

A model selection procedure for time series with seasonality

Philip Hans Franses

Econometric Institute, Erasmus University Rotterdam, Netherlands

Received May 1992

Abstract: In this paper a model selection test procedure for seasonal time series is proposed. It uses the estimated autocorrelations of the moving average part of the Box and Jenkins airline model. This ensures that the test statistics asymptotically follow standard normal distributions. The merits and limitations of the procedure are illustrated via simulations as well as by some empirical series.

Keywords: Time series; seasonality; model selection.

1. Introduction

There are several models which characterize a (non-) stationary univariate time series with seasonality. Examples are (a) models which assume the presence of seasonal unit roots, (b) models in which seasonality is deterministically modelled, and (c) the airline model proposed in Box and Jenkins (1970). The latter model assumes that a seasonally and first order differenced series can be described by a multiplicative moving average process. Recent contributions to the search for an appropriate model selection strategy are Dickey, Hasza and Fuller (1984), Osborn (1990), and Hylleberg et al. (1990). A characteristic of these approaches is that test procedures are developed for a certain nonstationary null hypothesis. This implies that only a subset of possible models is considered, and that the empirical distributions of the test statistics have to be tabulated. In the present paper an alternative model selection procedure is proposed which tries to cope with these difficulties.

Correspondence to: Philip Hans Franses, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, Netherlands.

In Section 2, it is argued that the airline model is linked with several common univariate models for seasonal time series. In Section 3, the estimated autocorrelations of the doubly differenced variable are used to construct test statistics. It is shown that the test statistics asymptotically follow standard normal distributions under the null hypotheses. In Section 4, the test procedure is evaluated by simulation experiments, and applied to some empirical time series. Section 5 contains the concluding remarks.

2. Models for time series with seasonality

Consider n observations on the possibly nonstationary process y_t , which is a univariate time series measured s times a year. Define ε_t as an uncorrelated zero mean process with variance σ_ε^2 , and let there be seasonal dummies D_{1t}, \dots, D_{st} which take a value of 1 in the corresponding season every year, and 0 in other periods. The airline model proposed in Box and Jenkins (1970), is

$$(1 - B)(1 - B^s)y_t = (1 + \theta_1 B)(1 + \theta_s B^s)\varepsilon_t, \quad (1)$$

where B is defined by $B^k y_t = y_{t-k}$. In many applications this model appears to be a reasonable approximation to empirical data, although the parameters θ_1 and/or θ_s may sometimes be close to -1 , see also Harvey (1984). This may indicate that one or both of the differencing filters on the left hand side of (1) is inappropriate. Other simple models, see also Ghysels (1990), are

$$(1 - B)(1 - B^s)y_t = \eta_{1t}, \tag{2}$$

$$(1 - B^s)y_t = \eta_{2t}, \tag{3}$$

$$(1 - B)y_t = \sum_{i=1}^s \alpha_i D_{it} + \eta_{3t}, \tag{4}$$

$$y_t = \sum_{i=1}^s \alpha_i D_{it} + \eta_{4t}, \tag{5}$$

where the error processes η_{it} , $i = 1, 2, 3, 4$, are stationary and invertible ARMA processes. Model (5) may also contain a linear deterministic trend.

The null hypothesis in Dickey, Hasza and Fuller (1984) is (3), and it is tested against a stationary AR model, $y_t = \phi y_{t-s} + \eta_{2t}$. The null model in Hasza and Fuller (1982) is given by (2), and versions of (3), (4) and (5) are the alternatives. A test procedure applied in Osborn (1990) considers similar null and alternative models. The vast literature on testing for unit roots in nonseasonal time series considers (4) against (5), where the latter model may contain a linear trend. The procedure proposed in Hylleberg et al. (1990) compares model (3) with (4) and (5) and some intermediate cases. From this brief summary it is clear that these procedures test distinct nonstationary null hypotheses. First, this implies that the empirical distributions of the test statistics have to be tabulated since these models are nonstationary. Second, the practical application of the tests can give rise to conflicting empirical evidence since the models are nonnested, see also Ghysels (1990) and Osborn (1990).

