

EXPONENTIAL SMOOTHING AND RESAMPLING TECHNIQUES IN TIME SERIES PREDICTION

MARIA MANUELA NEVES

*CEAUL and Mathematics Department, Instituto Superior de Agronomia Technical
University of Lisbon
Tapada da Ajuda, 1349–017, Lisboa, Portugal*

e-mail: manela@isa.utl.pt

AND

CLARA CORDEIRO

*Mathematics Department, Faculty of Science and Technology
University of Algarve, Faro, Portugal*

e-mail: ccordei@ualg.pt

Abstract

Time series analysis deals with records that are collected over time. The objectives of time series analysis depend on the applications, but one of the main goals is to predict future values of the series. These values depend, usually in a stochastic manner, on the observations available at present. Such dependence has to be considered when predicting the future from its past, taking into account trend, seasonality and other features of the data. Some of the most successful forecasting methods are based on the concept of exponential smoothing. There are a variety of methods that fall into the exponential smoothing family, each having the property that forecasts are weighted combinations of past observations. But time series analysis needs proper statistical modeling. The model that better describes the behavior of the series in study can be crucial in obtaining “good” forecasts.

Departures from the true underlying distribution can adversely affect those forecasts. Resampling techniques have been considered in many situations to overcome that difficulty. For time series, several authors have proposed bootstrap methodologies. Here we will present an automatic procedure built in  language that first selects the best exponential smoothing model (among a set of possibilities) for fitting the data, followed by a bootstrap approach for obtaining forecasts. A real data set has been used to illustrate the performance of the proposed procedure.

Keywords: time series; bootstrap; exponential smoothing; forecasting; accuracy measures.

2000 Mathematics Subject Classification: 62F40, 60G25, 37M10.

1. INTRODUCTION AND MOTIVATION

A time series is a set of observations $\{x_{t_1}, x_{t_2}, \dots, x_{t_N}\}$ each one being recorded at a specific time $t_1, t_2, \dots, t_N$. One specific feature in time series is that the records are usually dependent.

A time series is said to be *discrete* (the type here considered) if the set T_0 of times at which observations are made is a discrete set. Usually the records are done at equally spaced times and the time series is then represented by

$$\{x_t, t \in T_0\} \text{ where } T_0 = \{1, 2, \dots, N\} \text{ or } T_0 = N \text{ or } T_0 = Z.$$

If the observations are made continuously in time, the time series is said to be *continuous*.

The areas of application of time series cover any area where statistics is applied, the main one being perhaps Economics, Engineering, Social Sciences, Medical Sciences and Environmental Sciences.

The motivation for this work was to model a data set of the number of airplanes that per month cruises the Flight Information Region (FIR) of Lisbon for the period 1985–2009* and to predict values of the series. A plot of these data can be seen in Figure 1.

*Data gently given by the Portugal Navigation-NAV Portugal, E.P.E.

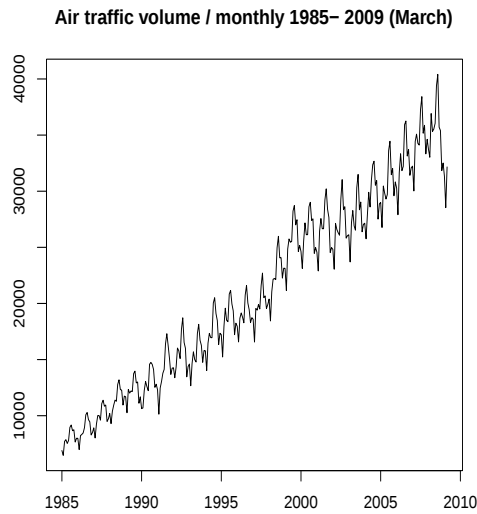


Figure 1. Number of planes in FIR of Lisbon per month.

In Cordeiro and Neves (2006) different approaches of resampling techniques for dependent data were applied to the first part of this data set (1985–2005) and prediction intervals were obtained.

The idea of this work is now to join exponential smoothing and bootstrap to built an automatic procedure in  language for modeling and forecasting time series.

