Working Paper 01-05
Statistics and Econometrics Series 03
January 2001

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 624-9849

BOOTSTRAP PREDICTION INTERVALS FOR POWER-TRANSFORMED TIME SERIES

Lorenzo Pascual, Juan Romo and Esther Ruiz*

Abstract

In this paper we propose a bootstrap resampling scheme to construct prediction intervals for future values of a variable after a linear ARIMA model has been fitted to a power transformation of it. The advantages over existing methods for computing prediction intervals of power transformed time series are that the proposed bootstrap intervals incorporate the variability due to parameter estimation, and do not rely on distributional assumptions neither on the original variable nor on the transformed one. We show the good behavior of the bootstrap approach versus alternative procedures by means of Monte Carlo experiments. Finally, the procedure is illustrated by analysing three real time series data sets.

Keywords: Forecasting; Non gaussian distributions; Box-Cox transformations; Resampling methods.

*Pascual, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-91-624.98.59, Fax: 34-91-624.98.49, e-mail: pascual@est-econ.uc3m.es; Romo, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Ph: 34-91-6249805, e-mail: romo@est-econ.uc3m.es; Ruiz, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Ph: 34-91-

Bootstrap Prediction Intervals for Power-transformed Time Series

Lorenzo Pascual, Juan Romo and Esther Ruiz*

Departamento de Estadística y Econometría

Universidad Carlos III de Madrid

WP 01 05

Statistics and Econometric Series 03

January 2001

*Corresponding author. Departamento de Estadística y Econometría. Universidad Carlos III de Madrid. C/ Madrid, 126, 28903 Getafe (Madrid), SPAIN. Tel: +34 91 624 9851, Fax: +34 91 624 9849, e-mail: ortega@est-econ.uc3m.es

Abstract

In this paper we propose a bootstrap resampling scheme to construct prediction intervals for future values of a variable after a linear ARIMA model has been fitted to a power transformation of it. The advantages over existing methods for computing prediction intervals of power transformed time series are that the proposed bootstrap intervals incorporate the variability due to parameter estimation, and do not rely on distributional assumptions neither on the original variable nor on the transformed one. We show the good behavior of the bootstrap approach versus alternative procedures by means of Monte Carlo experiments. Finally, the procedure is illustrated by analysing three real time series data sets.

KEY WORDS: Forecasting, Non Gaussian distributions, Box-Cox transformations, Resampling methods.

1. INTRODUCTION

Forecasting future values of time series data is one of the main objectives of time series analysis. Generally, predictions are given as point forecasts, although, it is even more important to provide as well interval forecasts; see, for example, Chatfield (1993).

In empirical time series analysis, it is common practice to transform the data using a power transformation prior to the estimation of the model used for forecasting. There are several reasons to transform the data before fitting a suitable model, for example, the necessity of stabilizing the increasing variance of trending time series, to reduce the impact of outliers, to make the normal distribution a better approximation to the data distribution, or because the transformed variable has a convenient economic interpretation; for example, first differenced log-transformed data correspond to growth rates.

The family of Box-Cox transformations is given by

$$\begin{aligned} g(X_t) &= \frac{X_t^{\lambda}-1}{\lambda}, \text{ for } 0<\lambda<1 \\ &= \ln(X_t), \quad \text{for } \lambda = 0, \end{aligned} \tag{1}$$

where the transformation for $\lambda = 0$ follows from the fact that $\lim_{\lambda\to 0}\frac{X_t^{\lambda}-1}{\lambda} = \ln(X_t)$; see Box and Cox (1964). Substracting 1 and dividing by λ does not influence the stochastic structure of X_t^{λ}, and hence one often considers the transformation suggested by Tukey (1957)

$$\begin{aligned} g(X_t) &= X_t^{\lambda}, \qquad \text{for } 0<\lambda<1 \\ &= \ln(X_t), \text{ for } \lambda = 0, \end{aligned} \tag{2}$$

instead of (1), without loss of generality. In both cases, $\{X_t\}$ denotes the observed time series with $X_t > 0$, λ is a real number and $\ln(\cdot)$ denotes the natural logarithm.

Once a model has been estimated, point and interval forecasts can be obtained for the transformed series $y_t = g(x_t)$ defined as in (2). We will focus on ARIMA models fitted to y_t. The specification of the model and the parameter λ will be assumed to be known. If the objective is to predict future values of X_t, the retransformed point forecasts induces bias in the forecasts, as is shown for linear models in Granger and Newbold (1976). When Y_{T+k} is normally distributed and the point forecast of X_{T+k} is just the inverse transformation of the forecast obtained for the transformed variable, this naive point prediction is not the minimum mean squared error (MMSE) forecast, but the minimum mean absolute error (MMAE), that is the median of the conditional probability density function (pdf) of X_{T+k}. Therefore, if the error loss function is quadratic, this naive prediction of X_{T+k} is not optimal; see Guerrero (1993).

Assuming Gaussianity of Y_t, Granger and Newbold (1976) propose a debiasing factor to reduce the transformation bias in the point forecast. Unfortunately, since they solve the problem using Hermite polynomials expansions, their procedure becomes very complicated for many fractional power transformations, making this approach not useful in practice. Latter, Taylor (1986) proposes a simpler expression for the debiasing factor, but for $\lambda = 0$ does not provide an adequate solution. Notice that the logarithmic transformation is one of the most usual in practice. Another alternative proposed by Pankratz and Dudley (1987) is complicated, and additionally, only admits a closed form expression when λ is a fractional positive integer. Finally, the method proposed by Guerrero (1993) avoids all the drawbacks found in previous approaches. His proposal is both simple and general. In a comparative study, Guerrero (1993) shows that his method has a performance similar to or better than the other procedures.

Although it is relatively well studied how to obtain a good estimate for the conditional mean in the original metric, there is no generally accepted method of con-

structing prediction intervals for the untransformed variable. One solution is based on a normal assumption on X_{T+k}, providing a symmetric interval. This seems not to be a good choice unless the distribution of X_{T+k} is close to be Gaussian; see Chatfield (1993). Another alternative is to construct prediction intervals for X_{T+k} by retransforming the upper and lower values of the corresponding prediction interval for Y_{T+k}. Finally, Guerrero (1993) suggests to correct for bias the endpoints of the latter prediction intervals using a procedure similar to the one he proposes for the point forecast.

In this paper, we propose a bootstrap resampling scheme to obtain an estimate of the pdf of X_{T+k} conditional on the available data when an ARIMA model has been fitted to y_t. Given this density, the required prediction intervals for X_{T+k} can be constructed. There are several advantages over the methods previously described. First of all, the bootstrap procedure does not rely on distributional assumptions neither on the transformed data nor on the original scale. The second advantage is that the bootstrap intervals incorporate the variability due to parameter estimation, which is not allowed by any of the alternative procedures. Finally, the method is very easy to implement.

The finite sample behavior of the bootstrap intervals is compared with the alternative intervals by means of an extensive simulation study. It is shown that the proposed procedure performs as well as the best alternatives when Y_t is Gaussian, and tends to outperform to its competitors when leaving this assumption.

The paper is organized as follows. Section 2 presents a description of the existing methods for obtaining prediction intervals for a variable in its original scale. In Section 3 we introduce the bootstrap approach. A Monte Carlo study comparing the finite sample performance of all existing methods is presented in Section 4. In Section 5, we illustrate the procedure analysing empirically three real data sets. Finally, we conclude with some remarks and suggestions for future research in Section 6.

