Universidad de la República
Facultad de Ciencias Sociales
DEPARTAMENTO DE ECONOMIA

Documentos de trabajo

Prediction with univariate time series models: The Iberia case

E. Ruiz
F. Lorenzo

Documento No. 02/98
Diciembre, 1998

PREDICTION WITH UNIVARIATE TIME SERIES MODELS: THE IBERIA CASE[1]

by

Esther Ruiz[2]
Fernando Lorenzo[3]

December 1997

---

1. We are grateful to Antoni Espasa and Eva Senra for helpful commnents on a previous version of this paper. All errors are our own responsability.

2. Departamento de Estadística y Econometría. Universidad Carlos III de Madrid. C/Madrid 126, 28903 Gestafe, Madrid, Spain.. Phone: (349 1) 6249851, Fax: (349 1) 6249849 E-mail: ortega@est-econ.uc3m.es.

3. Departamento de Economía, Facultad de Ciencias Sociales, Universidad de Uruguay. Centro de Investigaciones Económicas (CINVE-Uruguay) E-mail: cinve@chasque.apc.org.

ABSTRACT

In this paper we model the monthly number of passengers flying with the Spanish airline IBERIA from January 1985 to December 1992 and predict future values of the series up to October 1994. This series is characterized by strong seasonal variations and by having an upward trend which has a rupture during 1990 with the slope changing to be negative. We compare observed values with predictions made by a deterministic components model, the Holt-Winters exponential smoothing filter, an ARIMA model and a structural time series model. As expected, we show that the deterministic components model is too rigid in the presence fo breaks in trends although surprisingly the within-sample fit is better than for any of the other models considered. With respect to Holt-Winters predictions, they fail because they are not able to acommodate outliers. Finally, ARIMA and structural models are shown to have very similar prediction performance, being very flexible to predict reasonably well when there are changes in trend and outliers.

Key words: ARIMA models, Breaks in trends, Deterministic components, Holt-Winters algorithm, Outliers, Intervention analysis, Structural time series models, Unobserved components models.

RESUMEN

En este artículo se modelizan las observaciones mensuales del número de pasajeros en IBERIA desde enero de 1985 hasta diciembre de 1992 y se predicen valores futuros de la serie hasta octubre de 1994. Dicha serie se caracteriza por tener un marcado componente estacional y una tendencia positiva hasta 1990 cuando pasa a ser negativa. Se comparan predicciones realizadas mediante un modelo con componentes deterministas, el algoritmo de Holt-Winter, un modelo ARIMA y un modelo de series temporales estructural. Como era de esperar, el modelo determinista es demasiado rígido cuando existen rupturas en la tendencia aunque, sorpresivamente, tiene el mejor ajuste muestral de todos los modelos considerados. Las predicciones realizadas mediante el algoritmo de Holt-Winter son sesgadas por no ser capaces de acomodar la presencia de datos atípicos en la muestra. Finalmente, la capacidad predictiva de los modelos ARIMA y estructura es muy similar, siendo ambos muy flexibles para predecir adecuadamente tanto cuando hay cambios en la pendiente como observaciones atípicas.

INDEX

1. INTRODUCTION

Breaks in trends of economic time series can often be stated *a posteriori*. However, usually, it is interesting to predict a crisis or a recovery as soon as possible. Consider, for example, the monthly series of passengers in the Spanish airline IBERIA which is represented in figure 1 after being transformed into logarithms. In this figure, it seems rather clear that there has been a change in the long-run trend of the series from being positive to negative and one may be tempted to fit a deterministic trend model with a break in the slope around December 1990. However, if we were in 1992, could we detect a change in the trend? In this paper, we fit several univariate time series models to the IBERIA series to analyse which of them adapts sooner after the crisis. We are considering models which assume that the components of a series are deterministic and models with stochastic components. Between the latter, we are first using the Holt-Winters algorithm. This method is very simple but is not based on a statistical model and, consequently, has fundamental limitations. There are two main classes of statistical models with stochastic components: ARIMA and structural time series models which will also be fitted to the same data set.

The empirical analysis of the IBERIA airline data indicates that the ARIMA and structural time series models, are rather similar being both flexible enough as to adapts after the change in the long-run trend of the series. However, both the deterministic model and the Holt-Winters algorithm have problems when use for prediction purposes.

The paper is organized as follows. In section 2, the main statistical properties of each of the models considered are described. In section 3, we present a brief description of prediction theory. The models are fitted to the IBERIA series in section 4. Then, the estimated models are used for prediction purposes. Finally, section 5 presents the conclusions and some suggestions for further research.

2. UNIVARIATE TIME SERIES MODELS

In this section we describe four different approaches to describe the dynamic behaviour and forecast future values of univariate time series. The time series we are focusing on are monthly series exhibiting long-run trend and seasonal dynamics as, for example, the IBERIA series represented in figure 1. First, we are considering models with deterministic trend and seasonal components. These models can be too rigid in many situations, specially when dealing with long time series. In general, it will be desirable to allow the components to evolve over time. There are several alternatives to allow the components to change over time. First, the components can be deterministic in subsamples with changes in discrete points of time. In this case, it is necessary to specify the mechanism by which the components change, and this could be in most cases a rather difficult task. Consequently, in this paper we consider models which contain unit roots, allowing the components to evolve continously in time following a stochastic process.

Between the procedures which consider that the components of a time series are stochastic, we are first considering some filters based on discounting past observations. These filters are *ad hoc* in the sense that they are not based on a statistical model. Alternatively, we are considering two statistical models with unit roots, namely ARIMA and structural time series models. Latter on, we will see that structural models can be seen as ARIMA models with restrictions on the parameters.