2. TIME SERIES ANALYSIS

When the analysis of a time series is performed there are several objectives we want to attain. Some of the main ones are (Chatfield, 2004):

- **Description** The first step in the analysis is usually to plot the data and to obtain simple descriptive measures of the main properties. It begins with plotting a graph, looking for trend and seasonal effects, possible *outliers*, possible turning points, etc.
- **Explanation** When there are observations on two or more variables, we can try to use the variation in one time series to explain the variation in the other. This needs to know the process which generates a given time series.

- **Forecasting** Given an observed time series one of the main goals is to predict the future values of the series. Adequate models for describing the series are now very important.
- **Control** It is important in many engineering and industrial applications after predicting values of a series to adjust various control parameters.

A time series can be thought as a combination of various components:

- **Trend** (T) – describes the long term direction of the series;
- **Seasonal** (S) – describes the short term recurring pattern of change, that repeats with a known periodicity (e.g. 12 months per year or 7 days per week);
- **Cycle** (C) – describes a pattern that repeats with some regularity but with varying amplitude and duration;
- **Error** (ϵ) – describes the erratic movement in the series.

The graphical representation of a time series is a very important auxiliary tool, because it allows us to look for those patterns that the time series can exhibit, described above. These patterns need to be considered to forecast future values.

Those components can be combined in several ways, two examples are:

- an additive model $y_t = T_t + S_t + C_t + \epsilon_t$,
- a multiplicative model $y_t = T_t \times S_t \times C_t \times \epsilon_t$,

where y_t is the observation, T_t the trend, S_t the seasonality, C_t the cycle component and ϵ_t the random error at time t .

It is of primary importance to develop mathematical models that provide plausible explanations for sample data in a time series and can be used for modeling and forecasting. Theory of stochastic processes is then absolutely necessary. Most of the probability theory of time series is concerned with *stationary time series* and for this reason several procedures were developed to turn a non-stationary series to a stationary one, see Chatfield (2004).

3. EXPONENTIAL SMOOTHING METHODS

Forecasting future values of a time series is one of the main objectives in the analysis of a time series. Forecasting methods have been developed based on well known models: AR, ARIMA, etc.

Around since the 1950s another class of forecasting methods appeared. Exponential smoothing (EXPOS) refers to a set of methods that, in a versatile way, can be used to model and to obtain forecasts. These methods are based on the concept of exponential smoothing, i.e., they have ... *the property that forecasts are weighted combinations of past observations, with recent observations given relatively more weight than older observations. The name “exponential smoothing” reflects the fact that the weights decrease exponentially as the observations get older.* (Hyndman *et al.*, 2008).

Following very closely Hyndman *et al.* (2008) we can summarize some of the most well-known exponential smoothing methods:

- Simple exponential smoothing — Given a time series, let $\hat{y}_t$ be the forecast for the value y_t . Once observed this value, $\epsilon_t = y_t - \hat{y}_t$ is the forecast error. Brown (1959) considered to obtain the forecast for the next period, $\hat{y}_{t+1}$, as the forecast for the previous period adjusted by using the forecast error, i.e.,

$$(1) \quad \hat{y}_{t+1} = \hat{y}_t + \alpha(y_t - \hat{y}_t),$$

with α a constant between 0 and 1.

Developing relation (1) it is easy to show that $\hat{y}_{t+1}$ represents a weighted moving average of all past observations with the weights decreasing exponentially.

- Holt’s linear method — simple exponential smoothing works well when there are no trend, seasonality or other patterns. Holt (1957) extended it to linear exponential smoothing allowing forecasting of data with trend. The forecast for this method is found using two smoothing constants, α and β (with values between 0 and 1) and three equations:

$$\begin{aligned} - \text{Level} \quad & l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ - \text{Growth} \quad & b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \\ - \text{Forecast} \quad & \hat{y}_{t+h|t} = l_t + b_t h, \end{aligned}$$

where l_t denotes an estimate of the level of the series at time t and b_t denotes an estimate of the slope (growth) of the series at time t . Level l_t and slope b_t are considered as two components of the trend.