2. PREDICTION INTERVALS FOR TRANSFORMED TIME SERIES

There are two main alternatives proposed in the literature to obtain prediction intervals for X_{T+k} given the observed series $(x_t,\ t=1,\ldots,T)$ after an ARIMA model has been fitted to the power transformed variable Y_t. In this section these two procedures are described.

Consider that $\{x_1,\ldots,x_T\}$ is an available sequence of T observations such that, for any of the reasons previously mentioned, need to be transformed adequately by a function $g(\cdot)$ defined in (2), to obtain a new sequence $\{y_1,\ldots,y_T\}$. Let also assume that the transformed sequence is well fitted by an ARIMA(p,d,q) process given by

$$\nabla^d y_t = \phi_0 + \phi_1 \nabla^d y_{t-1} + ... + \phi_p \nabla^d y_{t-p} + a_t + \theta_1 a_{t-1} + ... + \theta_q a_{t-q}, \tag{3}$$

where a_t is a white noise process, ∇ is the difference operator such that $\nabla y_t = y_t - y_{t-1}$ and $\left(\phi_0, \phi_1, ..., \phi_p, \theta_1, ..., \theta_q\right)$ are unknown parameters. From the transformed series $\{y_1, y_2, ..., y_T\}$, these parameters can be estimated by a consistent estimator, for example, conditional quasi-maximum likelihood (QML). Given the estimates $(\widehat{\phi}_0, \widehat{\phi}_1, \ldots, \widehat{\phi}_p, \widehat{\theta}_1, \ldots, \widehat{\theta}_q)$, the residuals are calculated by the following recursion

$$\widehat{a}_t = \nabla^d y_t - \widehat{\phi}_0 - \widehat{\phi}_1 \nabla^d y_{t-1} - ... - \widehat{\phi}_p \nabla^d y_{t-p} - \widehat{\theta}_1 \widehat{a}_{t-1} - ... - \widehat{\theta}_q \widehat{a}_{t-q}, t = p+d+1, ..., T, \tag{4}$$

where the residuals corresponding to periods of time $t = 0, -1, -2, ...$ are set equal to zero.

Once the ARIMA model has been estimated, the optimal linear predictor of Y_{T+k}, denoted by $\widehat{Y}_T(k)$, is given by

$$\nabla^d \widehat{Y}_T(k) = \widehat{\phi}_0 + \widehat{\phi}_1 \nabla^d \widehat{Y}_T(k-1) + \cdots + \widehat{\phi}_p \nabla^d \widehat{Y}_T(k-p) + \widehat{a}_{T+k} + \widehat{\theta}_1 \widehat{a}_{T+k-1} + \cdots + \widehat{\theta}_q \widehat{a}_{T+k-q}, \tag{5}$$

where $\widehat{Y}_T(j) = Y_{T+j}$ for $j \leq 0$ and $\widehat{a}_{T+j} = 0$ for $j \geq 0$. The usual Box and Jenkins

(1976) prediction intervals for Y_{T+k} are given by

$$\left[\widehat{Y}_T(k) - z_{\alpha/2}\left(\widehat{\sigma}_a^2 \sum_{j=0}^{k-1} \widehat{\Psi}_j^2\right)^{1/2},\ \widehat{Y}_T(k) + z_{\alpha/2}\left(\widehat{\sigma}_a^2 \sum_{j=0}^{k-1} \widehat{\Psi}_j^2\right)^{1/2}\right], \tag{6}$$

where $z_{\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution, $\widehat{\sigma}_a^2$ is the usual estimate of the innovations variance and $\widehat{\Psi}_j$ are the estimated coefficients of the moving average representation.

2.1 Symmetric prediction intervals

Multistep symmetric prediction intervals have been widely used in linear time series models. These intervals are constructed under the assumption of normality for the variable of interest. Therefore, they provide a reasonable performance in terms of coverage and length if the density of the forecast error is well approximated by the normal distribution.

To obtain a symmetric prediction interval k periods into the future, it is needed first a point forecast $\widehat{X}_T(k)$ for X_{T+k}, usually corrected by bias using one of the methods previously mentioned to compute the debiasing factor, and secondly, an explicit expression for the k-step ahead conditional mean squared error, say $V_c(k)$. Then, given the assumption of normality, it follows that the conditional distribution of X_{T+k} given the available data is normal with mean $\widehat{X}_T(k)$ and mean squared error $V_c(k)$. In such a case, the k-step ahead prediction interval is given by

$$\left[\widehat{X}_T(k) - z_{\frac{\alpha}{2}} V_c(k)^{\frac{1}{2}},\ \widehat{X}_T(k) + z_{\frac{\alpha}{2}} V_c(k)^{\frac{1}{2}}\right]. \tag{7}$$

The expression of $V_c(k)$, given by Granger and Newbold (1976), is very dependent on the Gaussian assumption for the series Y_t. Furthermore, this expression is derived by using Hermite polynomials and, it is not easy to obtain for a general transformation $g(\cdot)$. In fact, Granger and Newbold (1976) only give the final expression of $V_c(k)$

for the logarithmic and square root transformations. For example, if the logarithmic transformation is considered, $V_c(k)$ is given by $\exp\left\{2\widehat{Y}_T(k)+\widehat{\sigma}^2(k)\right\}\left[\exp\left\{\widehat{\sigma}^2(k)\right\}-1\right]$, where $\widehat{\sigma}^2(k)=\widehat{\sigma}_a^2\sum_{j=0}^{k-1}\widehat{\Psi}_j^2$.

Furthermore, the prediction intervals in (7) ignore the skewness and all higher moments in the distribution of the forecast error by assuming that is approximately normal, and therefore, will only be accurate if the corresponding forecast error is approximately normally distributed.

Notice that usually Y_t is assumed to be normally distributed, and consequently, the untransformed variable X_t will be non-normally distributed unless the parameter λ in the Box-Cox transformation is equal to 1, i.e. X_t is not transformed.

2.2. Naive prediction intervals

Alternatively, prediction intervals for the variable in the original scale can be constructed by retransforming the upper and lower values of the corresponding prediction intervals for the transformed variable Y_t given by (6). If the prediction interval for Y_t has a prescribed probability, say (1-α), then the retransformed prediction interval for X_t should have the same prescribed probability; see Harvey (1989).

The corresponding prediction interval with nominal coverage of $1-\alpha$ is given by

$$\left[g^{-1}\left\{\widehat{Y}_T(k)-z_{\frac{\alpha}{2}}\widehat{\sigma}(k)\right\},\ g^{-1}\left\{\widehat{Y}_T(k)+z_{\frac{\alpha}{2}}\widehat{\sigma}(k)\right\}\right]. \tag{8}$$

Additionally, as proposed by Guerrero (1993), it is possible to correct for bias the previous confidence interval, by multiplying the end points of (8) by the following debiasing factor

$$\begin{aligned} C_\lambda(k) &= \left\{0.5+0.5\left[1+2\left(\lambda^{-1}-1\right)\sigma^2(k)/\widehat{Y}_T^2(k)\right]^{1/2}\right\}^{1/\lambda},\ \text{for } \lambda\neq 0 \\ &= \exp\left(\sigma^2(k)/2\right),\ \text{for } \lambda=0. \end{aligned} \tag{9}$$

Notice that the prediction intervals in (8) are able to cope with the potential

asymmetry of the distribution of X_t, although they still rely on the Gaussianity assumption for the transformed variable, Y_t, and do not incorporate the uncertainty due to parameter estimation.