2.1 Deterministic components models

The trend and seasonal components of a time series, y_t, can be modelled as deterministic functions of time. In particular, if the trend is linear, the deterministic component model for a monthly series is given by

$$y_t = \mu + \beta t + \sum_{i=1}^{11} \alpha_i D_{it} + u_t \tag{2.1}$$

where D_{it} are dummy seasonal variables taking value 1 when the observation t belongs to month i, -1 in December and zero otherwise and u_t is a stationary process. Notice that, in order to identify model (2.1), we are imposing the restriction that the seasonal component averaged over a year is zero. The noise, u_t, is assumed to be a stationary stochastic process. The dynamics implied by model (2.1) can be too rigid to represent the behaviour of most economic time series observed over a sufficiently large span of time.

2.2 Holt-Winters algorithm

The prediction of the unobserved components of a time series can be carried out by algorithms based on exponential smoothing which allow the components to evolve stochastically over time. These procedures have the attraction that they allow the forecast of each component to be updated very easily each time a new observation becomes available. To estimate the components at time *t*, these algorithms put relatively more weight in the most recent observations discounting past observations.

In this context, time series with trend and seasonal effects may be handled in two ways depending on whether the trend and seasonal components are thought to combine in an additive or multiplicative fashion. Given that the time series we are considering has been transformed taking logarithms, in this section we are describing the additive version. The interested reader may refer to Abraham and Ledolter (1983) for a description of the multiplicative procedure. For monthly time series with stochastic trend and seasonal components, the estimates of the level of the series, m_t, the slope of the trend, b_t, and the seasonal component, d_t, are given by the following Holt-Winters equations:

$$m_t = \lambda_1 (y_t - d_{t-12}) + (1 - \lambda_1)(m_{t-1} + b_{t-1}) \tag{2.2a}$$
$$b_t = \lambda_2 (m_t - m_{t-1}) + (1 - \lambda_2) b_{t-1} \tag{2.2b}$$
$$d_t = \lambda_3 (y_t - m_t) + (1 - \lambda_3) d_{t-12}, \tag{2.2c}$$

respectively, where λ_1, λ_2 and λ_3 are smoothing constants such that $0<\lambda_i \leq 1$, i=1,2,3. As λ_i tends to 0 the corresponding component is closer to be deterministic and all observations have similar weights in its estimation. As λ_i gets closer to 1, the discount of past observations increases and, in the limit, when λ_i=1, only the last observation is used to estimate the corresponding component at time *t*. These smoothing constants can be estimated by minimising the within-sample sum of squares of the one-step-ahead forecast errors. An obvious weakness of the Holt-Winters procedure is that each seasonal component is only updated every twelve periods and, consequently, in (2.2a) the estimated level is corrected by a seasonal component which is 12 periods out of date.

The Holt-Winters equations in (2.2) are extremely simple to apply and consequently their use is quite extended between practitioners. However, as they are not based on a statistical model, it is not possible to extend their use to estimate the trend and seasonal components in a model with explanatory variables. In particular, it is not possible to carry out an intervention analysis. Furthermore, it is not possible to construct prediction intervals for future values of the series.

2.3 ARIMA models

The autoregressive-moving average (ARMA) class of models provides a parsimonious representation of any stationary stochastic process. Although most economic time series are not stationary, ARMA models can be extended to encompass a much wider class of non-stationary series with trend and seasonal components. If after taking d differences a time series has an ARMA(p,q) representation, the series is said to follow an autoregressive-integrated-moving average process of order (p,d,q), denoted by ARIMA(p,d,q).

The multiplicative seasonal ARIMA$(p,d,q)\times(P,D,Q)_{12}$ model is a very general model to represent the dynamic behaviour of a wide number of monthly economic time series which exhibit long-run trend and seasonal movements. The model is given by

$$\Phi_p(L)\ \Phi_P(L^{12})\ \Delta^d \Delta^D{}_{12} y_t = \Theta_q(L)\ \Theta_Q(L^{12})\ a_t \tag{2.3}$$

where all the roots of the polynomials $\Phi_p(L)$, $\Phi_P(L^{12})$, $\Theta_q(L)$ and $\Theta_Q(L^{12})$ lie outside the unit circle and a_t is a white noise process with variance σ_a^2. One of the most useful models within the class of multiplicative ARIMA models is the "airline model" which was originally fitted to a monthly time series of airline passengers in the United Kingdom by Box and Jenkins (1976). The model is given by

$$\Delta\Delta_{12} y_t = (1-\theta_1 L)(1-\theta_{12} L^{12}) a_t. \tag{2.4}$$

The autocovariance function of $\Delta\Delta_{12} y_t$ in the "airline model" is given by

$$\begin{aligned}
\gamma_0 &= (1+\theta_1^2)(1+\theta_{12}^2)\,\sigma_a^2 \\
\gamma_1 &= -\theta_1 (1+\theta_{12}^2)\,\sigma_a^2 \\
\gamma_j &= 0, \qquad j=2,\dots,10 \\
\gamma_{11} &= \theta_1\,\theta_{12}\,\sigma_a^2 \\
\gamma_{12} &= -\theta_{12}(1+\theta_1^2)\,\sigma_a^2 \\
\gamma_{13} &= \theta_1\,\theta_{12}\,\sigma_a^2 \\
\gamma_j &= 0, \qquad j \geq 14.
\end{aligned} \tag{2.5}$$