- Holt-Winters Trend and Seasonality Method — the previous methods are not appropriate if data exhibit seasonal patterns. Holt (1957) proposed a method for seasonal data, later improved by Winters(1960), the reason why is known by “Holt-Winters method”. The method is based on three smoothing equations (for level, growth and seasonality). For additive seasonality the equations are:

$$\begin{aligned}
 - \text{Level} \quad & l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}) \\
 - \text{Growth} \quad & b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \\
 - \text{Seasonal} \quad & s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \\
 - \text{Forecast} \quad & \hat{y}_{t+h|t} = l_t + b_th + s_{t-m+h_m^+}
 \end{aligned}$$

$h_m^+ = [(h-1) \bmod m] + 1$ and parameters (α, β, γ) are usually restricted to lie between 0 and 1.

- Gardner and McKenzie (1985) proposed a modification of Holt’s linear and Holt-Winters to allow the “damping” of trends, i.e., the growth is dampened by a factor of ϕ for each additional future time period. For example, in Holt’s linear method the equation for level will become

$$l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + \phi b_{t-1}).$$

Methods for the best choice of starting values for the parameters in equations above, “the initialization problem”, have been studied for several authors.

Pegels (1969) classified exponential smoothing methods regarding the trend and seasonal patterns that a series reveals as: none, additive (linear) or multiplicative (nonlinear). Since then, many researchers such as Gardner(1985), Hyndman *et al.* (2002) and Taylor (2003) have investigated and developed EXPOS methods. Fifteen possibilities of exponential smoothing (ignoring the error component) are resumed in Table 1. For example (N, N) stands for the simple exponential smoothing, (A, N) the Holt’s linear method; see Hyndman *et al.* (2008) for the complete description.

Table 1. The exponential smoothing methods.

Trend Component	Seasonal Component		
	N (None)	A (Additive)	M (Multiplicative)
N (None)	N,N	N,A	N,M
A (Additive)	A,N	A,A	A,M
Ad (Additive damped)	Ad,N	Ad,A	Ad,M
M (Multiplicative)	M,N	M,A	M,M
Md (Multiplicative damped)	Md,N	Md,A	Md,M

In the past ten years the exponential smoothing methods has undergone a substantial revolution due to the consideration of the new state space framework.

Hyndman *et al.* (2008) presents exponential smoothing state space models for all methods showed in Table 1, allowing then to obtain forecasting intervals.

4. THE PROCEDURE

Cordeiro and Neves (2006, 2007) compared the following bootstrap methodologies for dependent data and obtained forecast intervals:

- Block bootstrap: Non-overlapping block bootstrap, Moving block bootstrap, Circular block bootstrap, Stationary block bootstrap.
- Sieve bootstrap.

Sieve bootstrap was proposed by Bühlman (1997) for dependent observations and extended by Alonso *et al.* (2002) for constructing prediction intervals

in stationary time series. In Cordeiro and Neves (2006, 2007) sieve bootstrap revealed to be better than the other block bootstrap methodologies for obtaining prediction intervals.

Based on those ideas the procedure **Boot.EXPOS** was developed and the steps are summarized and compared below with those of the sieve bootstrap. Step 0, step 5 and step 6 contain the differences we established to the sieve bootstrap.

Sieve bootstrap	Boot.EXPOS
Step 1. Adjust an autoregressive model with increasing order p using AIC criterion;	Step 0. Select the best EXPOS method by AIC; components are removed and the residuals obtained;
Step 2. Obtain the residuals;	Step 1. Adjust an autoregressive model with increasing order p using AIC criterion;
<i>For B replicates:</i>	<i>For B replicates:</i>
Step 3. Resample the centered residuals;	Step 2. Obtain the residuals;
step 4. Use AR for obtaining a new series by recursion;	Step 3. Resample the centered residuals;
Step 5. Fit $AR(p)$ to the new series;	Step 4. Use AR for obtaining a new series by recursion;
Step 6. Obtain the predicted values from the new series using the previous $AR(p)$ fit.	Step 5. Add the components in step 0 to the new series; fit EXPOS method (same type as in Step 0);
	Step 6. Obtain the predicted values from the new series using the previous EXPOS fit.