3. THE BOOTSTRAP APPROACH

In this section we describe a bootstrap procedure to obtain prediction densities and prediction intervals of future values of the series of interest X_t. The resampling scheme is based on the proposal by Pascual *et al.* (1998) to estimate prediction densities and intervals of series generated by ARIMA(p,d,q) processes.

Denote by $\widehat{F}_a$ the empirical distribution function of the centered residuals of the ARIMA model for y_t given in (4). Given a set of $p+d$ initial values of the variable y_t, say $\{y_1, \ldots, y_{p+d}\}$, a bootstrap replicate of the transformed series $\{y_1^*, \ldots, y_T^*\}$ is constructed by the following equation

$$\nabla^d y_t^* = \widehat{\phi}_0 + \sum_{j=1}^{p} \widehat{\phi}_j \nabla^d y_{t-j}^* + \sum_{j=1}^{q} \widehat{\theta}_j \widehat{a}_{t-j}^* + \widehat{a}_t^*, \ t = p+d+1, \ldots, T, \tag{10}$$

where $y_t^* = y_t$, $t = 1, ..., p+d$ and $\widehat{a}_{1+p+d-q}^*, \ldots, \widehat{a}_T^*$ are random draws from $\widehat{F}_a$. Once the parameters of this bootstrap series are estimated, say $(\widehat{\phi}_0^*, \widehat{\phi}_1^*, \ldots, \widehat{\phi}_p^*, \widehat{\theta}_1^*, \ldots, \widehat{\theta}_q^*)$, the bootstrap forecast k steps ahead is obtained as follows,

$$\nabla^d y_{T+k}^* = \widehat{\phi}_0^* + \sum_{j=1}^{p} \widehat{\phi}_j^* \nabla^d y_{T+k-j}^* + \sum_{j=1}^{q} \widehat{\theta}_j^* \widehat{a}_{T+k-j}^* + \widehat{a}_{T+k}^*, \ k = 1, 2, ... \tag{11}$$

where $y_{T+k-j}^* = y_{T+k-j}$, $j \geqslant k$, and $\widehat{a}_{T+k-j}^* = \widehat{a}_{T+k-j}$, $j \geqslant k$, i.e., the last $p+d$ observations of the series and the last q residuals are fixed in order to obtain the prediction density conditional on the observed data. Finally, in expression (11), $\widehat{a}_{T+k-j}^*$, $j < k$, are random draws from $\widehat{F}_a$.

Once B bootstrap replicates of Y_{T+k}^* are obtained, it is possible to construct a bootstrap estimate of the distribution of Y_{T+k} conditional on $\{y_1, \ldots, y_T\}$ and the

corresponding prediction intervals. Pascual *et al.* (1998) prove that for the transformed series $\{y_t\}$, $Y^*_{T+k} \rightarrow Y_{T+k}$ in conditional probability, in probability, as the sample size T goes to infinity. They also show that the finite sample properties of the bootstrap procedure just described outperforms other alternative bootstrap mechanisms proposed to compute prediction intervals in stationary AR(p) models.

However, the objective is to estimate the distribution of X_{T+k} conditional on $\{x_1, \ldots, x_T\}$. In this case, a new step has to be introduced in the described procedure. The bootstrap forecast k steps ahead for the variable in the original metric is then obtained as

$$x^*_{T+k} = g^{-1}\left(y^*_{T+k}\right), \ k = 1, 2, ... \tag{12}$$

This procedure is repeated B times to obtain a set of B bootstrap replicates for X_{T+k}, say $\left(x^{*(1)}_{T+k}, \ldots, x^{*(B)}_{T+k}\right)$. Then, the prediction limits are defined as the quantiles of the bootstrap distribution function of X^*_{T+k}, i.e., if $G^*(h) = \Pr(X^*_{T+k} \leq h)$ is the distribution function of X^*_{T+k} and its Monte Carlo estimate is $G^*_B(h) = \#(x^{*(b)}_{T+k} \leq h)/B$, a $100\alpha\%$ prediction interval for X^*_{T+k} is given by

$$[L^*_B, U^*_B] = \left[Q^*_B\left(\frac{1-\alpha}{2}\right), Q^*_B\left(\frac{1+\alpha}{2}\right)\right], \tag{13}$$

where $Q^*_B = G^{*-1}_B$.

Before summarizing the steps for obtaining bootstrap prediction densities and intervals for X_{T+k}, we illustrate the method with a simple example. Suppose that after taking an adequate power-transformation, the sequence $\{y_1, \ldots, y_T\}$ follows an ARIMA(0,1,2) model without constant term

$$\nabla y_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}. \tag{14}$$

Once the parameters of model (14) have been estimated and the bootstrap draws $\widehat{a}^*_0, \widehat{a}^*_1, \ldots, \widehat{a}^*_T$ are available, a bootstrap replicate of the transformed series is constructed by

$$y^*_t = y^*_{t-1} + \widehat{a}^*_t + \widehat{\theta}_1 \widehat{a}^*_{t-1} + \widehat{\theta}_2 \widehat{a}^*_{t-2}, \ t = 2, ..., T, \tag{15}$$

where $y_1^* = y_1$. Then, bootstrap estimates $\widehat{\theta}_1^*$ and $\widehat{\theta}_2^*$ are obtained for the bootstrap series and bootstrap replicates of future values of the transformed series are generated by

$$\begin{aligned} y_{T+1}^* &= y_T + \widehat{a}_{T+1}^* + \widehat{\theta}_1^* \widehat{a}_T + \widehat{\theta}_2^* \widehat{a}_{T-1} \\ y_{T+2}^* &= y_{T+1}^* + \widehat{a}_{T+2}^* + \widehat{\theta}_1^* \widehat{a}_{T+1}^* + \widehat{\theta}_2^* \widehat{a}_T \\ y_{T+3}^* &= y_{T+2}^* + \widehat{a}_{T+3}^* + \widehat{\theta}_1^* \widehat{a}_{T+2}^* + \widehat{\theta}_2^* \widehat{a}_{T+1}^*. \end{aligned}$$

It is important to note that since the predictions are conditional on the sample information available at time T, in the recursions above $\widehat{a}_T$ and $\widehat{a}_{T-1}$ are kept fixed in the different bootstrap replicates of y_{T+1}^* and y_{T+2}^* while $\widehat{a}_{T+1}^*$, $\widehat{a}_{T+2}^*$ and $\widehat{a}_{T+3}^*$ change from one to another replicate. Finally, bootstrap replicates of future values of the series in the original scale are generated by expression (12).

Now, we summarize all the steps needed for obtaining bootstrap prediction intervals for X_{T+k}.

Step 1. Compute the residuals $\widehat{a}_t$ as in (4) for the transformed series. Let $\widehat{F}_a$ be the empirical distribution function of the centered residuals.

Step 2. Generate a bootstrap series using the recursion in (10) and calculate the estimates $(\widehat{\phi}_0^*, \widehat{\phi}_1^*, \ldots, \widehat{\phi}_p^*, \widehat{\theta}_1^*, \ldots, \widehat{\theta}_q^*)$.

