The bootstrap methodology in time series forecasting

C. Cordeiro[1] and M. Neves[2]

[1] Mathematics Department
FCT, University of Algarve
Portugal
ccordei@ualg.pt

[2] Mathematics Department
ISA, Technical University of Lisbon
Portugal
manela@isa.utl.pt

Summary. The bootstrap methodology, initially proposed in independent situations, has revealed inefficient in the context of dependent data. Here, the estimation of population characteristics is more complex. This is what happens in the context of time series. There has been a great development in the area of resampling methods for dependent data. A revision of different approaches of this methodology for dependent data is presented, and the application to a problem of forecasting in time series of air traffic is discussed. The R software is used to obtain forecasts and confidence intervals monthly, for one year, using time series theory and comparing with procedures based on the Bootstrap methodology.

Key words: time series, bootstrap, confidence intervals, forecasts

1 Introduction

A time series is a set of observations usually ordered in equally spaced intervals. The first step in the analysis of any time series is the description of the historic series. It includes the graphical representation of the data. When a time series is plotted, common patterns are frequently found. These patterns might be explained by many possible cause-and-effect relationships. Common components are the trend, seasonal effect, cyclic changes and randomness. The more interesting and ambitious task is to forecast future values of a series on the basis of its recorded past, and more specifically to calculate prediction intervals. So the identification of these components are important in the choice of a forecast model. To see more about forecasting models consult [DeL98], for example.
The bootstrap methodology, initially proposed for independent situations, has revealed inefficient in the context of dependent data. Here, the estimation of population characteristics is more complex. This is what happens in the context of time

series. For dependent data the data generating process is often not fully specified. Then there exists no unique way for resampling. The resampling should be carried out in such a way that the dependence structure should be captured. There has been a great development in the area of resampling methods for dependent data. A revision of different approaches of this methodology for dependent data is presented, and the application to a problem of air traffic time series forecasting is discussed. Forecast intervals for one year are obtained using time series analysis. Forecast intervals based on the bootstrap methodology, using the *R* software, are also obtained and are compared one another for a case study. R packages and functions were used and for some procedures we constructed some functions that did not exist.

2 Holt-Winters Method

The majority of the forecast models assumes that the behavior of the series in the past will continue in the future. It is therefore, extremely important the choice of the model that better describes the behavior of the series in study. In our case, we adjust a model with trend and seasonality, as we see in Figure 1.

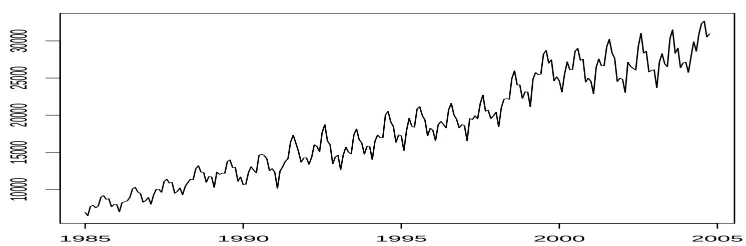

Fig. 1. The air traffic data

Using the theory of forecasting models, a model with these characteristics is the Holt-Winters. Thus, the Holt-Winters method was considered and it was shown that it fitted well the data set, Cordeiro and Neves [CN03]. The structure of this model has the following recursive equations to estimate the trend and the seasonal factor at time t:

(2.1) $T_t = \alpha(X_t/S_{t-s}) + (1-\alpha)(T_{t-1} + b_{t-1}),\ 0 \le \alpha \le 1$
(2.2) $b_t = \beta(T_t - T_{t-1}) + (1-\beta)b_{t-1},\ 0 \le \beta \le 1$
(2.3) $S_t = \gamma(X_t/T_t) + (1-\gamma)S_{t-s},\ 0 \le \gamma \le 1$

where

T_t smoothed value at end of period t after adjusting for seasonality
X_t value of actual demand at end of period t
S_{t-s} smoothed seasonal index, s periods ago

b_t smoothed value of trend through period t
α smoothing constant used for T_t
β smoothing constant used to calculate the trend (b_t)
γ smoothing constant used for calculate the seasonal index in period t

The forecast equation is:

$$\hat{X}_{t+h} = T_t + hb_t + S_{t+h-rs},$$

where h is the number of periods to be forecasted into the future, i.e., forecast horizon and $r = 1$, $1 \leq h \leq s$, $r = 2$, $s < h \leq 2s$, etc.
In contrast to other methods, this does not keep information on all last history and it is more adapted to the evolution of the series, giving bigger importance (weight) to most recent (in detriment of oldest) observations.