In some occasions there is however a possibility to circumvent these drawbacks. As already said, the airline model in (1) appears to be useful in many practical situations, albeit that sometimes the estimated θ_1 and/or θ_s may approach minus 1. This suggests that model selection with respect to the first and annual differencing filter can be carried out by checking the values of θ_1

and θ_s . In particular, assuming that the η_{it} in (2) through (5) can be restricted to be ε_t , it is clear that when $\theta_1 = \theta_s = 0$ model (1) reduces to (2). Restricting θ_1 in (1) to minus 1 gives a model like (3). In Bell (1987) it is shown that when in

$$(1 - B^s)y_t = (1 + \theta_s B^s)\varepsilon_t, \tag{6}$$

the value of the parameter θ_s is equal to minus 1, a model with seasonal dummy variables emerges. Of course, similar results apply when y_t is replaced by $x_t = (1 - B)y_t$. This implies that with $\theta_s = -1$ one obtains a model like (4), and with $\theta_1 = \theta_s = -1$ model (5) shows up. Hence it can be seen that, in some specific cases, the airline model is linked with several other seasonal time series models.

3. The model selection procedure

A natural model selection procedure is now given by testing whether $\theta_1 = -1$ and/or $\theta_s = -1$ in (1). These parameter values imply that the airline model is noninvertible, or, equivalently, that there are unit roots in the MA part of the model. This noninvertibility has a downward biasing effect on the estimates of the parameters, see Plosser and Schwert (1977). So, model selection should preferably not be based on the estimated parameters. It may however be more appropriate to base inference on the sample autocorrelations of the variable

$$x_t = (1 - B)(1 - B^s)y_t. \tag{7}$$

The relevant nonzero autocorrelations of the airline model at lags 1, $s - 1$, s and $s + 1$, when $s > 2$, are given by

$$\rho_1 = \theta_1 / (\theta_1^2 + 1), \tag{8}$$

$$\rho_{s-1} = \rho_{s+1} = \theta_1 \theta_s / ((\theta_1^2 + 1)(\theta_s^2 + 1)) = \rho_1 \rho_s, \tag{9}$$

$$\rho_s = \theta_s / (\theta_s^2 + 1), \tag{10}$$

which are estimated by r_1 , r_{s-1} , r_s and r_{s+1} . These expressions show that it is convenient for the construction of test statistics to consider ρ_1 and ρ_s only, since they are uniquely determined by θ_1 and θ_s , respectively.

The distributional results for the sample autocorrelations of MA processes, given and proved in Anderson and Walker (1964), are now useful. Consider n observations on the zero mean linear process

$$w_t = \sum_{i=-\infty}^{+\infty} \eta_i \varepsilon_{t-i}, \quad t = 0, \pm 1, \pm 2, \dots, \quad (11)$$

where $\sum_{i=-\infty}^{\infty} |\eta_i| < \infty$ and $\sum_{i=-\infty}^{\infty} |i| \eta_i^2 < \infty$. It can be shown that $n^{1/2}(r_k - \rho_k)$, with $1 \leq k \leq s$, asymptotically follows an s -variate normal distribution with mean zero and with covariances given by

$$\begin{aligned} n \text{Cov}(r_k, r_j) &= \sum_{j=-\infty}^{+\infty} (\rho_j \rho_{j+k-l} + \rho_j \rho_{j+k+l} + 2\rho_k \rho_l \rho_j^2 \\ &\quad - 2\rho_k \rho_j \rho_{j+l} - 2\rho_l \rho_j \rho_{j+k}). \end{aligned} \quad (12)$$

Note that the restrictions for η_i apply in the MA cases considered here.

Applying (12) to the present MA model gives, after some algebraic manipulation, the following distributional results,

$$n^{1/2} \begin{pmatrix} r_1 \\ r_s \end{pmatrix} \sim N \left\{ \begin{pmatrix} \rho_1 \\ \rho_s \end{pmatrix}, \begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix} \right\}, \quad (13)$$

where

$$\begin{aligned} A_{11} &= 1 + 6\rho_{s-1}^2 + 8\rho_1^2 \rho_{s-1}^2 - 16\rho_1 \rho_{s-1} \rho_s + 2\rho_s^2 \\ &\quad + 4\rho_1^2 \rho_s^2 + 2\rho_1 \rho_{s-1} - 3\rho_1^2 + 4\rho_1^4, \end{aligned}$$

$$\begin{aligned} A_{12} = A_{21} &= 4\rho_{s-1} + 8\rho_1 \rho_{s-1}^2 \rho_s - 8\rho_{s-1} \rho_s^2 \\ &\quad - 8\rho_1^2 \rho_{s-1} + 4\rho_1^3 \rho_s - 2\rho_1 \rho_s + 4\rho_1^3 \rho_s, \end{aligned}$$

$$\begin{aligned} A_{22} &= 1 + 6\rho_{s-1}^2 + 8\rho_{s-1}^2 \rho_s^2 - 16\rho_1 \rho_{s-1} \rho_s + 4\rho_s^4 \\ &\quad - 3\rho_s^2 + 2\rho_1^2 + 4\rho_1^2 \rho_s^2. \end{aligned}$$

