Bootstrap for correcting the mean square error of prediction and smoothed estimates in structural models

T. R. Santos[a]** and G. C. Franco[a]

[a]*Department of Statistics, Universidade Federal de Minas Gerais, Brazil*

February 17, 2014

Abstract

It is well known that the uncertainty in the estimation of parameters produces the underestimation of the prediction mean square error (PMSE). In the state space framework, this problem can affect confidence intervals for smoothed estimates and forecasts, which are generally built by state vector predictors that use estimated model parameters. In order to correct this problem, this paper proposes and compares parametric and nonparametric bootstrap methods based on procedures usually employed to calculate the PMSE. The comparisons are performed through an extensive Monte Carlo study which illustrates, empirically, the bias reduction in the estimation of PMSE using the bootstrap approaches. The finite sample properties of the bootstrap procedures are analyzed in the context of forecasting and smoothing, for Gaussian and non-Gaussian assumptions of the error term. The procedures are also applied to real time series, leading to satisfactory results.

State space models; hyperparameters; MLE; confidence and prediction intervals; parametric and nonparametric bootstrap.

**Corresponding author. Thiago Rezende dos Santos. Address: Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627, Belo Horizonte, MG, 31270-901, Brazil. E-mail: thiagords@ufmg.br

1 Introduction

One of the main objectives of modeling time series is to forecast future values. In classical inference, forecasting is usually performed using a plug-in approach, i.e., replacing the model parameters by their estimators. It is well known that this procedure leads to the underestimation of the prediction mean square error (PMSE), as it does not incorporate the uncertainty due to parameter estimation (see Pfeffermann & Tiller (2005), Ansley & Kohn (1986) and Yamamoto (1976)), resulting in prediction intervals with smaller widths.

In state space models, the underestimation of the PMSE can also influence inference for the state vector predictor, specially for short series. If model parameters are replaced by their estimators, the mean square error (MSE) and confidence intervals for the smoothed estimates of the state vector may be affected. The problem is worsened if the Gaussian distribution is assumed when there is no guarantee that this is the actual distribution of the error terms.

Many authors have proposed different approaches to overcome the problem of underestimating the PMSE. In the state space context, Queenville & Singh (2000) proposed a modification in the methods of Hamilton (1986) and Ansley & Kohn (1986) by incorporating the uncertainty in the estimation of the state vector MSE through the corrected hyperparameter estimator. A similar approach was adopted by Shephard (1993), who wrote the local level model as a normal mixed effects model in order to use the restricted maximum likelihood estimator (RMLE). Tsimikas & Ledolter (1994) presented an alternative way to build the restricted likelihood function, also using mixed effects models.

Another way to incorporate the uncertainty in the estimation of the parameters is through asymptotic sampling of the maximum likelihood estimator (MLE) (Hamilton, 1986; Queenville & Singh, 2000), which may be a poor approximation specially for small samples.

The bootstrap technique can also be employed to incorporate the uncertainty in the estimation of ψ, the hyperparameter vector. Wall & Stoffer (2002) proposed a complex and costly method in terms of computational implementation for the construction of the empirical distribution of the forecast error. Rodriguez & Ruiz (2010) presented a simpler nonparametric bootstrap procedure, compared to the proposal of Wall & Stoffer (2002), and the method is justified by the results of a Monte Carlo simulation in finite samples. They employ a different approach using a quasi-likelihood estimator of ψ, which has some interesting properties, such as consistency (Harvey, 1989). Pfeffermann & Tiller (2005) proposed parametric and nonparametric bootstrap methods for estimating the PMSE of the state vector.

The main objective of this paper is to propose and compare bootstrap methods to reduce the bias in the estimation of the PMSE. To this purpose, nonparametric bootstrap procedures are proposed by adapting some methods existing in the literature to estimate the PMSE of the state vector, such as those of Hamilton (1986) and Ansley & Kohn (1986). Forecast intervals using the nonparametric bootstrap correction of the PMSE are also provided, along with a new proposal of a parametric bootstrap interval. The procedures are compared, through Monte Carlo experiments, to the method of Pfeffermann & Tiller (2005) and to the standard procedures.

The paper is organized as follows. Section 2 presents the structural models and how to calculate forecasts and smoothed estimates of the state vector. Section 3 presents nonparametric bootstrap procedures for forecasts and smoothed state vectors while Section 4 presents the parametric bootstrap interval proposed here. Section 5 provides some simulation studies. Section 6 shows application to real data sets. Finally, Section 7 concludes the work.

2 Forecasting and smoothing in structural models

A univariate time series $\{y_t\}_{t \in Z}$ can be decomposed as the sum of its unobservable components, such as trend, seasonality and error. A procedure which has been widely used to model y_t is the state space model (SSM) (Harvey, 1989; West & Harrison, 1997), which can be written as

$$y_t = \mathbf{z}_t' \boldsymbol{\alpha}_t + \epsilon_t, \quad \epsilon_t \sim N(0, h_t) \tag{1}$$

$$\boldsymbol{\alpha}_t = \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{R}_t \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \mathbf{Q}_t) \tag{2}$$

where $\boldsymbol{\alpha}_t$ is the state vector, $\mathbf{z}_t$, $\mathbf{T}_t$ and $\mathbf{R}_t$ are the system matrices, ϵ_t are uncorrelated errors with variance h_t, $\boldsymbol{\eta}_t$ is a vector of serially uncorrelated errors with covariance matrix given by $\mathbf{Q}_t$ and $\boldsymbol{\eta}_t$ and ϵ_t are independent.

Covariates can be added to both equations to include information such as structural breaks, outliers or external variables.

The likelihood function can be obtained through the one-step-ahead forecast error, $\nu_t = y_t - \widetilde{y}_{t|t-1}$, calculated using the Kalman filter (KF) algorithm (Kalman, 1960), assuming $(y_t|Y_{t-1}) \sim N(\widetilde{y}_{t|t-1}, F_t)$, where $Y_{t-1} = \{y_1, ..., y_{t-1}\}$ and $\widetilde{y}_{t|t-1}$ is the one-step-ahead forecast. For a univariate series of size n, the logarithm of the likelihood function is given by

$$lnL(\boldsymbol{\psi}; Y_n) = ln \prod_{t=1}^{n} p(y_t|Y_{t-1}) = -\frac{n}{2} ln(2\pi) - \frac{1}{2}\sum_{t=1}^{n} ln|F_t| - \frac{1}{2}\sum_{t=1}^{n} \nu_t' F_t^{-1} \nu_t \tag{3}$$

where $\boldsymbol{\psi} = (\psi_1, \psi_2, \ldots, \psi_p)$ is the hyperparameter vector, which in this case are the variances of the errors. Since this is a nonlinear function of the hyperparameters, estimation must be done numerically. In this work, the optimization algorithm BFGS is used (details can be seen in Franco *et al.* (2008)).