Time series modelled by an ARIMA model can also be decomposed into components such as trend and seasonal. The advantages of the ARIMA modeling for signal extraction are, first, that it permits more flexibility than the fixed filters used by exponential smoothing procedures, as the Holt-Winters algorithm, and second, that the use of a statistical model offers a systematic framework for analysis. For example, the Holt-Winters equations in (2.2) are optimal to estimate the components of a series which follows a MA(13) model after beign differenced as $\Delta\Delta_{12} y_t$ (McKenzie, 1976). Later, Newbold (1988) showed that the parameters of the MA(13) model satisfy the following relationships:

$$\theta_1 = 1 - \lambda_1 (1 - \lambda_2) \tag{2.6a}$$
$$\theta_2 = \dots = \theta_{11} = -\lambda_1 \lambda_2 \tag{2.6b}$$
$$\theta_{12} = 1 - \lambda_1 \lambda_2 - (1 - \lambda_1)\lambda_3 \tag{2.6c}$$
$$\theta_{13} = -(1 - \lambda_1)(1 - \lambda_3) \tag{2.6d}$$

From equations (2.6), it is possible to see that the "airline model" results whenever both $\lambda_1\lambda_2$ and $\lambda_1\lambda_3$

are negligibly small.

However, ARIMA models still have some disadvantages when trying to estimate the unobserved components of a time series. Maravall (1985) points out that sometimes the models identified following the usual criteria offer unsatisfactory decompositions into components. In this sense, certain restrictions must be placed on the parameters of an ARIMA model for the decomposition into trend, seasonal and irregular components to exist (Hillmer and Tiao, 1982). Also, it is possible to obtain models with complicate analytical expressions for the components. Finally, it may be desirable to avoid the identification stage.

2.4 Structural time series models

A structural time series model is one which is set up in terms of components which have a direct interpretation. A univariate structural model aims to represent the "stylised facts" of a series in terms of a decomposition into components such as trend and seasonal. Following Engle (1978) and Nerlove *et al*. (1979), a particular dynamic structure is imposed on the components. The components of a structural model are often modelled as stochastic processes, and therefore each component may be driven by a different disturbance. As a consequence, several disturbances are involved in a structural model. Structural models have the limitation that the structure of the series is subject to certain constrains although they are more flexible than the exponential smoothing techniques. Their advantages versus ARIMA models are, first, that they skip the identification stage. Second, they provide components expected to behave properly as trend and seasonals. Furthermore, they provide estimates of the mean square error (MSE) of the estimated components. Harvey and Tood (1983) and Maravall (1985) compare the theoretical limitations and advantages of both ARIMA and structural models.

One useful model to represent the dynamic behaviour of time series with trend and seasonal component is the **basic structural model** (BSM). The BSM regards the observations, y_t, t=1,...,T, as being made up of an underlying stochastic trend, denoted by μ_t, a seasonal component, denoted by δ_t, and an irregular or transitory component denoted by ε_t. The BSM for monthly series is given by

$$y_t = \mu_t + \delta_t + \varepsilon_t \tag{2.7a}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{2.7b}$$

$$\beta_t = \beta_{t-1} + \xi_t \tag{2.7c}$$

$$\delta_t = -\sum_{i=1} \delta_{t-i} + \omega_t \tag{2.7d}$$

where ε_t, η_t, ξ_t and ω_t are assumed to be mutually independent white noise processes with variances σ^2_ε, σ^2_η, σ^2_ξ, and σ^2_ω respectively. The esential feature of the BSM is that both, the level and slope of the trend, change slowly over time according to a random walk mechanism. The seasonal pattern is also slowly changing but by a mechanism that ensures that the sum of the seasonal components over any year has an expected value of zero. Harvey (1984) proposes an alternative trigonometric form for the seasonal component which may have some advantages. The stochastic trigonometric seasonal component at time *t* is given by

$$\delta_t = \sum_{j=1}^{6} \delta_{jt} \tag{2.8a}$$

where

$$\delta_{jt} = \delta_{jt-1} \cos \lambda_j + \delta^*_{jt-1} \sin \lambda_j + \omega_t \tag{2.8b}$$
$$\delta^*_{jt} = -\delta_{jt-1} \sin \lambda_j + \delta^*_{jt-1} \cos \lambda_j + \omega^*_t. \tag{2.8c}$$

$\lambda_j = \frac{2\pi j}{12}$, j=1,...,6, are the seasonal frequencies and ω_{jt} and ω^*_{jt} are white noise processes mutually uncorrelated with common variance σ^2_ω for j=1,...,6. The larger this variance, the more past observations are discounted in estimating the seasonal pattern. The element δ^*_{jt} appears as a matter of construction. It is possible to assign different variances to each harmonic allowing them to evolve at varying rates. However, from a computational point of view, it is usually desirable to let these variances be the same.

The stationary form of the BSM with seasonal dummies is given by

$$\Delta\Delta_{12}\, y_t = S(L)\, \xi_{t-1} + \Delta_{12}\, \eta_t + \Delta^2\, \omega_t + \Delta\Delta_{12}\, \varepsilon_t, \tag{2.9}$$

where $S(L)=1+L+...+L^{11}$. Using the trigonometric formulation of the seasonal component, the stationary form is more complicated (Harvey, 1989). From (2.9), it is easy to see that $\Delta\Delta_{12}y_t$ has zero mean and its autocovariance function (*acf*) is given by

$$\begin{aligned}
\gamma_0 &= 12\sigma^2_\xi + 2\sigma^2_\eta + 6\sigma^2_\nu + 4\sigma^2_\varepsilon \\
\gamma_1 &= 11\sigma^2_\xi - 4\sigma^2_\nu - 2\sigma^2_\varepsilon \\
\gamma_2 &= 10\sigma^2_\xi + \sigma^2_\nu \\
\gamma_k &= (12-k)\sigma^2_\xi, \qquad 3 \le k \le 10 \\
\gamma_{11} &= \sigma^2_\xi + \sigma^2_\varepsilon \\
\gamma_{12} &= -\sigma^2_\eta - 2\sigma^2_\varepsilon \\
\gamma_{13} &= \sigma^2_\varepsilon \\
\gamma_k &= 0, \qquad k > 13.
\end{aligned} \tag{2.10}$$

Harvey (1989) derives the autocovariance function of the BSM with trigonometric seasonality for quarterly data.