Under the null hypothesis that $\theta_1 = -1$ it can easily be verified that $\rho_1 = -\frac{1}{2}$, that $\rho_{s-1} = -\frac{1}{2}\rho_s$, and that A_{11} equals $\frac{1}{2} + \rho_s^2 + \frac{1}{2}\rho_s^4$. The most likely range of values of ρ_s is between $-\frac{1}{2}$ and 0, hence implying that in this case A_{11} theoretically varies from $\frac{1}{2}$ and $\frac{7}{16}$. This last value is the minimum of A_{11} established by $\rho_s = -\frac{1}{4}$. Under the null hypothesis that $\theta_s = -1$ it can similarly be derived that A_{22} becomes equal to

$\frac{1}{2} + \rho_1^2$, the range of which is from $\frac{1}{2}$ to $\frac{3}{4}$. Hence, the variance of each estimated autocorrelation under either of the null hypotheses contains a function of the other true but unknown autocorrelation. One might now consider estimates of the latter, and use these to estimate the variances. The risk involved in this approach is that in small samples the asymptotic results may not be valid. However, considering the ranges of values for A_{11} and A_{22} , a simple alternative strategy may be to set the variances at their highest values. Note that this is equivalent to testing $\theta_1 = -1$ while assuming that $\theta_s = -1$, and vice versa. This may ensure a somewhat conservative procedure in the sense that the null hypothesis may now be accepted more often. Given the results in, e.g., Franses (1991) and Osborn (1990), where it is shown that incorrect differencing can have a substantial impact on several types of inference, it is not harmful to be somewhat conservative in this case. Setting $\rho_1 = \rho_s = -\frac{1}{2}$ and thus $\rho_{s-1} = \rho_{s+1} = \frac{1}{4}$, (13) becomes

$$n^{1/2} \begin{pmatrix} r_1 \\ r_s \end{pmatrix} \sim N \left\{ \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & \frac{1}{8} \\ \frac{1}{8} & \frac{3}{4} \end{pmatrix} \right\}. \quad (14)$$

Simple test statistics for the null hypotheses $\theta_1 = -1$, $\theta_s = -1$, and $\theta_1 = \theta_s = -1$ are respectively given by

$$T_1 = (2n)^{1/2} (r_1 + \frac{1}{2}), \quad (15)$$

$$T_s = (\frac{4}{3}n)^{1/2} (r_s + \frac{1}{2}), \quad (16)$$

$$T_{1,s} = (\frac{2}{3}n)^{1/2} (r_1 + r_s + 1), \quad (17)$$

the asymptotic marginal distributions of which are $N(0, 1)$ distributions. In Franses and Kloek (1990) it is shown that the estimated first order autocorrelation of a noninvertible MA(1) process is indeed symmetrically distributed. Hence, the test statistics in (15) through (17) will be considered in a two-sided test procedure.

The first step in the model selection strategy for a seasonal time series y_t is now to transform it with the $(1 - B)(1 - B^s)$ filter into x_t . Then, the estimated autocorrelations of x_t at lags k other than 1, $s - 1$, s and $s + 1$ are checked to be zero, otherwise the airline model may not be

appropriate. It is easy to derive from (12) that for these k it holds that

$$n \text{ Var } r_k = \sum_{j=-\infty}^{+\infty} (\rho_j^2 + \rho_j \rho_{j+2k}). \tag{18}$$

Under the null hypothesis $\theta_1 = \theta_s = -1$ it can be derived that the expression in (18) is equal to $2\frac{1}{4}$, with some exceptions. For example, for quarterly data $n \text{ Var } r_2$ is $\frac{3}{4}$, and for monthly data $n \text{ Var } r_5$, $n \text{ Var } r_6$ and $n \text{ Var } r_7$ are equal to 2 , $\frac{3}{4}$ and 2 , respectively. When this null hypothesis is not valid, (18) can not be reduced to such simple expressions, and one strategy may now be to apply the rough approximation $n^{1/2}r_k \sim N(0, 1)$. The next step for model selection consists of testing whether ρ_1 and/or ρ_s are equal to $-\frac{1}{2}$, and to see which of the seasonal models may be appropriate.