- **Smoothing**

The smoothed estimator of the state vector is the inference performed on any particular date t, using the whole sample information, $\{y_1, ..., y_n\}$. In linear Gaussian models, the smoothed estimator of the state vector is denoted by

$$\mathsf{a}_{t|n}(\psi) = E(\alpha_t|Y_n, \psi). \tag{4}$$

The smoothed estimate can be calculated running the KF and storing the conditional means, $\mathsf{a}_t(\psi) = E(\alpha_t|Y_t, \psi)$ and $\mathsf{a}_{t|t-1}(\psi) = E(\alpha_t|Y_{t-1}, \psi)$, and variances, $P_t(\psi) = \mathrm{Var}(\alpha_t|Y_t, \psi)$ and $P_{t|t-1}(\psi) = \mathrm{Var}(\alpha_t|Y_{t-1}, \psi)$, of the state vector. The sequence of smoothed estimates $\mathsf{a}_{t|n}(\psi)$ is then calculated in reverse order, for $t = n, n-1, n-2, ..., 1$. The corresponding mean squared error is given by

$$\mathrm{MSE}(\mathsf{a}_{t|n}(\psi)) = P_{t|n}(\psi) = E\left[\left(\alpha_t - \mathsf{a}_{t|n}(\psi)\right)\left(\alpha_t - \mathsf{a}_{t|n}(\psi)\right)' |Y_n\right]. \tag{5}$$

This work employs a moment smoothing algorithm, which is described in Koopman, Shephard & Doornik (1999), using the KF output. More details about smoothing algorithms can be found in de Jong (1989) and Harvey (1989).

- **Forecasting**

The forecast of a future value y_{n+k}, $k > 0$, based on all available data $\boldsymbol{Y}_n$, can be obtained through the k-step-ahead forecast of the state vector, $\mathsf{a}_{n+k|n}(\boldsymbol{\psi})$, which is given by,

$$\mathsf{a}_{n+k|n}(\boldsymbol{\psi}) = E(\alpha_{n+k}|\boldsymbol{Y}_n, \boldsymbol{\psi}) = \left(\prod_{i=1}^{k} \boldsymbol{T}_{n+i}\right) \mathsf{a}_n(\boldsymbol{\psi}), \tag{6}$$

where $\mathsf{a}_n(\psi) = E(\alpha_n|Y_n, \psi)$.

By combining Equation (1) in time $n+k$ with Equation (6), the k-step-ahead forecast for $\{y_t\}$, defined by $\tilde{y}_{n+k|n}(\boldsymbol{\psi}) = E(y_{n+k}|\boldsymbol{Y}_n, \boldsymbol{\psi})$, can be calculated as

$$\tilde{y}_{n+k|n}(\boldsymbol{\psi}) = \boldsymbol{z}'_{n+k}\left(\prod_{i=1}^{k} \boldsymbol{T}_{n+i}\right) \mathsf{a}_n(\psi) = \boldsymbol{z}'_{n+k}\mathsf{a}_{n+k|n}(\boldsymbol{\psi}). \tag{7}$$

The corresponding mean square error is given by

$$\mathrm{MSE}(\tilde{y}_{n+k}|\boldsymbol{Y}_n, \boldsymbol{\psi}) = \mathbf{z}'_{n+k}\left(\prod_{i=1}^{k} \boldsymbol{T}_{n+i}\right) \mathbf{P}_n(\psi) \left(\prod_{i=1}^{k} \boldsymbol{T}'_{n+i}\right) \mathbf{z}_{n+k} + \\ \mathbf{z}'_{n+k}\sum_{i=1}^{k}\left(\prod_{j=1}^{k-i} \boldsymbol{T}_{n+k-j+1}\right) \boldsymbol{R}_{n+i}\boldsymbol{Q}_{n+i}\boldsymbol{R}'_{n+i}\left(\prod_{j=1}^{k-i} \boldsymbol{T}'_{n+k-j+1}\right) \mathbf{z}_{n+k} + h_{n+k}, \tag{8}$$

where $P_n(\psi) = \mathrm{Var}(\alpha_n|Y_n, \psi)$.

2.1 Asymptotic confidence intervals for smoothed estimates and forecasts

Let $\mathsf{a}_{\tau|n}(\psi)$ denote the optimal inference about α_t conditional on the whole sample. Thus, for $\tau \leq n$, $\mathsf{a}_{\tau|n}(\psi)$ is the smoothed inference given in (4), while for $\tau > n$, $\mathsf{a}_{\tau|n}(\psi)$ is the forecast of the state vector given in (6).

When the true value of ψ is known, the confidence interval of level $1-\kappa$ for the smoothed estimate is given by

$$\left[\mathsf{a}_{\tau|n}(\psi) \pm \left|z_{\kappa/2}\right| \sqrt{\mathrm{MSE}(\mathsf{a}_{\tau|n}(\psi))}\right], \quad \text{for} \quad \tau \leq n \tag{9}$$

while the prediction interval of level $1-\kappa$ is given by

$$\left[\tilde{y}_{\tau|n}(\boldsymbol{\psi}) \pm \left|z_{\kappa/2}\right| \sqrt{\mathrm{MSE}(\tilde{y}_{\tau|n}(\boldsymbol{\psi}))}\right], \quad \text{for} \quad \tau > n \tag{10}$$

where $\tilde{y}_{\tau|n}(\boldsymbol{\psi}) = \boldsymbol{z}_{\tau}' \mathsf{a}_{\tau|n}(\psi)$ and $\left|z_{\kappa/2}\right|$ is the quantile $\kappa/2$ of the standard normal distribution.

The value of ψ, however, is frequently unknown. In this case, it should be replaced by its MLE, $\hat{\psi}$, and the obtained interval is called natural or standard (Brockwell & Davis, 1996). The problem is that this interval does not incorporate the uncertainty related to ψ, which inevitably leads to underestimation of the PMSE in the classical inference (Harvey, 1989). Another question that arises is related to the normality assumption, which can be unrealistic in practice. For example, if the future observations assume an asymmetric distribution, the interval coverage rates in the tails may be affected (see Rodriguez & Ruiz (2010)).

In the next sections, some alternative to correct these problems are proposed, using the bootstrap.

3 Nonparametric bootstrap for estimating the PMSE

As stated in Section 2, the standard procedure to estimate the PMSE replaces the unknown parameter vector ψ by its MLE, $\hat{\psi}$. In order to correct the bias introduced in the MSE of the state vector and the predictions by this practice, some alternative approaches using the bootstrap are described in this section. As the procedures are very similar in both forecasting and smoothing contexts, only the notation for the forecast of a future value is used here to present the proposed algorithms. The procedures are easily adapted to the smoothed inference by replacing y_τ by $\mathbf{a}_\tau$ with $\tau \leq n$.

Consider the model given in (1 - 2). Initially the hyperparameters, which are the unknown variances of the errors ϵ_t and η_t, must be estimated. Then the KF is run to obtain the values of the estimated innovations, $\hat{\nu}_t$, and their variances, $\hat{F}_t$. It should be noted that these quantities are functions of the unknown parameters.

The standardized innovations, $\hat{e}_t = (\hat{\nu}_t - \bar{\nu})/\sqrt{\hat{F}_t}$, $t = 1, 2, \ldots, n$, where $\bar{\nu} = \sum_{j=1}^{n} \hat{\nu}_j / n$, are resampled to construct the bootstrap series (see Stoffer

& Wall (1991)). Then samples can be taken with replacement, from $\hat{e}_t$, to obtain the bootstrap innovations, $\hat{e}_t^*$.