The *acf* in (2.10) is the *acf* of a MA(13) process with restrictions on the parameters. In many empirical applications σ^2_ξ=0; see, for example, Harvey and Tood (1983). In this case, the reduced form of the BSM is very similar to the "airline model" (Maravall, 1985 and Harvey, 1989). Ansley (1983) also shows how a modification of the BSM may approximate the "airline model". Consequently, the dynamic properties of a series implied by the "airline model", the BSM and the Holt-Winters algorithm are very similar.

Model (2.7) is cast in what is known as state space form and, therefore, the Kalman filter can be used to get optimal one-step-ahead estimates of the unobserved components of the series. Under conditional normality of the series, these estimates of the components are optimal in the sense of minimising the MSE.

3. PREDICTION OF UNIVARIATE TIME SERIES

3.1 Prediction theory

Let $\{y_t\}$ be a univariate stochastic process which has been observed from t=1 to T. Suppose that one wants to predict k periods ahead, i.e. to obtain predictions of y_{T+k}, k=1,2,..., using the information contained in past values of the series up to time T. It is well-known that the conditional expectation is the optimal forecast in the sense that it has the smallest MSE. Therefore, if we denote by $\hat{y}_{T+k}$, the forecast of y_{T+k} based on information available at time T, the optimal predictor is given by the conditional expectation

$$\hat{y}_{T+k} = E(y_{T+k}/y_1,...,y_T). \tag{3.1}$$

Assuming gaussianity of the process $\{y_t\}$, the conditional expectation in (3.1) is a linear combination of the information set, and consequently, we can focus on linear predictors.

If the ARIMA (p,d,q) or its seasonal variant, has been fitted to a time series y_t, t=1,...,T, then forecasts of future values can be obtained from the following expression:

$$\hat{y}_{T+k} = \phi_1^* \hat{y}_{T+k-1} + ... + \phi_{p+d}^* \hat{y}_{T+k-p} - \theta_1 \hat{a}_{T+k-1} - ... - \theta_q \hat{a}_{T+k-q} \tag{3.2}$$

where $\phi_{p+d}^*(L) = \Delta^d \phi_p(L)$ and

$$\hat{y}_{T+j} = \begin{cases} \hat{y}_{T+j}, & j > 0 \\ y_{T+j}, & j \leq 0, \end{cases}$$

$$\hat{a}_{T+j} = \begin{cases} 0, & j > 0 \\ a_{T+j}, & j \leq 0. \end{cases}$$

In practice, if the model has a MA component the a_{T+j}, $j \leq 0$ are not known but can be estimated.

The variance of the forecast error, $e_{T+k}=y_{T+k}-\hat{y}_{T+k}$, can be obtained using the infinite MA representation of the model as

$$E(e_{T+k})^2 = \sigma_a^2 \sum_{i=0}^{k-1} \psi_i^2 ; \tag{3.3}$$

see, for example, Granger and Newbold (1986). Consequently, assuming normality of forecast errors, a 95% prediction interval is given by

$$\hat{y}_{T+k} \pm 1.96 \sigma_a \left(\sum_{i=0}^{k-1} \psi_i^2 \right)^{\frac{1}{2}} . \tag{3.4}$$

In practice, as the parameters of the model are unknown, one need to estimate them and the expression for the prediction error variance in (3.3) has to be modified to take into account parameter uncertainty.

Forecast of future values based on the deterministic model are given by

$$\hat{y}_{T+k} = \mu + \beta(T+k) + \sum_{i=1}^{11} \alpha_i D_{it} + \hat{u}_{T+k} \tag{3.5}$$

where $\hat{u}_{T+k}$ are the forecasts of future values of the perturbation which depend upon the stationary process followed by u_t. If u_t follows an ARMA process its predictions are given by an expression similar to (3.2). The MSE of e_{T+k} depends on the MSE corresponding to the prediction of u_{T+k}.

If the Holt-Winters algorithm has been used to estimate the unobserved components of a series, running equations (2.2) recursively from t=13 to T, forecasts of future values can be obtained by extrapolating the estimates of the components at time T into the future:

$$\begin{aligned} \hat{y}_{T+k} &= m_T + b_T k + d_{T+k-12}\,, \quad k = 1, ..., 12 \\ &= m_T + b_T k + d_{T+k-24}\,, \quad k = 13, .., 24 \\ &\vdots \end{aligned} \tag{3.6}$$

Given that the exponential smoothing methods are not based on a statistical model, there are not analytical expressions for the MSE corresponding to the forecast errors from these algorithms.

Finally, forecast of future values given by the BSM model can also be obtained as in (3.6) by extrapolating the estimates of μ_t, β_t and δ_t available at time T into the future. This can easily be done by repeatedly applying the prediction equations of the Kalman filter without using the updated equations. Also, the Kalman filter prediction equations give an expression of the MSE of the predictions (Harvey, 1989).