4. Empirical performance

To investigate the performance of the above procedure, 100 Monte Carlo replications have been carried out for the case where $s = 4$ and $n = 100$. The rejection frequencies of the three tests are reported in Table 1. The first four data generating processes (DGPs) are

- (i) $y_t = y_{t-1} + y_{t-4} - y_{t-5} + \varepsilon_t$
($\theta_1, \theta_s \neq -1$, P, P),
- (ii) $y_t = y_{t-4} + \varepsilon_t$
($\theta_1 = -1$, S, P),
- (iii) $y_t = y_{t-1} + D_{1t} - 0.5D_{2t} + 1.5D_{3t} - D_{4t} + \varepsilon_t$
($\theta_s = -1$, P, S),
- (iv) $y_t = D_{1t} - 0.5D_{2t} + 1.5D_{3t} - D_{4t} + \varepsilon_t$
($\theta_1, \theta_s = -1$, S, S),

where the null hypothesis, as well as whether the outcomes for the DGP convey information about the empirical size (S) or power (P), are given between parentheses. From the results in Table 1 it is clear that for these DGPs the empirical size and power of the tests are almost ideal values. To illustrate that the first step of checking the auto-

Table 1

Rejection rates of testing for the noninvertibility of the airline model, based on 100 Monte Carlo replications and sample size 100

DGP	Nominal size	T_1	T_s	$T_{1,s}$
(i)	0.05	1.00 (P)	1.00 (P)	1.00 (P)
	0.10	1.00 (P)	1.00 (P)	1.00 (P)
(ii)	0.05	0.05 (S)	0.99 (P)	0.99 (P)
	0.10	0.11 (S)	1.00 (P)	0.99 (P)
(iii)	0.05	0.99 (P)	0.01 (S)	0.95 (P)
	0.10	1.00 (P)	0.01 (S)	0.98 (P)
(iv)	0.05	0.06 (S)	0.04 (S)	0.03 (S)
	0.10	0.11 (S)	0.11 (S)	0.09 (S)
(v)	0.05	0.50 (S)	0.99 (P)	0.89 (P)
	0.10	0.70 (S)	1.00 (P)	0.92 (P)
(vi)	0.05	0.08 (S)	0.13 (P)	0.11 (P)
	0.10	0.08 (S)	0.26 (P)	0.17 (P)

(S) and (P) correspond to the empirical size and power. Expressions for the test statistics can be found in (15), (16) and (17).

correlations at lags other than 1, $s - 1$, s and $s + 1$ is important, consider the DGP,

$$(v) \quad y_t = y_{t-4} + \varepsilon_t - 0.6\varepsilon_{t-1} \\ (\theta_1 = -1, S, P).$$

Doubly differencing results in $x_t = (1 - B)(1 - \delta B)\varepsilon_t$, which theoretically would show an auto-correlation at lag 2, and so the airline model is not appropriate. When δ equals 0.6, this correlation is about 0.153, and its estimate may often appear to be significant in case $n = 100$ (see (18)). From the fifth row of Table 1 it is clear that now the size is largely affected. Finally, when DGP is

$$(vi) \quad y_t = y_{t-4} + \varepsilon_t - 0.6\varepsilon_{t-4} \\ (\theta_1 = -1, S, P),$$

it can be expected that the power of the test for θ_s becomes low. This is reflected by the last row of Table 1.

To empirically illustrate the above test procedure, I have chosen to consider the seasonal series discussed in Granger and Newbold (1986), the Box and Jenkins (1970) airline data, the Dutch car sales series given in Franses (1990), the monthly car sales in the U.S., given in Nelson (1973), and the price of steers series for the U.S.

in Nerlove et al. (1979). The results are summarized in the Table 2. It appears that the airline model may be appropriate, at least in the sense that there is no autocorrelation other than those at lags 1, $s - 1$, s and $s + 1$, for 6 of the 7 series. Only the Dutch car sales series shows estimated values for r_2 , r_5 and r_9 of -0.28 , 0.24 and 0.24 , which may be significant under an hypothesis other than $\theta_1 = \theta_s = -1$. The GN110 series is clearly doubly over-differenced, and model (5), possibly with the inclusion of a trend, may be useful. For both the GN108 and GN106 series the double filter can be used, which corresponds to the successful estimation results for both series in Granger and Newbold (1986, p. 111). Curiously enough, the airline data themselves are over-differenced when the airline model is used, and the most suitable model seems to be a model of type (4). This corresponds with the findings in Franses (1991), where the Hylleberg et al. (1990) procedure is extended and applied to the case where $s = 12$. The FRCAR series illustrates the probable size distortions of the test statistics in case the error process in the airline model is not white noise. In Franses (1991) it is found that for this series a type (4) model with several ARMA terms is appropriate. Although the doubly differenced U.S. car sales series shows a correlation at lag 2 of 0.20 which is even significant under the null

hypothesis, there seems to be an indication that the $(1 - B^s)$ filter is not suitable. This may also be confirmed by the estimated value of θ_s which is -0.85 (cf. Nelson, 1973, p. 187). Finally, the price of steers series may also be modeled with a model like (4), a result which might have been expected by the estimate of -0.96 for θ_s in Nerlove et al. (1979, p. 211).