The bootstrap series, y_t^*, is built, recursively, using the bootstrap innovations, $\hat{e}_t^*$, and the quantities $\hat{F}_t$ and $\hat{\mathbf{K}}_t$ obtained from the KF, where $\hat{\mathbf{K}}_t = \mathbf{T}_{t+1}\hat{\boldsymbol{P}}_{t|t-1}\mathbf{z}_t^{'}\hat{F}_t^{-1}$ and $\hat{\mathbf{P}}_{t|t-1} = Var(\boldsymbol{\alpha}_t|Y_{t-1})$. Initially, the KF equations for $\mathbf{a}_{t+1|t} = E(\boldsymbol{\alpha}_{t+1}|Y_t)$ and y_t are written in function of the innovations,

$$\mathbf{a}_{t+1|t} = \mathbf{T}_t\mathbf{a}_{t|t-1} + \mathbf{K}_t\nu_t$$

$$y_t = \mathbf{z}_t^{'}\mathbf{a}_{t|t-1} + \nu_t.$$

Next, the vector $\mathbf{S}_t = \begin{bmatrix} \mathbf{a}_{t+1|t} \\ y_t \end{bmatrix}$ is defined as

$$\mathbf{S}_t = \mathbf{A}_t\mathbf{S}_{t-1} + \mathbf{B}_t\nu_t, \quad t = 1, 2, ..., n, \tag{11}$$

where $\mathbf{A}_t = \begin{bmatrix} \mathbf{T}_t & 0 \\ \mathbf{z}_t & 0 \end{bmatrix}$ and $\mathbf{B}_t = \begin{bmatrix} \mathbf{K}_t\sqrt{F_t} \\ \sqrt{F_t} \end{bmatrix}$.

The bootstrap series, y_t^*, $t = 1, 2, ..., n$, is obtained by solving Equation (11), replacing ν_t by $\hat{e}_t^*$, $\mathbf{K}_t$ by $\hat{\mathbf{K}}_t$ and F_t by $\hat{F}_t$.

The procedures described in the next subsections use the bootstrap series y_t^* to obtain the forecasts and the prediction intervals for future values. They are based on the methods of Hamilton (1986), Ansley & Kohn (1986) and Pfeffermann & Tiller (2005), which calculate the PMSE for the state space model with estimated parameters.

3.1 Hamilton procedure with bootstrap resampling

Following the proposal of Hamilton (1986), which incorporates the uncertainty of the parameter estimation in estimating the state vector, the effect of the estimation of ψ in the forecasts can be eliminated using Monte Carlo integration. The predictive distribution of future observations is given by

$$\begin{aligned} p(y_\tau, \psi|Y_n) &= \int p(y_\tau, \psi|Y_n)d\psi \\ &= \int p(y_\tau|Y_n, \psi)p(\psi|Y_n)d\psi \\ &\cong \frac{1}{MC}\sum_{i=1}^{MC} p(y_\tau|Y_n, \psi^i), \end{aligned}$$

where ψ^i are samples from $p(\hat{\psi}|Y_n) \cong N(\hat{\psi}, I^{-1}(\hat{\psi}))$, the asymptotic distribution of the MLE of ψ, where $I(\cdot)$ is the Fisher information matrix.

The mean of the predictive distribution is given by:

$$\begin{aligned}\bar{y}_{\tau|n} &= \int \tilde{y}_{\tau|n}(\psi)p(\psi|Y_n)d\psi \\ &\cong \frac{1}{MC}\sum_{i=1}^{MC}\tilde{y}_{\tau|n}(\psi^i). \end{aligned} \tag{12}$$

Following the results in Hamilton (1986), with some modifications, the PMSE can be calculated as

$$\begin{aligned}\text{MSE}^{Ha}_{\tau|n} &= E\left[(y_\tau - \bar{y}_{\tau|n})^2\right] \\ &= \text{MSE}_{p(\psi|Y_n)}(\tilde{y}_{\tau|n}(\psi)) + E_{p(\psi|Y_n)}[(\tilde{y}_{\tau|n}(\psi) - \bar{y}_{\tau|n})^2]\end{aligned}$$

Then,

$$\widehat{\text{MSE}}^{Ha}_{\tau|n} \cong \frac{1}{MC}\sum_{i=1}^{MC}\widehat{\text{MSE}}(\tilde{y}_{\tau|n}(\psi^i)) + \frac{1}{MC}\sum_{i=1}^{MC}(\tilde{y}_{\tau|n}(\psi^i) - \bar{y}_{\tau|n})^2, \tag{13}$$

where $\bar{y}_{\tau|n}$ is given in equation (12), $\tilde{y}_{\tau|n}(\psi^i)$ in (7), $\widehat{\text{MSE}}(\tilde{y}_{\tau|n}(\psi^i))$ in (8) and ψ^i are samples obtained from $p(\hat{\psi}|Y_n) \cong N(\hat{\psi}, I^{-1}(\hat{\psi}))$.

Pfeffermann & Tiller (2005) argue that sampling ψ^i from $N(\hat{\psi}, I^{-1}(\hat{\psi}))$ may result in several problems, such as parameters being close to their boundary values, the distribution of $\hat{\psi}$ can be asymmetric in small samples and the calculation of the Fisher information matrix may become unstable for complex models. To avoid this, a simple and efficient procedure is to use bootstrap resamples $\psi^{\star(i)}$. Thus, the first procedure employed in this work is a variation of the Hamilton approach discussed above, using sampling bootstrap.

Procedure 1 (HaB):

1. Generate B nonparametric bootstrap series, y_t^*, using $\hat{\psi}$ estimated from the original series;
2. Calculate $\hat{\psi}^{*(b)}$ based on the bootstrap series $y_t^{*(b)}$, for $b = 1, \ldots, B$;
3. Using $\hat{\psi}^{*(1)}, \ldots, \hat{\psi}^{*(B)}$, compute:

- $\bar{y}_{\tau|n}$ given by equation (12);
- the PMSE given in equation (13).

4. The forecast interval of level $(1-\kappa)$ for $y_{\tau|n}$ is given by,

$$\left[\bar{y}_{\tau|n} \pm |z_{\kappa/2}| \sqrt{\widehat{\text{MSE}}^{HaB}_{\tau|n}}\right].$$

Assuming the state space form with Gaussian errors given in (1-2), Stoffer & Wall (1991) ensure that $p(\hat{\psi}^*|Y_n) \cong p(\hat{\psi}|Y_n)$ when B (the number of bootstrap replicates) is sufficiently large. That is, the distribution of $\hat{\psi}$ is approximated by the bootstrap distribution of the MLE and they are equivalent when the sample size is sufficiently large. Therefore, the distribution of $p(\hat{\psi}|Y_n)$ can be replaced by the distribution of $p(\hat{\psi}^*|Y_n)$ in the above procedure. This procedure is similar to the one described in Rodriguez & Ruiz (2012), for the prediction of the state vector.

3.2 Ansley and Kohn procedure

Ansley & Kohn (1986) proposed a procedure for incorporating the uncertainty in the estimation of ψ through a conditional mean square error for the estimate of the state vector. According to (Harvey, 1989, page 151), the idea can be used in the forecasting context through the approximation of $\text{MSE}(\tilde{y}_\tau|Y_n, \psi)$ by $\text{MSE}(\tilde{y}_\tau|Y_n, \hat{\psi})$ and the expansion of $\tilde{y}_{\tau|n}(\hat{\psi})$ around $\tilde{y}_{\tau|n}(\psi)$ up to the second term. The PMSE estimator is given by

$$\widehat{\text{MSE}}^{\text{AK}}_{\tau|n} = \widehat{\text{MSE}}(\tilde{y}_{\tau|n}(\hat{\psi})) + \left[\frac{\partial \tilde{y}_{\tau|n}(\psi)}{\partial \psi}\right]'_{\psi=\hat{\psi}} \left[I^{-1}(\hat{\psi})\right] \left[\frac{\partial \tilde{y}_{\tau|n}(\psi)}{\partial \psi}\right]_{\psi=\hat{\psi}}, \quad (14)$$

where $\hat{\psi}$ is the MLE of ψ and $I^{-1}(\hat{\psi})$ is the Fisher information matrix evaluated in $\hat{\psi}$.