3.2 Forecast evaluation

The variance of the one-step-ahead prediction errors, can be used as a basic measure of goodness of fit within sample. Since, in this paper, several specifications are tryed, there is the danger of "data mining". Consequently, we use post-sample observations as a yardstick by which to judge the forecasting accuracy of the models estimated.

The post-sample one-step-ahead prediction errors can be used to compute the post-sample predictive test statistic given by

$$\xi(\tau) = \frac{\sum_{j=0}^{\tau-1} e^2_{(T+j)+1}}{\tau \tilde{\sigma}^2_a} \tag{3.7}$$

where $\tilde{\sigma}_a^2$ is the estimated variance of the within sample innovations and τ is the number of post-sample predictions. If the model is correctly specified and assuming normality of the forecast errors, Box and Tiao (1976) show that the statistic $\xi(\tau)$ has a $F(\tau,T-r)$ distribution where r is the number of observations used as initial conditions in the estimation procedure.

Also, we may compute the *CUSUM* quantities which are defined as

$$CUSUM(T,h) = \tilde{\sigma}_*^{-1} \sum_{j=0}^{h-1} e_{(T+j)+1}, \quad h=1,...,\tau, \tag{3.8}$$

where $\tilde{\sigma}_*$ is the sample standard deviation of the standardized innovations; see, for example, Harvey (1989). The *CUSUM* plot is valuable for detecting structural changes, depending on whether or not it crosses either of two predefined lines given by $\pm[a\sqrt{\tau} + 2ah/\sqrt{\tau}]$, where a=0.948 for a significance level of 5%. The *CUSUM* is particularly useful when the model is systematically under or overpredicting in the post-sample period.

In practice, it is desirable to subject models to post-sample multi-step predictive testing; see the paper by Clements and Hendry (1993) and the discussion of it. Looking at predictions several steps ahead is useful to check that the form of the forecast function is sensible, although the corresponding errors are not independent of each other and a valid post-sample predictive test would have to take account of this dependence.

3.3 Forecast comparisons

We compare the forecasting performance of the univariate models comparing both one-step-ahead and multi-step-ahead predictions in terms of some usual measures as their mean, mean absolute error (MAE), minimum and maximum absolute error and the root MSE (RMSE). If the predictor is optimum it should be unbiased and its MSE should be minimum. Consequently, we are, in an initial stage, rejecting models with biased predictions. Then, we compare models with unbiased forecasts by means of RMSE. We will also compare the forecast errors which are unbiased using the following tests proposed by Diebold and Mariano (1995):

$$S_{2a} = \frac{\sum_{i=1}^{\tau} I_+(d_i) - 0.5\tau}{\sqrt{0.25\tau}} \tag{3.9a}$$

$$S_{3a} = \frac{\sum_{i=1}^{\tau} I_+(d_i)\, rank(|d_i|) - \frac{\tau(\tau+1)}{4}}{\sqrt{\frac{\tau(\tau+1)(2\tau+1)}{24}}} \tag{3.9b}$$

where d_i is the difference between the one-step-ahead forecast errors made with each model and $I_+(d_i)$

is an indicator function which takes value one when d_i is positive and cero otherwise. Both statistics have an asymptotic standard normal distribution.

Following Fair and Shiller (1990), we also test for forecast encompassing which concerns whether the one-step ahead forecasts of one model can explain the forecast errors made by another (Chong and Hendry, 1986). As we mentioned before, there is autocorrelation between successive step ahead forecast errors and, consequently, forecast encompassing can not be made operational for multi-step-ahead forecasts. The test we are carrying out is based on the following regressions:

$$y_{(T+i)+1} - \hat{y}_{(T+i)+1(1)} = \alpha_1 + \beta_1 (\hat{y}_{(T+i)+1(2)} - \hat{y}_{(T+i)+1(1)}) + u_{1i} \qquad (3.10a)$$

$$y_{(T+i)+1} - \hat{y}_{(T+i)+1(2)} = \alpha_2 + \beta_2 (\hat{y}_{(T+i)+1(1)} - \hat{y}_{(T+i)+1(2)}) + u_{2i}, \qquad (3.10b)$$

where i=1,..., τ, $\hat{y}_{(T+i)+1(1)}$ denotes the one-step-ahead prediction made with model 1 and $\hat{y}_{(T+i)+1(2)}$ is the one-step-ahead prediction made with model 2.

Finally, we consider whether forecast combination can be of any help in improving the forecast performance of the models considered. To estimate the parameters of the combination, we are considering the following regression:

$$y_{(T+i)+1} = \alpha_0 + \delta_1 \hat{y}_{(T+i)+1(1)} + \delta_2 \hat{y}_{(T+i)+1(2)} + u_{0i}, \quad h=1,...,\tau. \qquad (3.11)$$

If both predictors are unbiased, δ_2=1-δ_1. In this case, model 1 is favoured if δ_1=0, model 2 if δ_1=0 and neither model if δ_1 is neither zero nor unity. A predictor is conditionally efficient if the variance of the noise in (3.11) is not smaller than the variance of their own predictions.