3 Bootstrap for Dependent Data

The bootstrap is a computer-intensive method introduced by Efron [Efr79] to present solutions in situations where the traditional methods failed. But, Efron's bootstrap classical approach does not contemplate structures of dependent data, such as a time series, where the dependence data arrangement should be kept during the resampling scheme.
Recently, resample methods for dependent data have been developed. Most of them consider segments of the data to define blocks, such that the dependence structure within each block can be kept. Different versions of blocking differ in the way as blocks are constructed.

Let $X = \{X_1, \cdots, X_n\}$ be the observations from a stationary process. Let us denote by $l \in \{1, \cdots, n\}$ and $b \geq 1$ the length and the number of blocks, respectively, such that $l \times b \leq n$. Let n and m be the initial data size and the bootstrap sample size, $m \leq n$ and k the number of blocks chosen.

We will consider a review of the following resampling techniques for dependent data: the Nonoverlapping Block Bootstrap, Carlstein [Car92]; the Moving Block Bootstrap, Künsch [Kun89]; the Circular Block Bootstrap and the Stationary Block Bootstrap, Politis and Romano, [PR92] and [PR94]; and for a large class of stationary processes, Bühlmann [Bul97] presented the Sieve Bootstrap method based on a sieve of autoregressive processes of increasing order. For some theoretical comparisons of those methods see Lahiri [KK04] and [Lah99].

3.1 The Block Bootstrap

Nonoverlapping Block Bootstrap

The nonoverlapping block bootstrap(NBB) method consists of dividing the time series into b blocks of consecutive observations denoted by

$$B_i = (X_{(i-1)l+1}, \cdots, X_{il}) \; i = 1, \ldots, b.$$

A random sample of k blocks, $k \geq 1$, $B_1^*, \cdots, B_k^*$ is selected with replacement from $\{B_1, \cdots, B_b\}$. Joining the k blocks the *bootstrap* sample is constructed with $m = k \times l$ observations. So, we get the sample

$$(X_1^*, \cdots, X_l^*, \cdots, X_{(k-1)l+1}^*, \cdots, X_m^*).$$

Notice that, for this bootstrap procedure, the correlation is considered strong inside the blocks of observations and relatively weak between blocks.

Moving Block Bootstrap

The moving block bootstrap(MBB) was proposed by Künsch [Kun89] and Liu and Singh [LS92]. This procedure resamples blocks of consecutive observations at a time. As a result, the dependence structure of the original observations is preserved within each block. Let

$$B_i = (X_i, \cdots, X_{i+l-1})$$

denote the block of length l starting with X_i, $1 \leq i \leq b$ where $b = n - l + 1$. The MBB sample is obtained by randomly selecting a suitable number of blocks from the collection $\{B_1, \cdots, B_b\}$. Accordingly, let $B_1^*, \cdots, B_k^*$ denote a simple random sample drawn with replacement from $\{B_1, \cdots, B_b\}$. Denoting the elements in B_i^* by $(X_{(i-1)l+1}^*, \cdots, X_{il}^*)$, $i = 1, \cdots, k$, the MBB sample of size m=k×l is $X_1^*, \cdots, X_m^*$. The special case where each block consists of a single element (i.e., l=1) is the Efron's bootstrap classic method [Efr79]. For other formulations and details see [KK04].

Circular Block Bootstrap

This bootstrap procedure is an extension of the previous method. The idea is to wrap the data around a circle and form additional blocks using the "circularly defined" observations. So, for $i > n$, we define $X_i = X_{i_n}$, where $i_n = imodn$ and $X_0 = X_n$. The circular block bootstrap(CBB) method, proposed by Politis and Romano [PR92], resamples overlapping and periodically extended blocks of length l.
Notice that each X_i appears exactly l times in the collection of blocks, and since the CBB resamples the blocks from this collection with equal probability, each of the original observations $X_1, \cdots, X_n$ receives equal weight under the CBB. This property distinguishes the CBB from previous methods, the MBB and the NBB, which suffer from edge effects.
Like the previous methods, k blocks of length l are selected and organized in a sequence of observations $X_1^*, \cdots, X_m^*$.