The procedure of Ansley & Kohn (1986) incorporates the uncertainty in the estimation of hyperparameters in an elegant way, although the calculation of the Fisher information matrix and the derivation of the forecast function with respect to ψ can be a difficult task. The first problem can be solved in two ways:

1. Using a numerical approximation, as in Franco *et al.* (2008) and Santos & Franco (2010);

2. Using a bootstrap approximation for the Fisher information matrix $I_n^{-1}(\hat{\psi}) \cong \widehat{Cov}(\hat{\psi}^*)$. This approximation is supported by the same argument of the asymptotic validity of Procedure 1.

For the second problem, the calculation of $\left.\frac{\partial \tilde{y}_{\tau|n}(\psi)}{\partial \psi}\right|_{\psi=\hat{\psi}}$ can be done using numerical derivatives.

Using the empirical bootstrap distribution of the MLE of ψ, the following Ansley and Kohn procedure with bootstrap sampling (AKB) is proposed. This proposal is a variation of the Ansley and Kohn procedure, where the Fisher information matrix is approximated by the covariance matrix of the empirical bootstrap distribution for the MLE of ψ.

Procedure 2 (AKB):

1. The first steps are Steps (1) and (2) in Procedure 1.
2. With $\hat{\psi}^{*(1)}, \ldots, \hat{\psi}^{*(B)}$, compute:
   - $Cov(\hat{\psi}^*)$;
   - $\bar{y}_{\tau|n}$ given by equation (12);
   - The prediction MSE given by equation (14).
3. The forecast interval of level $(1-\kappa)$ for $y_{\tau|n}$ is given by,

$$\left[\bar{y}_{\tau|n} \pm |z_{\kappa/2}| \sqrt{\widehat{\text{MSE}}^{\text{AK}}_{\tau|n}}\right].$$

3.3 Pfeffermann and Tiller procedure

Pfeffermann & Tiller (2005) proposed a procedure that estimates the PMSE by incorporating the uncertainty in the hyperparameters estimation. An advantage of this procedure is that its PMSE estimator is $O(1/n^2)$ under certain conditions. Therefore, it is expected that the Pfeffermann and Tiller method produces better estimates for the PMSE. The procedure, called here PT, is described below.

Procedure 3 (PT):

1. The first steps are Steps (1) and (2) in Procedure 1.

2. Compute

$$\widehat{\text{MSE}}^{PT}_{\tau|n} = \widehat{\text{MSE}}^{bs}(\tilde{y}_{\tau|n}(\psi)) + 2\widehat{\text{MSE}}(\tilde{y}_{\tau|n}(\hat{\psi})) - \text{MSE}^{bs}(\tilde{y}_{\tau|n}(\psi)),$$

where $\widehat{\text{MSE}}^{bs}(\tilde{y}_{\tau|n}(\psi)) = \frac{1}{B}\sum_{b=1}^{B}\left[y^{\star(b)}_{\tau|n}(\hat{\psi}^{*b}) - y^{\star(b)}_{\tau|n}(\hat{\psi})\right]^2$

and

$\text{MSE}^{bs}(\tilde{y}_{\tau|n}(\psi)) = \frac{1}{B}\sum_{b=1}^{B}\widehat{\text{MSE}}(\tilde{y}_{\tau|n}(\hat{\psi}^{*b}))$.

It is important to observe that $y^{\star(b)}_{\tau|n}(\hat{\psi})$ and $y^{\star(b)}_{\tau|n}(\hat{\psi}^{*b})$ are the bootstrap estimates of the future observations with the estimates $\hat{\psi}$ and $\hat{\psi}^{*b}$ of ψ, respectively. $\widehat{\text{MSE}}(\tilde{y}_{\tau|n}(\hat{\psi}^{\star b}))$ is the natural estimator in Equation (8) with the original series, but using the estimate $\hat{\psi}^{*b}$.

3. The forecast interval of level $(1-\kappa)$ for $y_{\tau|n}$ is given by,

$$\left[\bar{y}_{\tau|n} \pm |z_{\kappa/2}|\sqrt{\widehat{\text{MSE}}^{PT}_{\tau|n}}\right],$$

where $\bar{y}_{\tau|n}$ is given in (12).

4 Parametric bootstrap forecast interval

Unlike the nonparametric bootstrap, in which the residuals of the fitted model are resampled with replacement, the parametric bootstrap uses only the parameter estimates from the original series. The procedure is performed as follows. The bootstrap errors ϵ^*_t and η^*_t are sampled from the Gaussian distribution with zero mean and variance replaced by the estimates obtained from the original series.

Then the bootstrap series y^*_t is constructed as:

$$\begin{aligned} y^*_t &= \mathbf{z}'_t\alpha^*_t + \hat{\epsilon}^*_t, \\ \alpha^*_t &= \mathbf{T}_t\alpha^*_{t-1} + \mathbf{R}_t\hat{\eta}^*_t, \quad t = 1, 2, \ldots, n. \end{aligned}$$

In the above equations, α^*_0 can be initialized using the same values specified when running the KF for the original series.

A new parametric bootstrap forecast interval, called here PBFI, which incorporates the uncertainty due to parameter estimation is presented in Procedure 4. Although this method can be performed for observations not necessarily Gaussian, the Gaussian distribution is assumed for the future observations. The procedure works well in the context of Gaussian state space models and is simpler than the sophisticated procedures proposed by Wall & Stoffer (2002) and Rodriguez & Ruiz (2010), preserving good asymptotic properties.

Procedure 4 (PBFI):

1. The first steps are Steps (1) and (2) in Procedure 1.
2. Using $\hat{\psi}^{*(1)}, \ldots, \hat{\psi}^{*(B)}$, run the KF for the original series and obtain $\mathsf{a}_n^{(b)}(\psi)$ and $\mathbf{P}_n^{(b)}(\psi)$, for $b = 1, \ldots, B$;
3. Generate the k-steps-ahead bootstrap forecasts using the following equations:
   (a) For $b = 1, \ldots, B$, calculate
   $\alpha_{n+k}^{*(b)} = \left(\prod_{i=1}^{k} \mathbf{T}_{n+i}^{*(b)}\right) \alpha_n^{*(b)} + \sum_{i=1}^{k} \left(\prod_{j=1}^{k-i} \mathbf{T}_{n+k-j+1}^{\prime *(b)}\right) \mathbf{R}_{n+i}^{\prime *(b)} \eta_{n+i}^{*(b)}$,
   where $\eta_{n+i}^{*(b)} \sim F(\mathbf{0}, \hat{\mathbf{Q}}_t^{*(b)})$ and $\alpha_n^{*(b)}$ is generated from the Normal distribution with mean $\mathsf{a}_n^{(b)}(\psi)$ and covariance matrix $\mathbf{P}_n^{(b)}(\psi) + (\mathsf{a}_n^{(b)}(\psi) - \mathsf{a}_n(\psi)) \times (\mathsf{a}_n(\psi)^{(b)} - \mathsf{a}_n(\psi))'$.
   (b) $y_{n+k}^{*(b)} = \mathbf{z}_{n+k}^{\prime *(b)} \alpha_{n+k}^{*(b)} + \epsilon_{n+k}^{*(b)}$, where $\epsilon_{n+k}^{*(b)}$ is generated from a G distribution with zero mean and variance $\hat{h}_t^{*(b)}$.
4. The bootstrap forecast interval of level $(1 - \kappa)$ for y_{n+k} is given by,
$$\left[y_{n+k}^{*(\kappa/2)}; y_{n+k}^{*(1-\kappa/2)} \right],$$
   where $y_{n+k}^{*(\kappa)}$ is the quantile of order κ.