4. EMPIRICAL ANALYSIS

4.1 The data

The time series under study consists of monthly observations of the number of passengers in the Spanish airline IBERIA from January 1985 to October 1994. The observations, after been transformed into logarithms, appear in figure 1. A visual inspection of the data suggests that the trend component presents a rupture during 1990 with its slope changing sign from positive to negative. It is also apparent that the dynamic behaviour of the series is characterized by seasonal oscilations. Moreover, during the sample period under study there have been several extraordinary events which affected temporaly the level of the series. First, in January 1989 the ASETMA (mechanic's union) went on strike and the level of the series was reduced up to February. Secondly, after February 1991, due to the Golf War, the number of passengers dropt drastically. The war seems to affect the series up to May 1991. Finally, during the Summer of 1992, there were two exceptional events in Spain, the Olympic Games at Barcelona and the Universal Exposition at Seville which increased the level of the series during July, August and September. The effects on the level of the series of all of these extraordinary events have been modelled using intervention variables. The intervention variables considered in this paper are pulse variables. See Box and Tiao (1975) for a detailled description of intervention analysis in univariate ARIMA models. Harvey (1989) gives details on the especification, estimation and diagnosis of intervention effects on structural time series models.

There is another deterministic effect related with the Easter vacation which affects the level of the series. The presence of the Easter holiday in a particular month has a positive effect over the number of passengers in this month. Consequently, we include a dummy variable having value zero except in March and April when it takes the value of the proportion of the Easter vacation in each month.

4.2 Estimated models

In order to analyse the predictive performance of each of the methods described in section 2 and their ability to detect the slope change, each model has been fitted to the observations up to December 1992. The last 22 observations, corresponding to 1993 and part of 1994, have been kept to make post-sample comparisons. The deterministic and ARIMA models as well as the Holt-Winters recursions have been estimated using the SCA program. The structural time series models have been estimated using the STAMP program.

a. <u>Deterministic model (DET)</u>

Fitting model (2.1) to the IBERIA series, the estimates of the parameters μ and β are highly negatively correlated (-0.89). This could be due to the fact that the trend in the series is changing at the end of the sample period and, consequently, a deterministic model with a constant slope may be inappropiate to represent the dynamics of the series. Alternatively, we consider trends with polynomials of order greater than one in time but the results were not significant. Consequently, we try a deterministic model with β fixed to be zero, i.e. the series has irregular and seasonal oscilations around a constant level. The estimated parameters together with some diagnostics and measures of fit appear in table 1. It is worth noting that the estimate of the autoregressive parameter of order one is rather close to unity, suggesting the possible presence of a unit root. We could not find any explication for the residuals corresponding to December 1986, March 1987 and July 1991 which are greater than 2.5 $\hat{\sigma}_a$.

Figure 2.a represents the *CUSUM* quantities of the within sample innovations of the deterministic model. The *CUSUM* are well within the significance lines and, consequently, they are not detecting any structural change in the data.

In practice, as the series analysed in this paper has been transformed into logarithms we are computing the forecast errors as

$$e_{T+k} = \frac{\exp(y_{T+k}) - \exp(\hat{y}_{T+k})}{\exp(y_{T+k})}.$$

In panels a to d of figure 3 we represent the *k*-step-ahead prediction errors obtained using the deterministic model to predict values of y_{T+k} for k=1,3,6 and 12 respectively. Looking at these plots, it is possible to observe that the deterministic model systematicaly over predict the true values when predictions are made for k=3, 6, 12. In figure 4, we represent the *CUSUM*s of the one-step-ahead forecast errors. We can observe that these quantities cross the significance lines by July 1993 showing that the model is not any longer appropriate. In table 4, we report some measures of the forecast errors. Looking at the value of the statistic $\xi(22)$ we reject the correct specification of the model. Also, we may observe that the predictions have a negative significant bias which goes from a 2% when predicting one month ahead to a 9% when predicting one year ahead.

Consequently, although the within sample fit of the deterministic model was adequate, the predictions are biased and we reject the model.

b. Holt-Winters algorithm (HW)

The smoothing constants which minimise the sum of squares of the within sample one-step-ahead prediction errors of the Holt-Winters algorithm in (2.2) applied to the IBERIA series are λ_1=0.51, λ_2=0.01 and λ_3=0.45. The value of λ_2 implies a slope of the trend close to be constant over time which is in corcondance with the results for the deterministic model. The *CUSUM* quantities represented in figure 2.b do not give any indication of structural change.

The prediction errors made by using the Holt-Winters algorithm with such smoothing constants have been plotted in figure 3. These errors are, in general, greater than those obtained using the deterministic model. However, except for k=12, the Holt-Winters predictions are not systematicaly overpredicting future values of y_t. In figure 4 where we represent the *CUSUM*s of the one-step-ahead prediction error, we can observed that although these quantities are close, they do not cross the significance lines. In table 4, we may observe that the $\xi(22)$ statistic does not reject the model. However, the Holt-Winters predictions are bised and worse than the deterministic predictions for all the other criteria reported. The bad behavior of the Holt-Winters forecasts could be due to the presence of the extraordinary events in the sample used to estimate the smoothing constants. As we noted before, the Holt-Winters algorithm is not able to cope with explanatory variables such as interventions and during the sample period analysed, there are important extraordinary events which affect fundamentaly the levels of the series.

c. Airline model (ARIMA)

The results of fitting a multiplicative ARIMA$(0,1,1)\times(0,1,1)_{12}$ model with interventions and the Easter variable to the IBERIA series appear in table 2. Notice that the estimates of the interventions and the Easter effect are rather similar to the ones obtained in the deterministic model. The standard deviation of the innovations of the "airline model" is greater than the one obtained for the deterministic model. We obtain residuals greater than $2.5\hat{\sigma}_a$ at the same dates as in the deterministic model plus one in January 1991. Once more, the *CUSUM* of the residuals, in figure 2.c, are well inside the confidence lines showing no indication of structural change.