Stationary Block Bootstrap

The stationary block bootstrap(SBB) differs from the earlier block bootstrap methods because it uses blocks of random lengths rather than blocks of a fixed length l. Using Politis and Romano [PR94] formulation, this method can be described as: let X_1^* be picked at random from $\{X_1, \cdots, X_n\}$. To select the next observation X_2^*, we further randomize and perform a binary experiment with a p probability of "success". If the binary experiment results in a "success", then we select X_2^* again at random from $\{X_1, \cdots, X_n\}$. Otherwise we set X_2^* the observation next to X_1^* in the periodically extended series. And so on, until contain the n observations in the bootstrap sample.
An important property of the SBB method is that the bootstrap observations $\{X_i^*\}$ $i \in 1, 2, \cdots, n$ are stationary, the reason why it is called the "stationary" bootstrap (see [KK04]).

3.2 The Sieve Bootstrap

Recently Bühlmann [Bul97] proposed a new bootstrap scheme called sieve bootstrap(SB). This method is based on the idea of fitting parametric models first and resampling from the residuals. However the model is chosen adaptively rather than considering a pre-fixed model. This approach is different from the previous bootstrap methods, the sample bootstrap is (conditionally) stationary and does not present structure of dependence. Another different feature is that the sieve bootstrap sample is not a subsample from the original data, as in the previous methods.
Given a sample $X_1^*, \cdots, X_m^*$, an autoregressive process is estimate. The residuals are then centered and the empirical cumulative distribution function of these residuals is obtained. From this distribution we get an i.i.d. residuals resample. So we generate the bootstrap error series and than a bootstrap series. Observe that even if the sieve bootstrap is based on a parametric model it is basically non parametric in its spirit. The AR(p) model here is just used to filter the residuals series.
Zagdanski [Zag99] consider several applications of sieve bootstrap. He obtained prediction intervals for future observations of stationary time series and described the construction of the best linear predictor for future value of time series. The sieve bootstrap is shown as providing consistent estimators of the conditional distribution of future values given the observed data.

4 The Case Study

The air traffic increased too much in the last decade, specially in the recent years. Portugal has two Flight Information Regions (FIRs)-Lisbon and Santa Maria (Azores). In these regions the air traffic management is ensured both in the air and on the ground. The real case study refers to the air traffic of Lisbon's FIR and we intended to get monthly forecasts intervals for one year.
Our first approach is to see how the methods perform, so we use the time series (1985-2004) to predict for 2005. Since we have the true values for 2005, we intended to compare their performance.
In Figure 1 the data are plotted in order to observe its behavior through time horizon and to identify its components. Cordeiro and Neves [CN03] showed that the best forecasting model for this data is that obtained by Holt-Winters multiplicative method. Bootstrap methods for dependent data, described in section 3, were also considered and extensive computational work was performed using *R* software. Some R packages and functions were used and for some procedures we constructed some functions that did not exist.
In order to apply the block bootstrap methodologies the trend was first estimated and removed from the time series. Tests of stationarity are performed and it is accepted for the transformed time series. Then blocks of length l=12 were formed because this choice was revealed as being the best. The results for NBB, MBB, CBB, SBB and SB methods were obtained using 1000 replicates. In what concerns SB case, the Holt-Winters model was previously fitted. For each bootstrap sample an autoregressive model was fitted to the residuals.
After all these computational work, we obtain for each method 95% monthly forecasts intervals. In table 1 the results are presented and when the T. Value is inside the interval it is in **bold**. We can observe that the NBB, MBB and CBB present

the wider intervals and SB are narrower and don't contain often the true value. It seems us that the best compromise will be the SB. Some work is now in progress concerning this aim.

We can observe that the Sieve Bootstrap and Stationary Block Bootstrap have

Table 1. Monthly Forecasts Intervals.