The G distribution is not necessarily Gaussian, despite being the most frequent case. The term $(\mathsf{a}_n^{(b)}(\psi) - \mathsf{a}_n(\psi)) \times (\mathsf{a}_n^{(b)}(\psi) - \mathsf{a}_n(\psi))'$ was introduced in the covariance matrix to incorporate the uncertainty in the estimation of ψ when generating values of the state vector.

The asymptotic validity of Procedure 4 under Gaussian errors can be easily obtained, although it is not difficult to prove the non-Gaussian case. If the distribution of the observation equation is not Gaussian, the MLE of ψ becomes the maximum quasi-likelihood estimator of ψ, which is consistent and preserves good properties (Harvey, 1989).

5 Simulation results

The bootstrap procedures described in Sections 3 and 4 are compared through a Monte Carlo study and the results are presented separately for smoothing and forecasting.

The experiments are performed for the simplest structural model, known as the Local Level Model (LLM) (Harvey (1989)), defined as,

$$\begin{aligned} y_t &= \alpha_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2) \\ \alpha_t &= \alpha_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2). \end{aligned} \tag{15}$$

Different values for the signal-to-noise ratio, $q = \sigma_\eta^2/\sigma_\epsilon^2$, are used in the simulations, with $\sigma_\epsilon^2 = 1$. To check the robustness of the methods with respect to the Gaussian assumption, a Gamma distribution for the error terms, re-centered and scaled so that the errors have zero mean and variance 1, was also assumed. All models were fitted using the KF and the Gaussian likelihood function and the simulations were implemented in the Ox Software ?.

5.1 Smoothing

The PMSE of the smoothed state vector is estimated using the nonparametric bootstrap procedures proposed in Section 3. To this purpose, 1000 time series are generated for the LLM, with Gaussian and non-Gaussian distributions, for three sample sizes $n = 40$, $n = 100$ and $n = 500$ and with signal-to-noise ratio $q = 0.25$. These values and distributions were chosen according to a simulation study in Queenville & Singh (2000). The number of bootstrap replications was 2000.

The estimators of the PMSE compared here are: Standard (S), obtained by substituting the hyperparameters in equation (15) by their MLE to calculate the PMSE of $\mathsf{a}_{t|n}(\psi)$; Hamilton with bootstrap resample (HaB) described in Procedure 1; Ansley and Kohn with bootstrap resample (AKB), defined in Procedure 2; and Pfeffermann & Tiller estimator (PT) described in Procedure 3. For the first three estimators, a numerical approximation for the Fisher information matrix was used.

Following the Pfeffermann & Tiller (2005) and Queenville & Singh (2000) experiments, the procedures are compared using the relative bias (Rel-Bias) and relative square root of the mean square error (Rel-SMSE), given respectively by,

$$\text{Rel-Bias} = \frac{100}{n}\sum_{t=1}^{n}\frac{\bar{d}_t(\psi)}{MSE_t(\psi)}$$

and

$$\text{Rel-SMSE} = \frac{100}{n}\sum_{t=1}^{n}\frac{(\bar{d}_t^{(2)}(\psi))^{1/2}}{MSE_t(\psi)},$$

where $\bar{d}_t(\psi) = \sum_{i=1}^{1000}\frac{d_{i,t}}{1000}$, $\bar{d}_t(\psi) = \sum_{i=1}^{1000}\frac{d_{i,t}^2}{1000}$, $d_{i,t} = [M\hat{S}E_t^{(i)}(\psi) - MSE_t(\psi)]$ and MSE_t is the true PMSE of $\mathsf{a}_{t|n}(\psi)$ calculated for each time $t = 1, \ldots, n$ by simulating 10000 series for each lenght ($n = 40, 100, 500$) according to

$$MSE_t(\psi) \cong \sum_{i=1}^{10000}\frac{(\mathsf{a}_{t|n}(\hat{\psi})^{(i)} - \alpha_t^{(i)})^2}{10000}.$$

Table 1 shows the Rel-Bias and Rel-SMSE for the PMSE of the smoothed estimator in the LLM with Gaussian and non Gaussian errors. In the case of Gaussian errors, it can be seen that the bootstrap procedures are much better than the standard (S) with respect to the bias, especially for small sample sizes. Among the bootstrap procedures, the PT is the one which presents the smallest mean percent Rel-Bias for all sample sizes. With respect to the nonparametric bootstraps proposed here, the AKB is superior to the HaB procedure. In spite of possessing higher Rel-Bias, the S estimator has a similar or inferior Rel-SMSE compared to the bootstrap estimators. This is not an unexpected result, once the bias correction can increase the variance (see Pfeffermann & Tiller (2005)). Although the PT presents a better performance with respect to the Rel-Bias compared to the other methods, the nonparametric bootstrap procedures proposed here are simpler and require less computational efforts. This can bring a significant gain when the model complexity increases.

Table 1: Relative bias (%) and relative MSE square root (%) of the smoothed estimators for a local level model with Gaussian (G) and Non-Gaussian (NG) errors and signal-to-noise ratio $q = 0.25$.

		n = 40		n = 100		n = 500	
		Rel-Bias	Rel-SMSE	Rel-Bias	Rel-SMSE	Rel-Bias	Rel-SMSE
G	S	-23.422	19.413	-8.677	10.315	-2.235	4.021
	HaB	-8.224	24.204	-5.726	10.407	-2.158	3.917
	AKB	-4.460	28.131	-4.003	11.224	-1.396	3.940
	PTB	0.110	19.828	0.742	10.108	0.598	3.903
NG	S	-24.498	28.389	-10.294	15.093	-1.545	6.007
	HaB	-13.536	34.921	-10.102	16.990	-1.592	6.269
	AKB	-10.934	31.445	-4.906	15.621	-0.547	6.084
	PTB	2.567	31.125	0.742	10.108	0.519	5.987

In the non-Gaussian case, Gamma$(16/9, 3/4)$ and Gamma$(25/16, 4/5)$ distributions for ϵ_t and η_t, respectively, were assumed. The Rel-Bias and Rel-SMSE are larger for non-Gaussian errors, although the behavior of the PMSE estimators is similar to the Gaussian case.

5.2 Forecasting

The performance of the bootstrap methods presented in Sections 3 and 4 are studied here in the forecasting context. The following procedures: Natural (N) shown in Equation (11), Procedure 1 (HaB), Procedure 2 (AKB), Procedure 3 (PT) and the parametric bootstrap from Procedure 4 (PBFI) are compared through the average width and coverage rates of the forecast intervals. For this study, series of size $n = 50$, 100 and 500 were generated with a burn-in equal to 100. The values of the hyperparameters were chosen to equal the following signal-to-noise ratio: $q = 0.1$ and $q = 1.0$. The forecasts were calculated k-steps-ahead for $k = 1, 5$ and 15. The number of Monte Carlo replications was $MC = 1000$, the number of bootstrap resamples was $B = 2000$ and the nominal level of the prediction intervals was fixed at 95%. These values were chosen according to a simulation study in Rodriguez & Ruiz (2010). For the non-Gaussian case, a Gamma$(1/9, 1/3)$ distribution for ϵ_t was assumed.