Step 3. Obtain a bootstrap future value for the transformed series by expression (11). Note that the last $p+d$ values of the transformed series and the final q residuals are fixed in this step but not in the previous one.

Step 4. Obtain a bootstrap future value for the series in the original scale by expression (12).

Step 5. Repeat the last four steps B times and then go to Step 6.

Step 6. The endpoints of the prediction interval are given by quantiles of G_B^*, the bootstrap distribution function of X_{T+k}^*, given by expression (13).

Alternatively, the bootstrap procedure just described could be also applied to con-

struct prediction intervals conditional on the parameter estimates; hereafter CB (conditional bootstrap). This procedure has been previously proposed by Cao *et al.* (1997) for series $\{x_1, \ldots, x_T\}$ following an AR(p) processes, and has been generalized by Pascual *et al.* (2001) for the general ARIMA(p,d,q) processes. With this method, the parameters are estimated once and these estimates are used in the calculation of all bootstrap forecasts x^*_{T+k}. The steps to obtain bootstrap forecasts are similar to those presented above except that *Step-2* is avoided since now it is not necessary to generate bootstrap replicates of the transformed series. Then, the expression to obtain bootstrap future values for the transformed series in *Step-3*, is replaced by

$$\nabla^d y^*_{T+k} = \widehat{\phi}_0 + \sum_{j=1}^{p} \widehat{\phi}_j \nabla^d y^*_{T+k-j} + \sum_{j=1}^{q} \widehat{\theta}_j \widehat{a}^*_{T+k-j} + \widehat{a}^*_{T+k}, k = 1, 2, ...,$$

where y^*_{T+k-j} and $\widehat{a}^*_{T+k-j}$ are defined as in (11). Since the parameter estimates are kept fixed in all bootstrap replicates of future values, the CB prediction intervals do not incorporate the uncertainty due to parameter estimation.

Notice that the estimated bootstrap density of X_{T+k} can also be used to obtain a bootstrap estimate of the expected value and/or the median of X_{T+k} conditional on the available series. These estimates can then be taken as point forecasts of X_{T+k}.

Finally, using the asymptotic results in Pascual *et al.* (1998) and since $g(\cdot)$ is a known continuous invertible function, it is straightforward to prove using the bootstrap version of the Continuity Theorem, that $g^{-1}\left(Y^*_{T+k}\right) \rightarrow g^{-1}\left(Y_{T+k}\right)$, i.e., $X^*_{T+k} \rightarrow X_{T+k}$, in conditional probability, in probability, as $T \rightarrow \infty$.

4. SMALL SAMPLE PROPERTIES

4.1 Monte Carlo design

We now describe the results of several Monte Carlo experiments carried out to study the small-sample performance of the prediction intervals built by the alterna-

tives considered in this paper. Prediction intervals built by the proposed bootstrap procedure (PRR) are compared with CB intervals and with the non-bootstrap methods described in Section 2. As previously mentioned, PRR is the only approach that does not condition on parameter estimates and then, introduces the variability due to parameter estimation in the intervals. Comparing PRR with CB intervals, we are studying the effect of parameter estimation variability on the shape of estimated prediction densities. The basic symmetric prediction intervals in (7) will be denoted hereafter by STD1, the intervals based on retransforming the ends in (8) will be denoted by STD2, and finally, the corrected by bias prediction intervals using (9) by STD3.

The focus of the simulation experiments is on prediction of future values of a series x_t, such that a linear ARIMA(p,d,q) model is fitted to a power transformation of its original values, say y_t. We consider the following ARIMA processes,

$$y_t = 0.95y_{t-1} + a_t \tag{16}$$

$$y_t = 1.75y_{t-1} - 0.76y_{t-2} + a_t \tag{17}$$

and

$$y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}. \tag{18}$$

The first two models considered are pure autoregressive with orders one and two respectively, and the third one includes a moving average component. The AR(1) model was chosen because the autoregressive polynomial has a root close to the non-stationarity region. The AR(2) model was selected because it was one of the models used by Thombs and Schucany (1990) in their seminal paper on using the bootstrap to approximate prediction densities for AR(p) processes. Finally, the ARMA(1,1) model was chosen to analyse the finite sample properties of the proposed procedure in models with moving average components.

For each model considered, we generate artificial series with several choices of error distributions, in particular, Gaussian, Student-t with 5 degrees of freedom, and two asymmetric distributions, exponential (exponential^{+}) and minus exponential (exponential^{-}) respectively. In all cases, we have centered the errors to have zero mean. With respect to the variance of the simulated errors, its value was chosen to have reasonable values of the original series x_t when the inverse transform was taken to the y_t series. These values are σ_a^2 =0.1, 0.01 and 0.5 for the AR(1), AR(2) and ARMA(1,1) models respectively. Note that the coverage properties are exactly the same whichever the value of the variance, and the only difference appears in the lengths of the prediction intervals.

We only report the results obtained for the logarithmic transformation, i.e. $y_t = \log(y_t)$. It is important to note that the conclusions with other power transformations and models considered are the same and, therefore, are not reported in this paper to save space.

All the models for the log-transformed series are estimated by conditional QML. In all cases, the sample sizes considered are 50 and 100. The prediction horizons under study are k=1, 2 and 3, and the corresponding intervals are constructed with a nominal coverage 1-α equal to 0.80, 0.95 and 0.99. For each particular series generated by any of the models considered, with a particular sample size and error distribution F_a, we generated R=1000 future values of x_{T+k} from that series and obtain $100(1-\alpha)\%$ prediction intervals, denoted by (L_X, U_X) by each of the five procedures considered. PRR and CB prediction intervals are constructed based on B=999 bootstrap replicates. The conditional coverage of each procedure is computed by

$$\widehat{\alpha}_X^* = \#\left\{L_X^* \leq x_{T+k}^r \leq U_X^*\right\}/R,$$

where x_{T+k}^r ($r = 1,\ldots,R$) are future values of the variable generated previously.

The Monte Carlo results are based on 1000 replicates. Prediction intervals are

compared in terms of average coverage and length, and the proportion of observations lying out to the left and to the right through all Monte Carlo replicates.

Computations have been carried out in a HP-UX C360 workstation, using Fortran 77 and the corresponding subroutines of Numerical Recipes by Press *et al.* (1986).

4.2 Results of Monte Carlo experiments

The results of the Monte Carlo experiments for model (16) with Gaussian innovations appear in table 1. First of all, as measured by interval content, the STD1 intervals appear to be about as accurate as STD2 and PRR at the 95% level. However, the interval content is somewhat misleading, because STD1 generate rather biased one-sided prediction intervals; see the average of observations lying out to the left and to the right. The results for the 80% prediction intervals are reported in table 2, where it can be observed that the symmetric intervals have even a worse performance, since in all cases the average coverage is over the nominal values. Additionally, notice that the accuracy of the STD1 intervals does not improve with sample size. Therefore, the symmetric STD1 intervals seem to be not adequate to predict future values of transformed variables. This effect was also observed by Collins (1991) in the context of regression models. As an illustration, figure 1 plots the prediction density corresponding to the STD1 intervals together with the empirical density. It is rather obvious that the symmetric density is not adequated to represent the empirical density of X_{T+k}. Furthermore, notice that in tables 1 and 2 the average length of the STD1 intervals is systematically larger than the empirical length.