Month	T.Value	HW	NBB	MBB	CBB	SBB	SB
Jan	27142	**26176;27737**	**25408;32270**	**25552;32910**	**25375;32524**	**27133;30682**	**25550;27375**
Feb	25760	23561;25552	25989;33088	**25487;32996**	**25194;33007**	27244;31069	23550;25372
Mar	27897	**26845;29462**	**25687;33387**	**25367;33312**	**25147;32849**	**27273;31183**	**25903;27783**
Apr	29901	**27642;30661**	**25344;33100**	**25583;33416**	**25271;33074**	**27467;31304**	**27351;29193**
May	28596	**27063;30326**	**25972;33321**	**25540;33843**	**25294;33204**	**27505;31241**	**27152;28969**
June	30966	26822;30338	**25534;33129**	**25565;34032**	**25266;33333**	**27603;31478**	26366;28284
July	32349	**30209;34379**	**26060;33239**	**25648;34100**	**25416;33383**	27605;31389	28872;30779
Aug	32671	**31293;35818**	**26008;32978**	**25980;34029**	**25676;33644**	27697;31644	30076;32002
Sep	30530	**28853;33255**	**26247;33318**	**26189;34182**	**25800;33448**	**27948;31542**	28047;30033
Oct	30950	**28558;33126**	**26860;34072**	**26278;34256**	**25994;33501**	**28048;31706**	26878;28695
Nov	27516	**25343;29641**	**27045;33700**	**26420;33854**	**26188;33603**	28145;31948	24510;26507
Dec	28867	**25966;30279**	**27588;33377**	**26363;33729**	**26338;33502**	**28177;32029**	25536;27740

a worst performance compared with the other methods. Moving blocks bootstrap and Circular block bootstrap methods, have a better behavior than the Holt-Winters(HW).

5 Remarks

For an air traffic time series different bootstrap methods for dependent data were employed to obtain forecasts intervals, after adequate transformation in order to get stationarity. Although the Moving Block Bootstrap and the Circular Block Bootstrap contain the true value they have very large intervals compared with the Sieve Bootstrap and the Stationary Block Bootstrap. It seem to us that the Sieve Bootstrap is a good compromise for obtaining forecasts intervals.

Acknowledgements

The authors are grateful to the company **Portugal Navigation-NAV Portugal, E.P.E.** for the data that made this study possible and for providing data that will allow continuing this work.

References

[APR02] Alonso, A.M., Peña, D. and Romo, J.: Forecasting time series with sieve bootstrap. Journal of Statistical Planning and Inference ., **100**, 1–11 (2002)

[AK03] Archibald and Koehler: Normalization of seasonal factors in Winters methods. International Journal of Forecasting.,**19**, 143–148 (2003)

[Bul97] Bühlmann, P.: Sieve Bootstrap for Time series. Bernoulli., **3**, 123-148 (1997)

[Car86] Carlstein, E.: The use of subseries values for estimating the variance of a general statistic from a stationary sequence. The Annals of Statistics., **14**, 1171–1179 (1986)

[Car92] Carlstein, E.: Resampling Techniques for Stationary Time-Series: Some Recent Developments. New Directions in time series Analysis., 75–85 (1992)

[CN03] Cordeiro, C. e Neves, M.: Séries temporais e modelos de previsão. Aplicação da metodologia Bootstrap. Estatística com Acaso e Necessidade. Actas do XI Congresso Anual da SPE., 153–164 (2003)

[DeL98] DeLurgio,S.A.: Forescasting Principles and Applications. McGraw-Hill International Editions (1998)

[Efr79] Efron, B.: Bootstrap methods: another look at the Jackknife. The Annals of Statistics., **7**, 1–26 (1979)

[Kun89] Künsch, H.: The Jackknife and the Bootstrap for General Stationary Observations. The Annals of Statistics., **17**, No.3, 1217–1241 (1989)

[Lah03] Lahiri, S.N.: Resampling Methods for Dependente Data. Springer Verlag Inc.(2003)

[Lah99] Lahiri, S.N.: Theoretical Comparisons of Block Bootstrap Methods. The Annals of Statistics., **27**, No.1, 386–404 (1999)

[LS92] Liu, R. Y. and Singh, K.: Moving blocks jackknife and bootstrap capture weak dependence. Exploring the Limits of Bootstrap., Wiley New York, 225–248 (1992)

[PR92] Politis, D.N. and Romano, J.P.: A circular Block-Resampling Procedure for Stationary Data. Exploring the Limits of Bootstrap., Wiley New York, 263–270 (1992)

[PR94] Politis, D.N. and Romano, J.P.: The Stationary Bootstrap. Journal of the American Statistical Association., **89**, No. 428, 1303–1313 (1994)

[Zag99] Zagdanski, A.: Prediction Intervals for Stationary Time Series Using the Sieve Bootstrap Method. Preprint (1999)