5. THE COMPUTATIONAL WORK

Cordeiro and Neves (2008, 2009) present a computational algorithm considering a choice among only four methods: single exponential smoothing, Holt's linear trend, Holt-Winters seasonal smoothing with additive and with

multiplicative seasonality. Now that algorithm was extended to the fifteen methods pointed out in Section 3. and several accuracy measures were considered.

The `R` procedure we built considers statistical tests, transformations and differentiation whenever it is necessary for studying and/or obtaining stationarity of the random part before the AR adjustment.

Some `R` packages are used: **car**, **forecast** and **tseries**. The new procedure using bootstrap and EXPOS methods, **Boot.EXPOS()**, was constructed in `R` language.

The performance of our procedure is evaluated considering the in-sample performance and the out-sample performance. The in-sample performance is based on the Mean Squared Error (MSE) and the out-sample performance is based on the forecasts for a given period. Denoting by $e_t = y_t - \hat{y}_t$ the forecasting error, the following accuracy measures are here considered:

Table 2. Accuracy measures.

Acronyms	Definition	Formula
RMSE	Root Mean Squared Error	$\sqrt{\text{mean}(e_t^2)}$
MAE	Mean Absolute Error	$\text{mean}(e_t)$
MAPE	Mean Absolute Percentage Error	$\text{mean}(100 \left \frac{e_t}{y_t} \right)$

6. CASE STUDY

The data set here studied is the Australian monthly gas production in the period of 1956 until 1995 (the data set is *gas* in `R`, Figure 2). It is a monthly time series with 476 observations in 39 years. The series exhibits periodic behavior with a seasonal cycle, $s=12$ months. The Australian gas production has a constant production from 1956 to 1970, and since then an increase is observed, due to the augment of gas consumption.

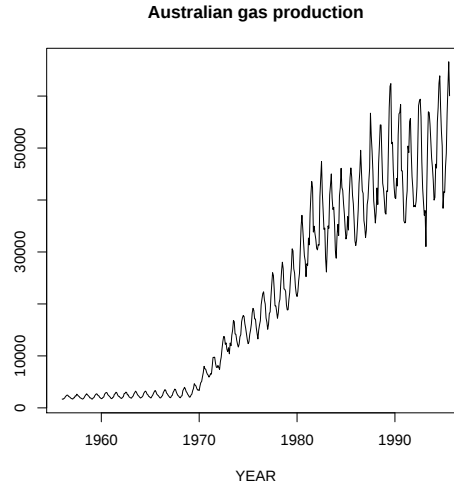


Figure 2. Australian gas production.

This time series was separated into two blocks:

- the first $\{y_1, \dots, y_{464}\}$ from January 1956 to December 1994 as a sample set for model estimation;
- the remaining $h=12$ observations, that is $\{y_{465}, \dots, y_{476}\}$ from January 1995 to December 1995 as post-sample for forecast evaluation.

The first step in our procedure is to fit the best EXPOS model (among the 15 methods with two sources of errors: additive and multiplicative) selected by AIC criterion minimization. The parameters are optimized using the mean squared error (MSE). For this time series the EXPOS model chosen is that one with multiplicative damped trend, multiplicative seasonality and multiplicative error term, $\text{ets}^\dagger(\text{M}, \text{Md}, \text{M})$. Figure 3(a) shows level, trend and seasonality components.

[†]ets stands for error, trend and seasonality.