Next, analyzing the behavior of the intervals based on (8) and (9) in tables 1 and 2, it is interesting to note that the use of the bias corrected STD3 intervals do not improve in any case the results of the STD2 intervals. They have larger average length than STD2 and the average observations left on the right and on the left is clearly asymmetric. This means that using the debiasing factor (designed to obtain

a better estimation of the conditional expectation in the original scale) for correcting the bias of the prediction intervals does not seem to work for the cases considered in this paper.

Comparing PRR and STD2 intervals, it is possible to observe that they have similar performance in terms of both average coverage and length. The reason for the good behavior of STD2 seems clear. In this case when taking logarithms of the original observations, the resulting transformed series has normal errors, and therefore, the usual Box and Jenkins (1976) and the PRR intervals have similar performance; see Pascual *et al.* (1998). Consequently, when going back to the original metric this similar behavior remains. In figure 1, we also plot the density of X_{T+k} corresponding to retransforming Y_{T+k} as is done when constructing the STD2 intervals. Notice that, although this density is closer to the empirical density than the one based on STD1 intervals, the shape is still slightly different.

Finally, we concentrate on the comparison of PRR with respect to CB which does not incorporate the parameter uncertainty variability. The results reported in table 1 show that CB intervals have lower average coverage than PRR, the latter having average coverage closer to the nominal value. Therefore, it seems to be important to include the uncertainty due to parameter estimation in prediction intervals in order to obtain coverages close to the nominal values. The necessity of using PRR is more evident for small sample sizes. As expected, since the conditional QML estimator is consistent, CB and PRR intervals get closer in terms of coverage and length as the sample size increases. The conclusions are essentially the same for predictions made one, two and three steps ahead. In figure 1, it is rather clear that the PRR prediction density is closer to the empirical density than the CB and STD2 densities. The densities plotted in figure 1 also show that the STD2 and CB densities are rather close. This could be due to the fact that, in this case, the improvement of PRR over STD2 intervals is not due to the distribution of the forecast errors but to the inclusion

of the variability due to parameter estimation.

Table 3 reports the Monte Carlo results for the 80% prediction intervals for log-transformed series generated by the AR(1) model with innovations generated by a Student-t with 5 degrees of freedom. The conclusions with respect the STD1, STD3 and CB intervals are the same as before. In this table it is also possible to observe the improvement of PRR with respect to STD2 intervals. In this case, the average coverage and lengths of the STD2 intervals are larger than nominal values, and what is even more important, this bad behavior does not improve as the sample size increases. Remember that the STD2 intervals are built assuming that the transformed variable Y_t is normal. Therefore, as soon as this assumption is not satisfied, the intervals do not have the usual properties. Figure 2 illustrates the results for a particular series generated by model (16) with Student-5 innovations with T=100. The conclusions from this picture are as previously. The PRR density is closer to the empirical density than any of the others. The symmetric density is clearly inadequated and the STD2 and CB densities are very similar.

Table 4 reports the results for the 80% prediction intervals for log-transformed series generated by the AR(2) model with exponential$^+$ innovations. The non-bootstrap methods have in general average coverage and length over nominal values and, as the sample size increases this bad behavior tends to be even worse. They are not able to cope with the asymmetry of the transformed series Y_t. Notice that the PRR intervals have an adequated performance and additionally, as the sample size gets bigger, its average coverage and, as expected, length measures get closer to nominal values, given the asymptotic properties mentioned in section 3. Also, notice that in this case the necessity of introducing the variability due to parameter estimation by the use of PRR is crucial even for large sample sizes. For this asymmetric to the right distribution, CB intervals have lower average coverage than PRR. Figure 3 plots kernel estimates of the densities obtained for a particular replicate in this case. It is

possible to observe that the non-bootstrap densities do not resemble the shape of the empirical density while the PRR density is able to mimic it.

Finally, table 5 shows the results obtained when generating series by the ARMA(1,1) model in (18) with exponential$^-$ errors for the 95% prediction intervals. Once more, we can observe that the STD1 intervals are not able to cope with the asymmetric shape of the density of X_{T+k}. Notice that in this case, the STD2 intervals have average lengths too large when compared with the empirical length and that this problem is still severe for samples as large as 100 observations. Figure 4 plots the kernel densities obtained for a particular replicate generated by this model with exponential$^-$ innovations and T=100. The conclusions are the same as in the previous pictures, and again, PRR density is the closest to the empirical one. Therefore, the proposed procedure seems to behave properly in models with moving average components.

Summarizing, PRR intervals perform as well as STD2 intervals when the innovations of the transformed data are well approximated by a normal distribution and, outperform the existing procedures when this distribution differs from the Gaussian one, a situation frequently found when working with real data. Furthermore, the symmetric intervals based on (7) are shown to have poor properties even when the transformed data are Gaussian. The bias correcting factor for the end of the prediction intervals in (8) proposed by Guerrero (1993) is also shown not to improve the properties of the non corrected intervals. Finally, we have shown that including the uncertainty due to estimation of the parameters of the model in the bootstrap prediction intervals may be crucial depending on the distribution of the transformed data.

5. REAL DATA APPLICATIONS

In this section, we illustrate empirically the use of the suggested bootstrap method to construct prediction intervals for transformed variables. We start considering the

Sales Data, studied firstly by Chatfield and Prothero (1973) and latter by Pankratz and Dudley (1987) and Guerrero (1993) among others. The series, plotted in figure 5a, consists on 77 observations of the monthly sales of an engineering product with a marked trend and a strong seasonal pattern. Since the size of the seasonal effect increases with the mean level of sales, Chatfield and Prothero used originally the log-transformation; however, this transformation was criticized by Wilson (1973) who found by maximum likelihood that a more convenient power transformation was $\widehat{\lambda} = 0.34$. This found was latter supported by Guerrero (1993). Therefore, we will consider $\lambda = \frac{1}{3}$ as known. Figure 5b represents the transformed observations. The model finally fitted to the transformed data is

$$(1 - \phi B)\, \Delta\Delta_{12} y_t = \left(1 - \Theta B^{12}\right) a_t \tag{19}$$

where $y_t = x_t^{1/3}$, x_t denotes the original series and B is the backshift operator. The first 65 observations of the series, corresponding to the period from January 1965 up to May 1970, are used to estimate the parameters of model (19). The last 12 observations, corresponding to the period from June 1970 up to May 1971, are used to assess the predictive performance of the STD2, STD3 and PRR prediction intervals. Note that for this particular transformation STD1 prediction intervals can not be computed since no formulas for the variance of the prediction error of x_t are available.

The QML estimates are $\widehat{\phi} = -0.5437$ and $\widehat{\Theta} = 0.5466$. A kernel estimate of the residuals density together with the normal density appear in Figure 6. The empirical distribution of the residuals has a long tail to the left. The skewness coefficient is -0.2946, and the excess kurtosis is 0.065, with the former significantly different from zero.

Then, we implement the procedure proposed by Pascual *et al.* (1998) to construct prediction densities of the transformed variable y_{65+k} for $k = 1, \ldots, 12$. The estimated density for lead time 1 together with the normal density, appears in figure 7. In this

figure, it can be observed that the bootstrap density has the same asymmetry to the left observed in the residuals distribution. Additionally, using the estimated bootstrap densities we construct prediction intervals for futures values of the variable in the transformed scale.