5. Concluding remarks

The model selection testing procedure for seasonal time series proposed in the present paper seems to meet its purpose. Advantages of the procedure are, first, that it selects a model by testing for restrictions in a general model, and, second, that the test statistics asymptotically follow standard normal distributions. The limitations to the procedure can be found in size distortions because of nonwhite error processes, and in power reduction in case the null and alternative hypotheses are in some sense close to each other. It may therefore be sensible to use the method together with related model selection strategies. Since the proposed procedure necessitates only simple calculations, it might however provide an initial quick check before any other methods are applied.

Table 2
Model selection for some empirical seasonal time series via testing for noninvertibility of the airline model

Series ^a	n	s	Autocorrelations		Test statistics		
			r_1	r_s	T_1	T_s	$T_{1,s}$
GN110 ^b	59	4	-0.34	-0.46	1.74	0.35	1.25
GN108 ^b	95	12	0.17	-0.14	9.24 *	4.05 *	8.20 *
GN106 ^b	109	12	0.06	-0.25	8.27 *	3.01 *	6.90 *
BJAIR ^c	83	12	-0.34	-0.44	2.06 *	0.46	1.64
FRCAR ^d	71	12	-0.15	-0.28	4.17 *	2.14 *	3.92 *
NECAR ^e	253	12	-0.09	-0.41	9.22 *	1.65	6.49 *
NEPST ^f	647	12	0.29	-0.46	28.42 *	1.18	17.24 *

* Significant at a 5% level.

^a The series refer to the doubly differenced series.

^b The sample autocorrelations of these series are given in Granger and Newbold (1986) on pages 110, 108 and 106, respectively.

^c Monthly airline data given in Box and Jenkins (1970). The last three years were excluded.

^d Monthly new car sales in the Netherlands, given in Franses (1991).

^e Monthly new car sales in the U.S., given in Nelson (1973).

^f Monthly price of steers for U.S., given in Nerlove et al. (1979).

Acknowledgements

Comments from Teun Kloek, Denise Osborn, and from participants of the EC² (1990) meeting in Amsterdam are gratefully acknowledged.

References

- Anderson, T.W. and A.M. Walker (1964), On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process, *Ann. Math. Statist.* **35**, 1296–1303.
- Bell, W. (1987), A note on overdifferencing and the equivalence of seasonal time series models with monthly means and models with $(0, 1, 1)_{12}$ seasonal parts when $\theta = 1$, *J. Bus. Econom. Statist.* **5**, 383–387.
- Box, G.E.P. and G.M. Jenkins (1970), *Time Series Analysis, Forecasting and Control* (Holden-Day, San Francisco, CA).
- Dickey, D.A., D.P. Hasza and W.A. Fuller (1984), Testing for unit roots in seasonal time series, *J. Amer. Statist. Assoc.* **79**, 355–367.
- Franses, P.H. (1991), Seasonality, non-stationarity and the forecasting of monthly time series, *Internat. J. Forecast.* **7**, 199–208.
- Franses, P.H. and T. Kloek (1990), Testing for non-invertibility of moving average models, unpublished manuscript.
- Ghysels, E. (1990), Unit-root tests and the statistical pitfalls of seasonal adjustment: The case of U.S. postwar real gross national product, *J. Bus. Econom. Statist.* **8**, 145–152.
- Granger, C.W.J. and P. Newbold (1986), *Forecasting Economic Time Series* (Academic Press, San Diego, CA, 2nd ed.).
- Harvey, A.C. (1984), A unified view of statistical forecasting procedures, *J. Forecast.* **3**, 245–275.
- Hasza, D.P. and W.A. Fuller (1982), Testing for nonstationary parameter specifications in seasonal time series models, *Ann. Statist.* **10**, 1209–1216.
- Hylleberg, S., R.F. Engle, C.W.J. Granger and B.S. Yoo (1990), Seasonal integration and cointegration, *J. Econometrics* **44**, 215–238.
- Nelson, C.R. (1973), *Applied Time Series Analysis for Managerial Forecasting* (Holden-Day, San Francisco, CA).
- Nerlove, M., D.M. Grether and J.L. Carvalho (1979), *Analysis of Economic Time Series* (Academic Press, New York).
- Osborn, D.R. (1990), A survey of seasonality in UK macroeconomic variables, *Internat. J. Forecast.* **6**, 327–336.
- Plosser, C.I. and G.W. Schwert (1977), Estimation of a non-invertible moving average process. The case of overdifferencing, *J. Econometrics* **6**, 199–224.