Figure 3 represents the prediction errors made by using the ARIMA model to forecast future values of y_t. Observing such figure, it seems rather clear that the predictive performance of this model is quite high compared with the deterministic model. The prediction errors are well behaved for all forecast horizons considered in this study. Even for k=12 the forecast errors are remarkably small. Figure 4 represents the *CUSUM*s of the one-step-ahead prediction errors, which are well inside the significance lines. Looking at table 4, it is possible to observe that both, the bias and the RMSE, have been reduced for all forecast horizons. The predictions are unbiased for all horizons and the reduction in RMSE is more important as k increases. As a consecuence, it seems that the airline model with interventions has been able to represent adecuately the change in the slope of the long-run trend of the IBERIA series being able to give good forecasts even one year ahead.

We also reestimate the model using rolling estimates which use a fixed sample size, implicitly allowing for changes in the parameters of the model, and forecasting with the resulting estimates. The rolling forecasts[4] do not improve significantly the ones obtained without updating the parameter estimates. Consequently, it seems that there is parameter stability and, as expected, in these circumstances there is not forecast improvement updating the estimates of the parameters. We can use the estimates obtained with the observations up to December 1992 without need to reestimate the model each time a new observation is available.

d. Basic structural model (BSM)

Finally, we fit the BSM with interventions and the Easter variable to the IBERIA series up to December 1992. The residuals from this fit have significant autocorrelations at lags 1, 3 and 4. Consequently, we add a cycle to the BSM. The cycle, denoted by ψ_t, has the following specification

$$\psi_t = \rho\, \psi_{t-1} \cos \lambda_c + \rho\, \psi^*_{t-1} \sin \lambda_c + \kappa_t \tag{4.1a}$$

$$\psi^*_t = -\rho\, \psi_{t-1} \sin \lambda_c + \rho\, \psi^*_{t-1} \cos \lambda_c + \kappa^*_t \tag{4.1b}$$

where $0 \leq \rho \leq 1$ and κ_t and κ^*_t are white noise disturbances mutually uncorrelated and with the same variance. ψ^*_t appears by construction in order to form ψ_t.

The estimation results appear in table 3 where we may observe that they are rather similar to the ones obtained for the "airline model" in table 2. Notice that the standard deviation of the innovations is slightly bigger than the one obtained for the ARIMA model and, of course, bigger than in the deterministic model. However, as we will see latter the forecast performance of the BSM is not worse, being an example of possible "data mining". The estimates of σ^2_η and σ^2_ξ imply that the trend is smoothly evolving over time. The seasonal and cyclical components are also stochastic. With respect to the cycle, the period corresponding to the estimated frequency is around 15 months. There are two extreme observations corresponding to January and July of 1991 which also appear in the ARIMA model. As we could not find any convincing explanation for these observations, and including intervention variables for them in the model makes the Easter variable not significant, we decide not to include them in the model. Figure 2.d represents the *CUSUM*s of the innovations of the BSM model. Once more, there is not evidence of any structural change during the sample period.

Looking at figures 3a and 4, we can observe that the one-step-ahead errors from the ARIMA and the BSM models have very similar behaviour. When forecasting 3, 6 and 12 steps ahead the BSM forecasts also have very similar patterns to the ARIMA forecasts. In table 4, we may see that both sets of errors have also very similar properties.

---

4. Available from the authors on request.

4.3 Forecasting performance

The predictions obtained with the deterministic and Holt-Winters models are biased for different reasons. Although the deterministic model has the best within sample fit measured in terms of the standard deviation of the residuals, it is not flexible enough to updated to changes in the long-run trend of the series. On the other hand, the Holt-Winters algorithm is not able to deal properly with extraordinary events and this affects both the within sample fit and the predictions. However, notice that there is not indication of failure in the fit. In this subsection we are comparing the forecast performance of the ARIMA and BSM models which produce forecasts which are unbiased. First, computing the S_{2a} and S_{3a} statistics, they take values of -0.43 and -0.89 respectively, being not significant for any of the usual confidence levels. This means that there are not significant differences between the forecast errors made with the ARIMA model and those made with the structural model. Secondly, we carried out the encompassing regressions in (3.10) with similar results. Neither of the regressions have significant coefficients. Consequently, neither the one-step-ahead forecast errors of the structural model can explain the forecast errors made by the "airline model" nor the other way round.

Finally, we consider combinations of predictions by means of the regression in (3.11). Estimating by Ordinary Least Squares such regression, we obtain the following results:

$$\hat{y}_{(T+i)+1} = \underset{(1.34)}{0.947} + \underset{(1.19)}{0.275}\, \hat{y}_{(T+i)+1(ARIMA)} + \underset{(2.85)}{0.657}\, \hat{y}_{(T+i)+1(BSM)}$$

$$\hat{\sigma}_u = 0.024$$

Looking at these results it seems that the improvement in forecast performance made by linear combinations is rather small because the residual standard deviation has been reduced only marginaly. However, there is a slight tendency towards the predictions obtained with the BSM.

It seems that for the IBERIA series analysed in this paper, ARIMA and structural time series models have very similar prediction performance. However, as we point out before when describing the univariate models, the structural approach allows us to easily obtain, by means of the Kalman filter, estimates of the components of a time series. For the IBERIA series the estimated components appear in figure 5. In this figure we may observe that the change in the trend of the series is quite smooth, with a break in the slope around December 1989.

Finally, it is important to notice the fundamental role of the interventions when they are close to the end of the sample period. As we mentioned before, there are three extraordinary events affecting the IBERIA series during the period we consider in this paper. To analyze the effect of these interventions on the performance of the models, we reestimate and predict with the ARIMA and BSM models without the Golf War and the Olympic Games and Expo interventions. The results appear in table 5.