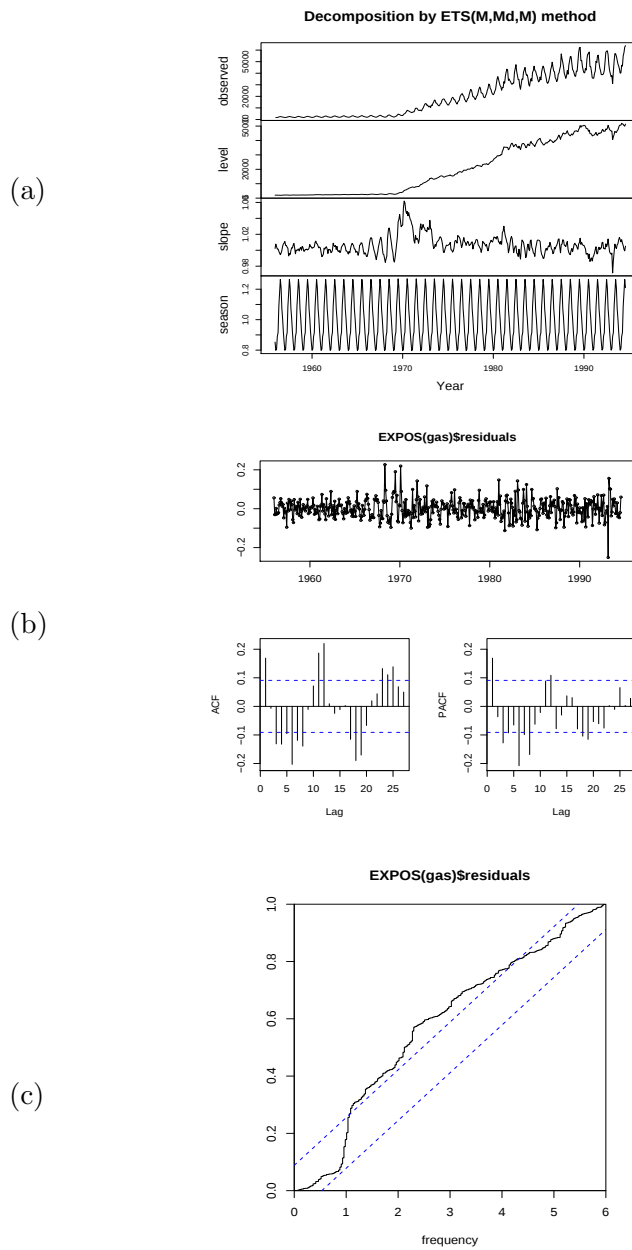


Figure 3. Time series gas: (a) EXPOS choice and EXPOS residuals; (b) correlogram and partial correlogram, (c) cumulative periodogram.

After fitting an EXPOS model, the residuals are extracted for white noise checking. Graphical approaches, correlogram and the cumulative periodogram are used, Figure 3(b) and (c), for diagnostic checking of the residuals. High correlation is evident. An autoregressive model is used to filter the EXPOS residuals. After fitting an AR model the correlogram, partial correlogram and cumulative periodogram are obtained, see Figure 4. Now the white noise hypothesis is not rejected so the bootstrap of the residuals is applied.

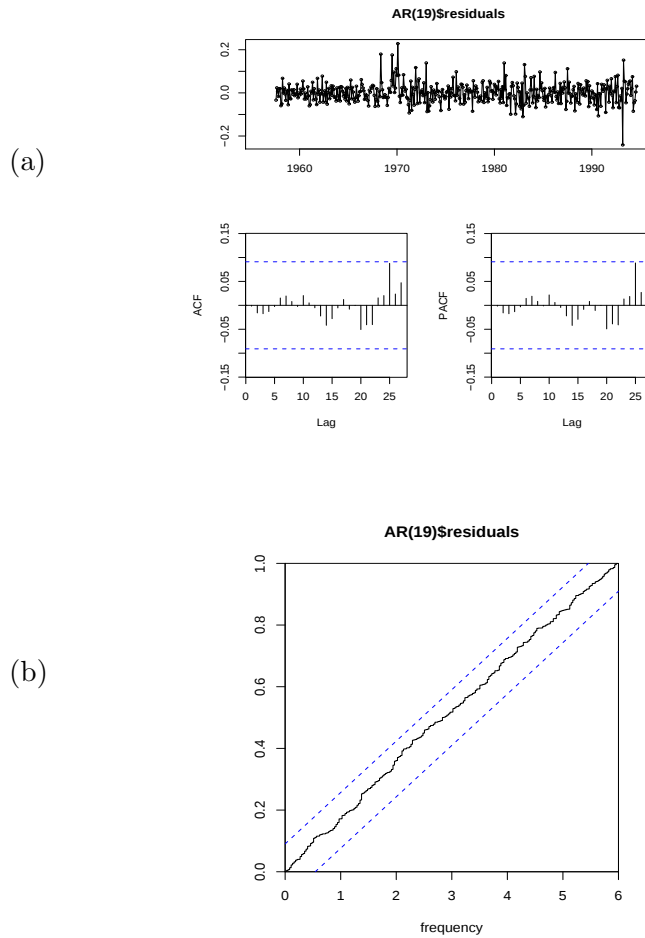


Figure 4. AR residuals: (a) correlogram and partial correlogram; (b) cumulative periodogram.