Tables 2 and 3 contain the results for Gaussian errors with $q = 0.1$ and $q = 1.0$, respectively. It can be seen that the proposed methods are consistent, as for all forecast lags, the coverage rates get close to the fixed 95% level assumed, as the sample size increases. It can be also noted that, even for small samples, all methods present very satisfactory results, with the AKB, PT and PBFI intervals showing coverage slightly better than the other procedures. Note that the bootstrap approaches present width greater than the N method, particularly for small samples, once they incorporate the uncertainty inherent in the estimation of ψ. The parametric bootstrap (PBFI) shows a better performance when $k = 1$, but its width is slightly larger than the other methods. The coverage on the tails appears to be symmetric for all methods, as expected. It can also be noted that the average width of the intervals increases as q increases.

Although the forecasting methods proposed here were able to correct the PMSE (the average width of the bootstrap intervals is larger than the N procedure), this correction does not seem to affect the coverage rates. Similar results can also be seen in Rodriguez & Ruiz (2010).

Tables 4 and 5 contain the results for Gamma errors with $q = 0.1$ and $q = 1.0$, respectively. Once again, it can be seen that the coverage rates get close to the fixed 95% level as the sample size increases, for all forecast lags. In this case, the coverage rate of the N method is slightly smaller than the bootstrap methods when the sample size is small ($n = 50$). It can be also seen that the bootstrap methods, which incorporate the uncertainty in the estimation of ψ, have larger width, as expected, especially for small sample sizes. Due to the asymmetry of the Gamma distribution, the coverage rates in the tails are not symmetric, but the asymmetry is more stressed when $q < 1$.

6 Application to real data sets

6.1 Electric energy consumption

To illustrate the smoothing procedures, this section presents a study on a time series of electric energy consumption in the Northeast region of Brazil. These data were obtained from a large study concerning the quantity of energy necessary to answer the maximum demand in the peak interval (from 6:00 pm to 9:00 pm). The series are monthly observations of electric consumption from CHESF (São Francisco Hydroelectric Company), in the period from May 1991 to December 1996 ($n = 68$). The data are shown in Figure 1

Table 2: 95% prediction intervals for y_{n+k} with $q = 0.1$ and Gaussian errors.

n	k	N Width Coverage Tails	PBFI Width Coverage Tails	HaB Width Coverage Tails	AKB Width Coverage Tails	PT Width Coverage Tails
50	1	4.462	4.707	4.536	4.590	4.601
		0.913	0.933	0.919	0.920	0.922
		0.041/0.046	0.035/0.032	0.038/0.043	0.037/0.043	0.037/0.041
100	1	4.533	4.631	4.574	4.561	4.603
		0.955	0.953	0.954	0.954	0.956
		0.021/0.024	0.019/0.028	0.020/0.026	0.020/0.026	0.020/0.024
500	1	4.582	4.566	4.579	4.586	4.592
		0.949	0.952	0.955	0.954	0.955
		0.022/0.029	0.020/0.028	0.017/0.028	0.016/0.030	0.016/0.029
50	5	5.099	5.340	5.174	5.229	5.215
		0.938	0.942	0.939	0.941	0.941
		0.031/0.031	0.028/0.030	0.030/0.031	0.029/0.032	0.029/0.030
100	5	5.147	5.240	5.180	5.174	5.213
		0.936	0.944	0.938	0.939	0.941
		0.031/0.033	0.025/0.031	0.03/0.032	0.030/0.031	0.030/0.029
500	5	5.214	5.198	5.210	5.218	5.225
		0.950	0.949	0.951	0.951	0.951
		0.023/0.027	0.026/0.025	0.023/0.026	0.023/0.026	0.023/0.026
50	15	6.331	6.606	6.410	6.448	6.406
		0.918	0.926	0.923	0.924	0.919
		0.044/0.033	0.045/0.029	0.047/0.030	0.044/0.032	0.049/0.032
100	15	6.409	6.504	6.429	6.440	6.472
		0.926	0.930	0.928	0.929	0.930
		0.037/0.037	0.037/0.033	0.038/0.034	0.037/0.034	0.037/0.033
500	15	6.485	6.464	6.479	6.489	6.498
		0.946	0.938	0.944	0.946	0.947
		0.028/0.026	0.033/0.029	0.029/0.027	0.028/0.026	0.028/0.025

Table 3: 95% prediction intervals for y_{n+k} with $q = 1.0$ and Gaussian errors.

n	k	N Width Coverage Tails	PBFI Width Coverage Tails	HaB Width Coverage Tails	AKB Width Coverage Tails	PT Width Coverage Tails
50	1	6.212	6.398	6.282	6.292	6.381
		0.924	0.933	0.927	0.929	0.932
		0.041/0.035	0.036/0.031	0.038/0.035	0.036/0.035	0.038/0.030
100	1	6.287	6.330	6.292	6.326	6.366
		0.952	0.952	0.952	0.955	0.956
		0.022/0.026	0.023/0.025	0.022/0.026	0.021/0.024	0.021/0.023
500	1	6.326	6.315	6.326	6.333	6.334
		0.949	0.946	0.949	0.949	0.949
		0.020/0.031	0.020/0.034	0.020/0.031	0.020/0.031	0.020/0.031
50	5	9.772	10.015	9.861	9.829	9.845
		0.925	0.933	0.927	0.927	0.929
		0.036/0.039	0.032/0.035	0.035/0.038	0.032/0.041	0.033/0.038
100	5	9.915	9.996	9.956	9.940	9.928
		0.945	0.938	0.944	0.944	0.944
		0.024/0.031	0.026/0.036	0.024/0.032	0.024/0.032	0.024/0.032
500	5	10.035	10.008	10.046	10.040	10.033
		0.949	0.943	0.948	0.948	0.950
		0.026/0.025	0.031/0.026	0.026/0.026	0.025/0.027	0.025/0.025
50	15	15.504	15.836	15.521	15.504	15.517
		0.911	0.913	0.910	0.911	0.913
		0.048/0.041	0.047/0.040	0.048/0.042	0.047/0.042	0.047/0.040
100	15	15.793	15.943	15.830	15.808	15.788
		0.928	0.927	0.925	0.925	0.925
		0.034/0.038	0.036/0.037	0.035/0.040	0.036/0.039	0.037/0.038
500	15	15.881	15.842	15.886	15.884	15.882
		0.955	0.952	0.955	0.955	0.955
		0.021/0.024	0.021/0.027	0.021/0.024	0.021/0.024	0.021/0.024

Table 4: 95% prediction intervals for y_{n+k} with $q = 0.1$ and Gamma errors.