Comparing these results with the ones reported in table 4, we may observe that, although the forecast are still unbiased, the RMSE are clearly bigger, being close to the values obtained with the deterministic model. The similarity between the deterministic model and the stochastic models without interventions can be understood observing figure 6 which represents the components estimated with the BSM without interventions. The estimate of the trend does not show the change in the slope observed in figure 5.a, when the interventions are included in the model. This could be one of the reasons why the predictions made with the stochastic models without interventions tend to overpredict the observed values.

5. CONCLUSIONS

Newbold and Granger (1974) compare the performance of ARIMA models and various exponential smoothing predictors over a large set of real time series. They note a tendency of ARIMA models to lose some of its advantage over exponential smoothing methods when forecasts over longer lead times are considered. However, the analysis of the IBERIA series carried out in this paper shows that the forecasting performance of the Holt-Winters algorithm relies heavily on the value of the smoothing constants. If such smoothing constants are estimated minimizing the one-step-ahead errors, the estimates depend crucially upon the presence of outliers in the data. For the IBERIA series, the forecasts made using the Holt-Winters algorithm behave poorly using any of the usual forecast evaluation criteria.

We also show that the use of deterministic components for the trend and the seasonal of a time series can be very dangerous with systematic forecast errors even for short horizons. Finally, the forecasting performance of ARIMA and structural models is, in this case, rather similar with unbiased forecast errors. However it is worth noting that the within sample fit of the deterministic model is better than the fit of either of the stochastic models. This result points out the danger in using solely within sample measures of fit to compare models which are going to be used for forecasting.

After testing for forecast encompassing between the "airline model" and the BSM plus cycle model, we find out that none of the errors made with one model can explain the errors made with the other. Also, for the IBERIA series it seems that the combined forecast may not constitute a worthwhile improvement over the individual procedures. The pattern of forecast errors for the airline model and the BSM model are very similar.

Finally, the results for the IBERIA series shows the fundamental importance that the adequate modelization of interventions may have for the performance of the predictions, specially if these interventions occur near the end of the sample period.

It could be interesting to compare the forecasting performance of the unit root models considered in this paper with models with segmented trends as the ones proposed by, for example, Rappoport and Reichlin (1989).

REFERENCES

Abraham, B. and J. Ledolter, 1983, *Statistical Methods for Forecasting*. Wiley, New York.

Ansley, C.F., 1983, "Comment on 'Forecasting Economic Time Series with Structural and Box-Jenkins models: A case Study'", *Journal of Business & Economic Statistics*, 1(4), 307-309.

Box, G.E.P. and G.M. Jenkins, 1976, *Time Series Analysis, Forecasting and Control*, 2nd ed. Holden Day, San Diego.

Box, G.E.P. and G.C. Tiao, 1975, "Intervention Analysis with Applications to Economic and Enviromental Problems", *Journal of the American Statistical Association*, 65, 1509-1526.

Box, G.E.P and G.C. Tiao, 1976, "Comparison of Forecasts with Actuality", *Applied Statistics*, 25, 195-200.

Chong, Y.Y. and D.F. Hendry, 1986, "Econometric Evaluation of Linear Macroeconomic Models", *Review of Economic Studies*, 53, 671-690.

Clements, M.P. and D.F. Hendry, 1993, "On the Limitations of Comparing Mean Square Forecast Errors" (with discussion), *Journal of Forecasting*, 12, 617-676.

Diebold, F.X. and R.S. Mariano, 1995, "Comparing Predictive Accuracy", *Journal of Business & Economic Statistics*, 13, 253-263.

Engle, R.F., 1978, "Estimating Structural Models of Seasonality", in: A. Zellner ed., *Seasonal Analysis of Economic Time Series*. U.S. Dept. of Commerce, Bureau of the Census, Washington D.C., 281-297.

Fair, R.C. and R.J. Shiller, 1990, "Comparing Information in Forecast From Econometric Models", *American Economic Review*, 80(3), 375-389.

Granger, C.W.J. and P. Newbold, 1986, *Forecasting Economic Time Series*, 2nd. ed. Academic Press, San Diego.

Harvey, A.C., 1984, "A Unified View of Statistical Forecasting Procedures", *Journal of Forecasting*, 3, 245-275.

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

Harvey, A.C. and P.H.J. Tood, 1983, "Forecasting Economic Time Series with Structural and Box-Jenkins Models: A case Study" (with discussion), *Journal of Business & Economic Statistics*, 1(4), 299-315.

Hillmer, S.C. and G.C. Tiao, 1982, "An ARIMA-Model-Based Approach to Seasonal Adjustment", *Journal of the American Statistical Association*, 77, 63-70.

Maravall, A., 1985, "On Structural Time Series Models and the Characterization of Components", *Journal of Business & Economic Statistics*, 3, 350-355.

McKenzie, E., 1976, "A Comparison of Some Standard Forecaasting Systems", *The Statistician*, 25, 3-14.

Nerlove, M., D.M. Grether and J.L. Carvalho, 1979, *Analysis of Economic Time Series*. Academic Press, New York.

Newbold, P., 1988, "Predictions Projecting Linear Trend Plus Seasonal Dummies", *The Statistician*, 37, 111-27.

Newbold, P. and C.W.J. Granger, 1974, "Experience with Forecasting Univariate Time Series and the Combination of Forecasts", (with discussion), *Journal of the Royal Statistical Society*, A, 137, 131-49.

Rappoport, P. and L. Reichlin, 1989, "Segmented Trends and Non-Stationary Tine Series", *The Economic Journal*, 99 (Conference), 168-177.