Boot.EXPOS steps shown in Section 4 are now executed and the accuracy measures proposed in Table 2 are calculated.

The results in Table 3 reveal the good performance of Boot.EXPOS.

Table 3. Accuracy measures for the *gas* time series.

				Accuracy measures			
Serie	n	s	h	Method	RMSE	MAE	MAPE
<i>gas</i>	476	12	12	ets	2773.72	2097.73	4.22
				Boot.EXPOS	2348.16	1908.15	3.84

7. CLOSING COMMENTS

An initial approach, Cordeiro and Neves (2008, 2009), for using exponential smoothing and bootstrap in time series forecasting is here extended by incorporating more models for selection and more accuracy measures for checking the performance of the procedure.

A case study illustrates the application of our procedure. Several other examples have been studied and the results obtained reveal this procedure as a very promising technique.

REFERENCES

- [1] A.M. Alonso, D. Peña and J. Romo, *Forecasting time series with sieve bootstrap*, Journal of Statistical Planning and Inference **100** (2002), 1–11.
- [2] R.G. Brown, *Statistical Forecasting for inventory control*, New York, McGraw-Hill 1959.
- [3] P. Bühlmann, *Sieve bootstrap for time series*, Bernoulli **3** (1997), 123–148.

- [4] C. Chatfield, *The Analysis of Time Series. An Introduction*, 6th ed. Chapman & Hall 2004.
- [5] C. Cordeiro and M.M. Neves, *The Bootstrap methodology in time series forecasting*, In "Proceedings of CompStat2006" (J. Black and A. White, Eds.), Springer Verlag (2006), 1067–1073.
- [6] C. Cordeiro and M.M. Neves, *The Bootstrap prediction intervals: a case-study*, In "Proceedings of the 22nd International Workshop on Statistical Modelling (IWSM2007)" (J. Castillo, A. Espinal and P. Puig, Eds.), Springer Verlag (2007), 191–194.
- [7] C. Cordeiro and M.M. Neves, *Bootstrap and exponential smoothing working together in forecasting time series*, In "Proceedings in Computational Statistics (COMPSTAT 2008)" (Paula Brito, Editor), Physica-Verlag (2008), 891–899.
- [8] C. Cordeiro and M.M. Neves, *Forecasting time series with Boot.EXPOS procedure*, *REVSTAT*, **7:2** (2009), 135–149.
- [9] E.S. Gardner, *Exponential smoothing: the state of the art*, *J. of Forecasting*, **4** (1985), 1–38.
- [10] E.S. Gardner and E. McKenzie, *Forecasting trends in time series*, *Management Science* **31** (1985), 1237–1246.
- [11] C. Holt, *Forecasting seasonals and trends by exponentially weighted averages*, O.N.R. Memorandum 52/1957, Carnegie Institute of Technology 1957.
- [12] R. Hyndman, A. Koehler, R. Snyder and S. Grose, *A state framework for automatic forecasting using exponential smoothing methods*, *International Journal of Forecasting* **18** (2002), 439–454.
- [13] R. Hyndman, A. Koehler, J. Ord and R. Snyder, *Forecasting with Exponential Smoothing: The State Space Approach*, Springer-Verlag Inc 2008.
- [14] C.C. Pegels, *Exponential smoothing: some new variations*, *Management Science* **12** (1969), 311–315.
- [15] R Development Core Team, *R: A language and environment for statistical computing*, R Foundation for statistical computing, Vienna, Austria 2008. <http://CRAN.R-project.org>.
- [16] J.W. Taylor, *Exponential smoothing with a damped multiplicative trend*, *International Journal of Forecasting/Management Science* **19** (2003), 273–289.

- [17] P.R. Winters, *Forecasting sales by exponentially weighted moving averages*, Management Science **6** (1960), 349–362.

Received 21 January 2010