n	k	N Width Coverage Tails	PBFI Width Coverage Tails	HaB Width Coverage Tails	AKB Width Coverage Tails	PT Width Coverage Tails
50	1	4.131	4.386	4.183	4.748	4.265
		0.921	0.929	0.925	0.929	0.927
		0.017/0.062	0.011/0.060	0.014/0.061	0.011/0.060	0.013/0.060
100	1	4.345	4.502	4.382	4.433	4.413
		0.962	0.968	0.963	0.963	0.962
		0.010/0.028	0.006/0.026	0.009/0.028	0.009/0.028	0.010/0.028
500	1	4.531	4.531	4.600	4.536	4.543
		0.956	0.956	0.956	0.956	0.956
		0.006/0.038	0.008/0.036	0.006/0.038	0.006/0.038	0.021/0.025
50	5	4.893	5.170	4.993	5.488	4.993
		0.934	0.941	0.934	0.945	0.937
		0.024/0.042	0.023/0.036	0.024/0.042	0.018/0.037	0.025/0.038
100	5	5.020	5.168	5.047	5.166	5.078
		0.960	0.959	0.959	0.959	0.960
		0.004/0.036	0.004/0.037	0.004/0.037	0.004/0.037	0.004/0.036
500	5	5.156	5.151	5.154	5.160	5.165
		0.952	0.952	0.952	0.952	0.952
		0.009/0.039	0.009/0.039	0.009/0.039	0.009/0.039	0.009/0.039
50	15	6.139	6.435	6.220	6.580	6.191
		0.917	0.924	0.923	0.927	0.923
		0.044/0.039	0.040/0.036	0.041/0.036	0.037/0.036	0.042/0.035
100	15	6.383	6.562	6.404	6.523	6.434
		0.942	0.940	0.943	0.946	0.945
		0.021/0.037	0.022/0.038	0.020/0.037	0.017/0.037	0.018/0.037
500	15	6.462	6.468	6.457	6.466	6.475
		0.957	0.960	0.957	0.958	0.959
		0.019/0.024	0.016/0.024	0.019/0.024	0.018/0.024	0.018/0.023

Table 5: 95% prediction intervals for y_{n+k} with $q = 1.0$ and Gamma errors.

n	k	N Width Coverage Tails	PBFI Width Coverage Tails	HaB Width Coverage Tails	AKB Width Coverage Tails	PT Width Coverage Tails
50	1	5.978	6.155	6.026	6.068	6.136
		0.936	0.939	0.935	0.938	0.940
		0.027/0.037	0.025/0.036	0.026/0.039	0.025/0.037	0.025/0.035
100	1	6.151	6.216	6.157	6.191	6.228
		0.943	0.951	0.945	0.947	0.945
		0.030/0.027	0.025/0.024	0.027/0.028	0.026/0.027	0.029/0.026
500	1	6.299	6.267	6.298	6.306	6.314
		0.948	0.954	0.948	0.948	0.949
		0.021/0.031	0.018/0.028	0.021/0.031	0.020/0.032	0.020/0.031
50	5	9.804	9.974	9.790	9.874	9.961
		0.927	0.929	0.927	0.926	0.929
		0.032/0.041	0.031/0.040	0.032/0.041	0.029/0.045	0.029/0.042
100	5	9.031	9.994	9.934	9.953	9.976
		0.942	0.949	0.944	0.944	0.944
		0.016/0.042	0.013/0.038	0.015/0.041	0.014/0.042	0.016/0.040
500	5	10.031	10.011	10.035	10.035	10.035
		0.942	0.940	0.942	0.942	0.940
		0.023/0.035	0.025/0.035	0.025/0.035	0.023/0.035	0.023/0.037
50	15	15.443	15.655	15.341	15.476	15.628
		0.918	0.915	0.914	0.919	0.920
		0.043/0.039	0.043/0.042	0.043/0.043	0.042/0.039	0.042/0.038
100	15	15.816	15.981	15.817	15.831	15.846
		0.925	0.923	0.927	0.926	0.927
		0.033/0.042	0.036/0.041	0.031/0.042	0.031/0.043	0.032/0.041
500	15	15.926	15.917	15.934	15.929	15.924
		0.941	0.943	0.942	0.941	0.943
		0.027/0.032	0.025/0.032	0.027/0.031	0.026/0.033	0.026/0.031

Table 6: MLE and confidence intervals for the hyperparameters of CHESF series.

ψ	MLE	Conf. Int. 95%
σ^2_η	0.14	[0.03; 0;31]
σ^2_ϵ	0.25	[0.11; 0.40]

Table 7: Percentual increasing of MSE compared to the Standard procedure for the CHESF series.

Methods	Percent
PT	9.56
HaB	7.71
AKB	6.40

and the LLM was fitted to the series.

Point and interval estimates for the hyperparameters are shown in Table 6. The signal-to-noise ratio is less than 1 (0.56). A residual analysis was carried out and no evidence of correlation across time in the error term was found. As the zero is not included in the intervals, the LLM can be an adequate model for this series.

Table 7 provides the percentage increase of MSE compared to the standard procedure. The PT method presents the highest increasing, while the AKB and HaB procedures have a similar performance. Figure 1 provides smoothed estimates (point and 95% confidence intervals) for the level component using the PT method, which follows well the series behavior.

6.2 Income of a small Brazilian city

This application deals with the net income series from a small Brazilian city, whose name is omitted here for confidentiality. It is important to forecast future values of income series for planning and control of the annual budget and costs.

The income series, shown in Figure 2, consists of 72 monthly observations in the period 2006/01 to 2011/12. The last twelve observations of the series were omitted to compare the forecast intervals to the future values. The data do not seem to follow a Normal distribution, although they are not too

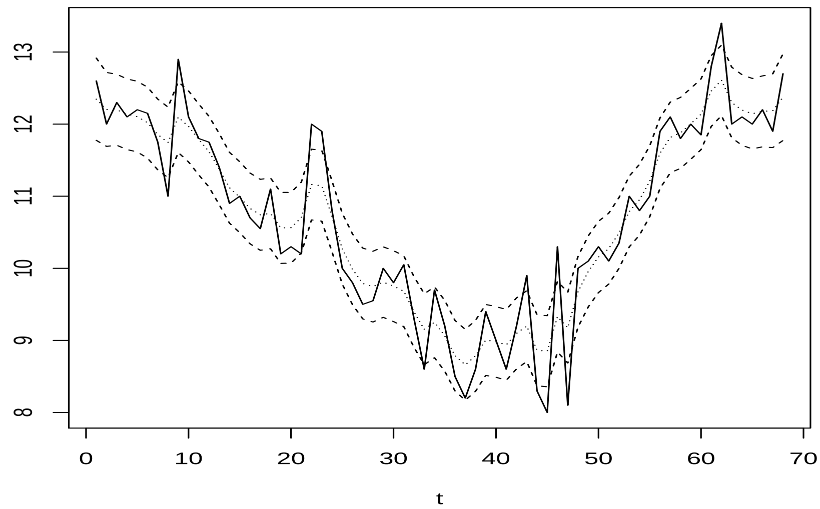

Figure 1: Electric energy consumption of CHESF. Values of the series were divided by 10. The dotted and dashed lines represent, respectively, the smoothing point estimates and the confidence intervals of level 95% for the level component using PT metohd.

asymmetric.

The series is non-stationary and has a seasonal component, thus a Structural Basic Model (SBM) is fitted to the series (Harvey, 1989). For the SBM the hyperparameter vector is given by $\psi_t = (\sigma^2_\epsilon, \sigma^2_\eta, \sigma^2_\xi, \sigma^2_\omega)$, where the components represent, respectively, the variance of the error, level, slope and seasonality.

The residuals of the fitted model do not seem to follow a normal distribution, but they are not autocorrelated. Table 8 presents the point and interval estimates for the hyperparameters. The point estimates for the slope (σ^2_ξ) and seasonal (σ^2_ω) hyperparameters are very close to zero, implying that these components should have a deterministic behavior. For this application, the forecast intervals N, HaB and PBFI were calculated and they are shown in Figure 2. It can be seen that the real future values are all of them inside the prediction intervals, but it is important to emphasize that the lower limit of the N interval is very close to the real future values. Note that the HaB and PBFI intervals present larger widths compared to N, as it was shown in the

Table 8: MLE and bootstrap confidence intervals of nominal level 95% for the hyperparameters of the SBM fitted to the income series.