Finally, we implement the new procedure to construct prediction densities and prediction intervals for future values of the variable in the original scale. Figure 8 plots the prediction intervals for X_{T+k} constructed using the bootstrap procedure and the retransformed ends. Notice that the bootstrap intervals are able to capture the asymmetry inherent in the prediction densities, and have lower length than the standard intervals in almost all forecast horizons. In table 6 that reports the interval lengths for some selected horizons, it can be observed that the bootstrap intervals are always thinner than the STD2 intervals.

In this case, using the mean or the median of the bootstrap density does not improve the mean squared prediction error over the retransformed point predictions.

Next, we analyze two economic time series, the U.S. dollar-pound real exchange rate (RXR) and the ratio of nonborrowed of total reserves (NBRX). These series are studied by Kilian (1998) in the context of VAR models, where can be seen that the residuals clearly reject the normality assumption. Both series consists on 197 observations, where the first 173 are used to estimate the parameters of the ARIMA model fitted to the log-transformed data, and the last 24 observations are used to asses the predictive performance of the methods considered in this section.

The final model fitted to the log-RXR series is

$$\Delta y_t = (1 + 0.401B)\,\widehat{a}_t \tag{20}$$

with $y_t = \log(x_t)$ and x_t been the original series. Figure 9 shows a kernel estimate of the residuals density together with the normal density. The skewness coefficient is 0.2112, and the excess kurtosis is 0.37, with the former different from zero. Figure

10 shows PRR and STD2 prediction intervals at 80%. It is clear how at the 80% level, the bootstrap intervals have a better behavior in terms of coverage since only one observation lies out of the bootstrap limits but three observations lie out of the STD2 intervals.

Finally, the model fitted to the log-NBRX data is

$$y_t = -0.031 + 0.8481 y_{t-1} \tag{21}$$

with $y_t = \log(x_t)$, where x_t is the original series. In figure 11 is shown a kernel estimate of the residuals density together with the normal density. The skewness coefficient is -0.9071, and the excess kurtosis is 6.50, both significantly different from zero. Therefore, the usual assumption of normality is clearly rejected. In figure 12 can be seen again how the bootstrap prediction intervals capture the asymmetry and kurtosis inherent in the residuals, and consequently in the prediction densities. It is important to note that the length of the bootstrap intervals are shorter than the STD2 ones.

6. SUMMARY AND CONCLUSIONS

This paper extends the bootstrap technique proposed by Pascual *et al.* (1998) to construct prediction intervals for a variable after a linear ARIMA model is fitted to a power transformation of it. In this situation, there is no generally accepted method of computing prediction intervals. The proposed resampling scheme does not assume any distribution for the errors neither in the original nor in the transformed metrics and, at the same time, allows to incorporate the variability due to parameter estimation. By means of Monte Carlo experiments, we compare the finite sample performance of alternative methods previously proposed in the literature to construct prediction intervals for power-transformed series with the bootstrap approach proposed in this paper. There are two main alternatives. The first one based

on Granger and Newbold (1976) assumes a symmetric distribution for both the original and the transformed variable and can only be implemented for logarithmic and root squared transformations. These prediction intervals ignore the skewness and all higher moments of the variable of interest. As a result, this approach will generate biased one-sided prediction intervals.

The second alternative is based on retransforming the ends of the prediction intervals for the transformed variable. In this case, only Gaussianity of the transformed variable is needed. None of these intervals are able to take into account the uncertainty due to estimation of the parameters. The intervals constructed simply by retransforming the upper and lower values of the usual prediction intervals for the transformed series have only good properties when the transformed series has normal errors. In this case, the usual Box-Jenkins intervals for the variable in the transformed metric have very good properties, and then, when going back to the original scale, this good behavior remains. The results show that for nonnormal innovations, these prediction intervals can be heavily distorted. The bias-correction proposed by Guerrero (1993) does not improve the results for the cases considered in this paper. The bootstrap intervals seem to have appropriate properties.

We also analyze how coverage and length of prediction intervals are affected by not taking into account the variability due to parameter estimation. We show that the average coverage of the intervals is closer to the nominal value when intervals are constructed incorporating parameter uncertainty. As expected, since we are considering consistent estimators, the effects of parameter estimation are particularly important for small sample sizes. Furthermore, these effects are more important when the error distribution is not Gaussian; see also Pascual *et al.* (2001).

To conclude, the bootstrap approach presented in this paper seems to have reasonable properties when prediction intervals are required for a variable after a power transformation is taken to its original values. This approach gives prediction intervals

with a reasonable finite sample performance in terms of average coverage and average length with both normal and nonnormal distributions of the innovations. Additionally, this method not only gives prediction intervals but also provides estimates of the prediction density function of the variable in its original scale. As expected, its behavior improves as the sample size increases.

Finally, the behavior of the PRR technique is illustrated with the analysis of three real time series. It is shown that the PRR intervals are shorter than the retransformed intervals having better coverage properties.

ACKNOWLEDGMENTS

The authors are very grateful to Lutz Kilian for providing the RXR and NBRX series. Financial support was provided by projects DGES PB96-0111 and DGICYT PB98-0026 from the Spanish Government and Catedra de Calidad from BBV.

REFERENCES

Box GEP, Cox DR. 1964. An analysis of transformations. *Journal of the Royal Statistical Society* **B-26**: 211-243.

Cao R, Febrero-Bande M, González-Manteiga W, Prada-Sánchez JM, García-Jurado I. 1997. Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes. *Communications in Statistics, Simulation and Computation* **26**: 961-978.

Chatfield C. 1993. Calculating interval forecasts. *Journal of Business & Economic Statistics* **11**: 121-135.

Chatfield C, Prothero DL. 1973. Box-Jenkins seasonal forecasting: problems in a case study. *Journal of the Royal Statistical Society* **A-136:** 295-315.

Collins S. 1991. Prediction techniques for Box-Cox regression models. *Journal of Business & Economic Statistics* **9:** 267-277.

Granger CWJ, Newbold P. 1976. Forecasting transformed series. *Journal of the Royal Statistical Society* **B-38:** 189-203.

Guerrero VM. 1993. Time-series analysis supported by power transformations. *Journal of Forecasting* **12:** 37-48.

Harvey AC. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press: Cambridge.

Kilian L. 1998. Confidence intervals for impulse responses under departures from normality. *Econometric Reviews* **17:** 1-29.

Pankratz A, Dudley U. 1987. Forecasts of power-transformed series. *Journal of Forecasting* **6**: 239-248.

Pascual L, Romo J, Ruiz E. 1998. Bootstrap predictive inference for ARIMA processes. Working Paper 98-86, Universidad Carlos III de Madrid: Spain.

Pascual L, Romo J, Ruiz E. 2001. Effects of parameter estimation on prediction densities: a bootstrap approach. *International Journal of Forecasting* **17**(1): 83-103.

Taylor JMG. 1986. The retransformed mean after a fitted power transformation. *Journal of the American Statistical Association* **81:** 114-118.

Thombs LA, Schucany WR. 1990. Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association* **85**: 486-492.

Tukey JW. 1957. On the comparative anatomy of transformations. *Annals of Mathematical Statistics* **28:** 602-632.

Wilson GT. Discussion of a paper by Dr Chatfield and Dr Prothero. *Journal of the Royal Statistical Society* **A-136**: 315-319.

