PBF_{zw}^{*}) for each data structure analyzed.

Data structure	Criteria
Low accuracy	$PBF_{12} = \ln\left(\frac{-0.0541}{-0.0875}\right) = -0.4808$
	$PBF_{23} = \ln\left(\frac{-0.0875}{-0.0035}\right) = 3.2188$
	$PBF_{13} = \ln\left(\frac{-0.0541}{-0.0035}\right) = 2.7380$
Median accuracy	$PBF_{12} = \ln\left(\frac{-0.0902}{-0.1255}\right) = -0.3302$
	$PBF_{23} = \ln\left(\frac{-0.1255}{-0.0011}\right) = 4.7369$
	$PBF_{13} = \ln\left(\frac{-0.0902}{-0.0011}\right) = 4.4067$
High accuracy	$PBF_{12} = \ln\left(\frac{-0.0857}{-0.0987}\right) = -0.1412$
	$PBF_{23} = \ln\left(\frac{-0.0987}{-0.0023}\right) = 3.7591$
	$PBF_{13} = \ln\left(\frac{-0.0857}{-0.0023}\right) = 3.6179$

cause probably the informative one will lead to better posterior results.

In the context of our data, at each new sire annual evaluation the researcher can increase the amount of information coming from the data, which means that the prior information used in this Bayesian analysis can be given by the posterior distribution obtained from the previous year. Thus, after several years, as the sample size increases, the point and interval Bayesian estimates of the parameters, including the EPD for a future year, will be driven more and more by the observed data rather than by the prior. Anyway, due to the great influence of the prior information in the first step of this proposed system, the Pseudo-Bayes Factor was used to objectively choose the prior distribution, and the independent Multivariate Student's t - Inverse Gamma (model 2) was indicated. Concluding, we believe that subjective prior choice can lead to results that are driven not by the data but by prior unconfirmed beliefs and this fact may compromise the reliability of the study findings.

After the choice of the best prior, independent multivariate Student's t – Inverse Gamma, the predictive ability produced by this model is shown in Table 7. It represents the percentage of sires whose credibility intervals contained the EPD true values, corresponding to year 2008. The residual variance of posterior estimates for each accuracy group is also presented in Table 6. The predictive ability of this proposed Bayesian methodology was efficient to forecast future EPD values, since the percentages were consistently high, around 80%.

The mean width of EPD interval estimation (95%) for the usual method (Bourbon, 2000) and proposed Bayesian forecasting are presented in Figure 1. These results demonstrate a great advantage of the Bayesian methodology, which provide narrow intervals for each accuracy group. In relation

Table 7 – Predictive capacity (PC), posterior mean estimate ($\hat{\sigma}_e^2$), 95% credibility intervals (LL and UL), and burn-in values.

Data structure	PC (%)	$\hat{\sigma}_e^2$	LL*	UL	Burn-in
Low accuracy	77.78	84.32	53.52	128.32	562
Median accuracy	83.33	31.00	19.25	47.33	320
High accuracy	85.71	37.99	26.35	58.97	128

to efficiency of point estimates, the Bayesian method also presented more efficiency based on RMSE (0.3022 versus 0.9345).

Conclusions

The proposed full Bayesian framework for analysis of panel data, using an exact likelihood function, comparative analysis of priors and predictive distributions of future observations worked very well in our study, for both, simulated and real data sets, since the predictive ability was nearly 90% for all considered models. Comparisons by the Pseudo-Bayes Factor indicated superiority of the independent Multivariate Student's t - Inverse Gamma prior in relation to the others. The real data analysis indicated the importance of grouping sires by accuracy values, and also showed a forecast efficiency around 80%.

References

Barreto, G.; Andrade, M.G. 2004. Robust Bayesian Approach for AR(p) Models Applied to Streamflow Forecasting. Journal of Applied Statistical Science 12: 269-292.

Bourdon, R.M. 2000. Understanding Animal Breeding. Prentice-Hall, Upper Saddle River, NJ, USA.

Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. 1994. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, CA, USA.

Broemiling, D.L.; Cook, P. 1993. Bayesian estimation of the mean of an autoregressive process. Journal of Applied Statistics 20: 25-38.

De Alba, E. 1993. Constrained forecasting in autoregressive time series models: a Bayesian analysis. International Journal of Forecasting 9: 95-108.

Gelman, A.; Carlin, J.B.; Stern, H.S.; Rubin, D.B. 2003. Bayesian Data Analysis. Chapman and Hall, Boca Raton, FL, USA.

Gelfand, A.E. 1996. Model Determination Using Sampling Based Methods. Chapman and Hall, London, UK.

Geweke, J. 1992. Evaluating the accurary of sampling-based approaches to the calculation of posterior moments. Oxford University Press, Oxford, UK.

Hay, J.L.; Pettitt, A.N. 2001. Bayesian analysis of a time series of counts with covariates: an application to the control of an infectious disease. *Biostatistics* 2: 433-444.

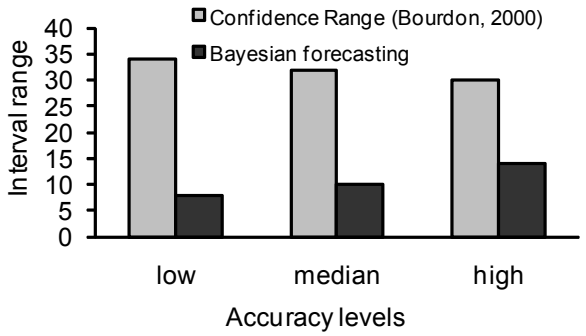

Figure 1 – Mean width for Confidence Range (CR) and Bayesian credibility intervals.

Heckman, J.; Leamer, E. 2001. Handbook of Econometrics. Elsevier Science, Amsterdam, The Netherlands.

Henderson, C.R. 1984. Applications of Linear Models in Animal Breeding. University of Guelph Press, Ontario, Canada.

Kass, R.E.; Carlin, B.P.; Gelman, A.; Neal, R.M. 1998. Markov Chain Monte Carlo in practice: a roundtable discussion. American Statistician 52: 93-100.

Kass, R.E.; Raftery, A.E. 1995. Bayes Factor. Journal of the American Statistical Association 90: 773-795.

Lambert, P.C.; Sutton, A.J.; Burton, P.R.; Abrams, K.R.; Jones, D.R. 2005. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Statistics in Medicine 24: 2401-2428.

Liu, L.M.; Tiao, G.C. 1980. Random coefficient first-order autoregressive model. Journal of Econometrics 13: 305-325.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A.E.; Lewis, S. 1992. How many iterations in the Gibbs sampler? Oxford University Press. Oxford, UK.

Smith, B.J. 2007. boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference. Journal of Statistical Software 21: 1-37.

Sun, D.; Ni, S. 2003. Noninformative priors and frequentist risks of Bayesian estimators of vector-autoregressive models. Journal of Econometrics 115: 159 -197.

Received September 03, 2009
Accepted August 16, 2010