ψ	Estimate	Bootstrap confidence interval
σ^2_η	0.001	$(5.61 \times 10^{-12}$; $0.015)$
σ^2_ϵ	0.005	$(2.75 \times 10^{-9}$; $0.001)$
σ^2_ξ	1.14×10^{-10}	$(7.15 \times 10^{-16}$; $1.157 \times 10^{-5})$
σ^2_ω	2.66×10^{-8}	$(5.15 \times 10^{-21}$; $0.0005)$

simulation studies.

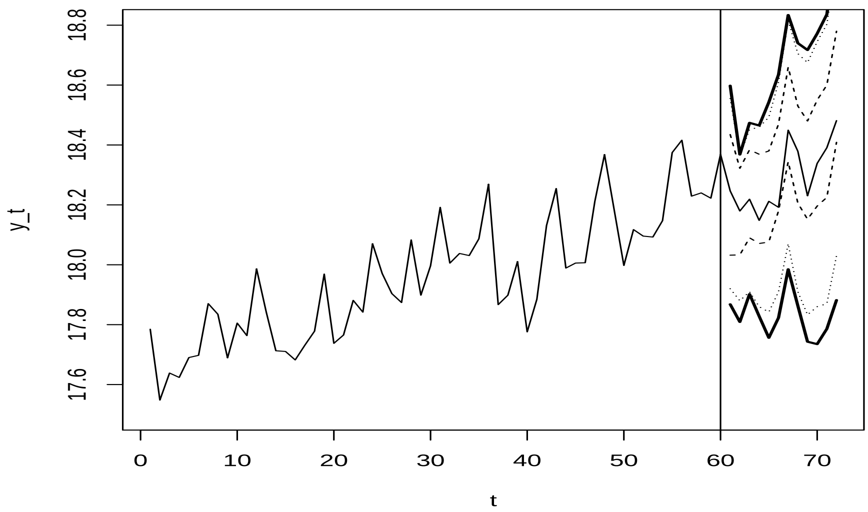

Figure 2: The full, dashed, dotted and full in bold lines indicate, respectively, the income series and the forecast intervals N, HaB and PBFI. The vertical line separates the data and the future values. The confidence level is 0.95.

7 Conclusion and Final Remarks

In this work, different bootstrap procedures to estimate the PMSE, which take into account the uncertainty associated to the hyperparameters estimation, have been proposed and compared for the SSM. These proposals are variations of the procedures developed in Hamilton (1986) and Ansley & Kohn (1986), along with a new parametric bootstrap procedure (PBFI). The methods were empirically compared to the standard procedure and to the proposal of Pfeffermann & Tiller (2005), under Gaussian and non-Gaussian assumptions for the errors, in the forecasting and smoothing contexts. The results confirmed that the performance of the bootstrap prediction intervals is slightly better than the natural (N) intervals with respect to the coverage rate, in the case of Gaussian distribution. The PMSE for the future observations was corrected using the bootstrap technique and had a larger value than the N procedure, but this correction does not seem to drastically affect the coverage rate of the intervals.

The results also showed that the presence of non-Gaussian errors with a very strong asymmetry directly interferes in the coverage rate in the tails. The major advantage of the bootstrap methods addressed here is their computational simplicity, particularly the parametric bootstrap interval, and the fact that they do not underestimate the variability of the forecast error.

Smoothing results were very satisfactory, especially for the AKB and HaB procedures, taking into account the quality of the estimates and the computational time. For more details see Santos (2012).

Future research includes the development of forecasting methods that do not employ the normality assumption for future values. This includes the use of asymmetric distributions for correcting the coverage rate of the intervals. New possibilities for the G distribution, other than the Gaussian, can also be studied in the PBFI method.

References

Anderson, B. D. O. & Moore, J. B. (1979). *Optimal Filtering*, Englewood Cliffs, New Jersey, Prentice Hall.

Ansley,C. F. & Kohn, R. (1986). Prediction mean square error for state space models with estimated parameters. *Biometrika*, **73**, 467-73.

Brockwell, P. J. & Davis, R. A. (1996). *Introduction to Time Series and Forecasting*, Springer text in Statistics. New York: Springer.

de Jong, P. (1988). A cross-validation filter for time series models. *Biometrika*, **75**, 594-600.

de Jong, P. (1989). Smoothing and interpolation with the state space model. *Journal of the American Statistical Association*, **84**, 1085-1088.

Doornik, J. A. (2002). *Ox: An Object-Oriented Matrix Language*. 3. ed. London: Timberlake Consultants Press.

Franco, G. C., Santos, T. R., Ribeiro, J. A. & Cruz, F. R. B.(2008). Confidence intervals for hyperparameters in structural models. *Communications in Statistics: Simulation and Computation*, **37**(3), 486-497.

Hamilton, J. D. (1986). A standard error for the estimated state vector of a state space model. *Journal of Econometrics*, **33**, 387-97.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: University Press.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.*, **82**, 35-45.

Koopman, S. J., Shephard, N. & Doornik, J. A. (1999). Statistical algorithms for inference in state space using SsfPack 2.2 *Econometrics Journal*, **47**, 129-139.

Pfeffermann, D. & Tiller, R. (2005). Bootstrap approximation to prediction MSE for state-space models with estimated parameters. *Journal of Time Series Analysis*, **26**, 893-916.

Queenville, B. & Singh, A. C. (2000). Bayesian prediction mean squared error for state space models with estimated parameters. *Journal of Time Series Analysis*, **21**, 2, 219-236.

Rodriguez, A. & Ruiz, E. (2010) Bootstrap Prediction Intervals in State Space Models. *Journal of the Time Series Analysis*, **30**, 167-178.

Rodriguez, A. & Ruiz, E. (2012) Bootstrap prediction mean squared errors of unobserved states based on the Kalman filter with estimated parameters. *Computational Statistics and Data Analysis*, **56**, 62-74.

Santos, T. R. & Franco, G. C. (2010). Inference for the hyperparameters of structural models under classical and Bayesian perspectives: a comparison study. *Communications in Statistics: Simulation and Computation*, **39**: 1671-1689.

Santos, T. R. (2012). Inferência, previsão e suavização em modelos estruturais Gaussianos e não-Gaussianos (in portuguese). PhD Thesis, Depto Statistics, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

Shephard, N. (1993) Maximum likelihood estimation of regression models with stochastic trend components. *Journal of the American Statistical Association*, **88**, 590-595.

Stoffer, D. S. & Wall, K. D. (1991). Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter. *Journal of the American Statistical Association*, **86**, 1024-1033.

Tsimikas, J. & Ledolter, J. (1994). REML and best linear unbiased prediction in state space models. *Communications in Statistics: Theory and Methods*, **23** (8), 2253-2268.

Wall, K. D. & Stoffer, D. S. (2002) A state space approach to bootstrapping conditional forecasts in ARMA models. *Jr. Time Series Analysis*, **23**, 733-751.

Watanabe, N. (1985) Note on the Kalman filter with estimated parameters. *Jr. Time Series Analysis*, **6**, 269-278.

West, M. & Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. New York: Springer.

Yamamoto, T. (1976). Asymptotic mean squared error prediction for an autoregressive model with estimated coefficients. *Appl. Statist.* 25, 123-7.