Authors' biographies:

Lorenzo Pascual is Ph.D. student at the Departamento de Estadística y Econometría of Universidad Carlos III de Madrid. He works on bootstrap methods for prediction in time series models.

Juan Romo received his Ph.D. in Mathematics from Texas A&M University. He has been Associate Professor both at Universidad Complutense and Universidad Carlos III de Madrid, where he is Professor of Statistics and Operations Research.

Esther Ruiz received her Ph.D. in Statistics from the London School of Economics in 1992. She is Associate Professor of Economics at Universidad Carlos III de Madrid.

Table 1

Monte Carlo results for AR(1) model with Gaussian errors.

Lead time	Sample size	Method	Average coverage	Coverage below/above	Average length
1	**n**	**Empirical**	**95%**	**2.5%/2.5%**	**2.04**
	50	STD1	94.96(.03)	.33/4.71	2.12(2.81)
		STD2	94.23(.03)	2.90/2.86	2.09(2.77)
		STD3	93.94(.03)	4.00/2.06	2.20(2.93)
		CB	92.43(.04)	3.77/3.80	2.02(2.74)
		PRR	93.45(.03)	3.32/3.23	2.06(2.76)
	100	STD1	95.24(.02)	.22/4.54	2.09(2.68)
		STD2	94.63(.02)	2.67/2.70	2.06(2.65)
		STD3	94.34(.02)	3.76/1.90	2.18(2.79)
		CB	93.83(.03)	3.08/3.09	2.05(2.71)
		PRR	94.10(.02)	2.92/2.98	2.04(2.62)
2	**n**	**Empirical**	**95%**	**2.5%/2.5%**	**2.86**
	50	STD1	94.54(.03)	.07/5.39	3.04(4.26)
		STD2	93.65(.03)	3.18/3.17	2.97(4.14)
		STD3	93.10(.03)	4.84/2.05	3.28(4.62)
		CB	92.81(.04)	3.53/3.65	2.89(4.06)
		PRR	93.32(.03)	3.32/3.36	2.87(4.04)
	100	STD1	94.93(.02)	.01/5.06	2.98(3.86)
		STD2	94.33(.02)	2.79/2.87	2.91(3.75)
		STD3	93.83(.02)	4.39/1.78	3.20(4.17)
		CB	93.91(.02)	2.93/3.16	2.85(3.68)
		PRR	94.04(.02)	2.89/3.06	2.82(3.51)
3	**n**	**Empirical**	**95%**	**2.5%/2.5%**	**3.44**
	50	STD1	94.15(.04)	.04/5.80	3.80(5.64)
		STD2	93.11(.04)	3.47/3.42	3.65(5.38)
		STD3	92.38(.04)	5.54/2.08	4.21(6.34)
		CB	92.31(.04)	3.79/3.90	3.55(5.31)
		PRR	92.87(.04)	3.56/3.57	3.50(5.18)
	100	STD1	94.69(.03)	.00/5.31	3.68(4.82)
		STD2	94.05(.03)	2.93/3.01	3.54(4.62)
		STD3	93.34(.03)	4.91/1.74	4.07(5.38)
		CB	93.65(.03)	3.07/3.28	3.48(4.58)
		PRR	93.78(.03)	3.03/3.19	3.40(4.32)

Quantities in parenthesis are standard deviations.

Table 2

Monte Carlo results for AR(1) model with Gaussian errors.

Lead time	Sample size	Method	Average coverage	Coverage below/above	Average length
1	**n**	**Empirical**	**80%**	**10%/10%**	**1.29**
	50	STD1	82.29(.05)	6.79/10.92	1.38(1.84)
		STD2	78.98(.05)	10.53/10.49	1.31(1.74)
		STD3	78.43(.05)	13.57/8.00	1.38(1.84)
		CB	77.98(.06)	11.04/10.98	1.29(1.67)
		PRR	78.97(.05)	10.58/10.45	1.30(1.70)
	100	STD1	82.83(.04)	6.44/10.73	1.37(1.75)
		STD2	79.51(.04)	10.18/10.30	1.30(1.66)
		STD3	78.98(.04)	13.23/7.79	1.37(1.75)
		CB	79.01(.04)	10.36/10.62	1.29(1.66)
		PRR	79.46(.04)	10.18/10.36	1.29(1.62)
2	**n**	**Empirical**	**80%**	**10%/10%**	**1.75**
	50	STD1	84.21(.06)	4.73/11.06	1.99(2.78)
		STD2	78.01(.06)	11.03/10.96	1.81(2.50)
		STD3	76.94(.06)	15.36/7.70	1.99(2.79)
		CB	77.23(.06)	11.37/11.39	1.77(2.44)
		PRR	78.06(.06)	11.05/10.89	1.76(2.42)
	100	STD1	85.27(.04)	4.10/10.63	1.95(2.52)
		STD2	79.01(.04)	10.45/10.55	1.77(2.28)
		STD3	78.01(.04)	14.75/7.25	1.95(2.53)
		CB	78.67(.04)	10.59/10.74	1.76(2.27)
		PRR	78.93(.04)	10.48/10.59	1.74(2.18)
3	**n**	**Empirical**	**80%**	**10%/10%**	**2.05**
	50	STD1	85.57(.07)	3.24/11.19	2.48(3.68)
		STD2	77.15(.06)	11.43/11.42	2.16(3.14)
		STD3	75.76(.06)	16.67/7.57	2.49(3.69)
		CB	76.38(.06)	11.81/11.81	2.12(3.06)
		PRR	77.20(.07)	11.46/11.33	2.07(2.99)
	100	STD1	87.05(.05)	2.40/10.55	2.40(3.15)
		STD2	78.59(.04)	10.62/10.79	2.10(2.71)
		STD3	77.11(.04)	15.92/6.98	2.41(3.16)
		CB	78.20(.05)	10.77/10.03	2.08(2.71)
		PRR	78.45(.05)	10.68/10.87	2.04(2.54)

Quantities in parenthesis are standard deviations.

Table 3

Monte Carlo results for AR(1) model with Student-5 errors.

Lead time	Sample size	Method	Average coverage	Coverage below/above	Average length
1	**n**	**Empirical**	**80%**	**10%/10%**	**1.12**
	50	STD1	84.82(.06)	6.01/9.17	1.55(2.75)
		STD2	82.40(.06)	8.76/8.83	1.47(2.58)
		STD3	81.89(.06)	11.15/6.97	1.55(2.76)
		CB	78.06(.06)	10.91/11.03	1.30(2.29)
		PRR	79.15(.06)	10.35/10.50	1.31(2.28)
	100	STD1	85.55(.04)	5.64/8.81	1.34(1.89)
		STD2	83.14(.04)	8.38/8.48	1.28(1.78)
		STD3	82.63(.04)	10.74/6.63	1.35(1.90)
		CB	79.21(.04)	10.23/10.55	1.14(1.56)
		PRR	79.54(.04)	10.15/10.31	1.14(1.59)
2	**n**	**Empirical**	**80%**	**10%/10%**	**1.58**
	50	STD1	84.98(.07)	4.92/10.11	2.28(4.73)
		STD2	80.05(.07)	9.95/10.00	2.05(4.18)
		STD3	79.02(.07)	13.71/7.26	2.29(4.75)
		CB	77.40(.07)	11.21/11.38	1.91(3.90)
		PRR	78.48(.08)	10.69/10.83	1.88(3.96)
	100	STD1	86.18(.05)	4.30/9.52	1.91(2.74)
		STD2	81.23(.05)	9.34/9.43	1.74(2.43)
		STD3	80.25(.04)	13.08/6.68	1.92(2.75)
		CB	78.71(.04)	10.55/10.74	1.62(2.27)
		PRR	79.01(.04)	10.46/10.53	1.60(2.19)
3	**n**	**Empirical**	**80%**	**10%/10%**	**1.89**
	50	STD1	85.52(.08)	3.85/10.63	2.93(7.47)
		STD2	78.51(.08)	10.68/10.81	2.50(6.22)
		STD3	77.05(.08)	15.47/7.48	2.93(7.49)
		CB	76.44(.08)	11.66/11.89	2.38(6.11)
		PRR	77.54(.08)	11.14/11.32	2.31(6.01)
	100	STD1	87.24(.05)	3.01/9.75	2.36(3.43)
		STD2	80.18(.05)	9.86/9.96	2.05(2.88)
		STD3	78.75(.05)	14.59/6.65	2.37(3.45)
		CB	78.36(.05)	10.72/10.91	1.95(2.77)
		PRR	78.38(.05)	10.75/10.87	1.90(2.62)

Quantities in parenthesis are standard deviations.

Table 4

Monte Carlo results for AR(2) model with Exponential$^+$ errors.

Lead time	Sample size	Method	Average coverage	Coverage below/above	Average length
1	**n**	**Empirical**	**80%**	**10%/10%**	**0.75**
	50	STD1	86.49(.07)	1.96/11.54	1.13(5.79)
		STD2	86.33(.07)	2.36/11.31	1.13(5.72)
		STD3	86.29(.08)	2.90/10.81	1.13(5.79)
		CB	75.86(.13)	12.76/11.37	.93(3.95)
		PRR	77.91(.12)	11.35/10.74	1.01(4.81)
	100	STD1	88.43(.04)	.56/11.01	.96(3.87)
		STD2	88.43(.04)	.81/10.76	.96(3.84)
		STD3	88.53(.05)	1.20/10.27	.96(3.87)
		CB	77.48(.11)	11.79/10.71	.84(3.40)
		PRR	78.77(.10)	10.82/10.41	.86(3.47)
2	**n**	**Empirical**	**80%**	**10%/10%**	**1.64**
	50	STD1	83.00(.09)	4.38/12.62	2.44(13.5)
		STD2	81.61(.10)	6.13/12.26	2.35(12.9)
		STD3	80.57(.10)	8.26/11.17	2.44(13.58)
		CB	74.77(.11)	13.31/11.91	2.25(12.9)
		PRR	77.09(.11)	11.75/11.16	2.35(13.5)
	100	STD1	85.87(.05)	2.46/11.67	2.10(9.57)
		STD2	84.54(.06)	4.15/11.30	2.05(9.27)
		STD3	83.41(.07)	6.37/10.22	2.11(9.58)
		CB	77.13(.08)	11.93/10.94	1.91(8.54)
		PRR	78.24(.09)	11.13/10.63	1.95(8.71)
3	**n**	**Empirical**	**80%**	**10%/10%**	**2.60**
	50	STD1	81.54(.10)	5.02/13.44	4.06(24.7)
		STD2	78.40(.11)	8.49/13.11	3.74(21.9)
		STD3	76.60(.12)	12.01/11.38	4.08(24.8)
		CB	73.53(.11)	13.88/12.59	3.80(24.35)
		PRR	76.08(.11)	12.21/11.71	3.76(22.7)
	100	STD1	85.00(.06)	2.91/12.09	3.53(18.0)
		STD2	81.82(.07)	6.40/11.77	3.31(16.6)
		STD3	79.74(.08)	10.22/10.04	3.53(18.04)
		CB	76.65(.08)	12.08/11.26	3.22(16.4)
		PRR	77.67(.08)	11.38/10.94	3.19(15.7)

Quantities in parenthesis are standard deviations.

Table 5

Monte Carlo results for ARMA(1,1) model with Exponential$^-$ errors.

Lead time	Sample size	Method	Average coverage	Coverage below/above	Average length
1	**n**	**Empirical**	**95%**	**2.5%/2.5%**	**1.99**
	50	STD1	99.67(.03)	.06/.27	4.43(2.35)
		STD2	94.10(.03)	5.81/.09	4.02(1.88)
		STD3	91.98(.03)	8.00/.02	5.38(3.29)
		CB	90.90(.09)	3.79/5.31	2.07(.78)
		PRR	94.27(.06)	3.44/2.28	2.28(.74)
	100	STD1	99.99(.01)	.01/.00	4.41(1.84)
		STD2	94.44(.02)	5.56/.00	4.05(1.58)
		STD3	92.25(.02)	7.75/.00	5.29(2.41)
		CB	93.18(.06)	3.10/3.72	2.07(.70)
		PRR	94.91(.05)	3.02/2.07	2.15(.68)
2	**n**	**Empirical**	**95%**	**2.5%/2.5%**	**2.37**
	50	STD1	99.57(.02)	.02/.41	5.44(3.27)
		STD2	94.46(.03)	5.44/.10	4.74(2.20)
		STD3	92.00(.04)	7.98/.02	6.90(5.09)
		CB	91.35(.08)	3.57/5.08	2.43(.77)
		PRR	93.75(.06)	3.31/2.94	2.60(.75)
	100	STD1	99.92(.01)	.00/.08	5.29(2.21)
		STD2	94.95(.02)	5.04/.01	4.70(1.70)
		STD3	92.43(.02)	7.57/.00	6.59(3.19)
		CB	93.20(.05)	3.03/3.76	2.41(.66)
		PRR	94.32(.04)	2.93/2.75	2.49(.64)
3	**n**	**Empirical**	**95%**	**2.5%/2.5%**	**2.58**
	50	STD1	99.28(.03)	.01/.71	5.56(3.42)
		STD2	94.29(.04)	5.53/.19	4.80(2.15)
		STD3	91.78(.03)	8.17/.05	7.12(5.49)
		CB	91.33(.07)	3.56/5.11	2.60(.77)
		PRR	93.48(.05)	3.33/3.19	2.77(.78)
	100	STD1	99.77(.01)	.00/.23	5.38(2.15)
		STD2	94.83(.02)	5.14/.03	4.75(1.59)
		STD3	92.19(.02)	7.81/.00	6.75(3.20)
		CB	93.06(.05)	3.03/3.91	2.59(.63)
		PRR	93.93(.04)	2.97/3.09	2.67(.63)

Quantities in parenthesis are standard deviations.

		Forecast horizon					
Nominal	**Method**	1	2	4	6	8	12
80%	STD2	108.28	132.86	268.63	400.23	399.38	258.86
	PRR	105.30	132.51	257.79	388.62	376.98	258.08
95%	STD2	166.18	203.94	412.48	614.75	614.31	401.13
	PRR	174.73	203.48	394.92	604.09	608.20	389.36
99%	STD2	221.09	271.38	549.16	818.80	819.73	540.45
	PRR	217.56	260.74	523.04	753.21	801.73	558.21

Table 6. Naive (STD2) and Bootstrap (PRR) interval lengths for the Sales Data.

6249851, e-mail: ortega@est-econ.uc3m.es.