Butterworth Filters: A New Perspective

Víctor Gómez[*]

D-98008

Abril 1998

(*) Ministerio de Economía y Hacienda

Abstract

In this document, we propose 1) a method for automatic ARIMA model identification, based on estimating first the unit roots and then applying the BIC criterium to specify an ARMA model for the stationary (differenced) series, and 2) an automatic procedure to handle four types of outliers. This last procedure automatically identifies the location and type of outliers and corrects the series for their effects. The outliers are obtained one by one, each time estimating jointly all model parameters and outlier effects so far detected. Possible masking effects are eliminated by performing multiple regressions. Only linear regression techniques can be used for both automatic model identification and automatic outlier treatment, thereby reducing considerably the computation time needed to perform all the calculations. In this way, the practical value of the proposed procedures is greatly increased. The two procedures can be combined so that they can easily and objectively handle the most difficult cases of time series modeling in the presence of outliers. Extensions of the basic procedures to the case when there are missing observations, and tests for the log–level specification and for the presence of Trading Day and Easter effects, are also proposed.

Contents

1 Introduction

Many of the time series found in industry, economics, social sciences and many other scientific fields can be represented, at least as a first approximation, by means of a stationary linear process, possibly after having applied to the data some suitable transformation and differencing. In this document, the term stationary means wide sense stationary, not strict stationary. See Brockwell and Davis (1992), pp. 11-14. Other synonyms for wide sense stationary are second order stationary and weakly stationary. It follows from the theorem of Stone–Weierstrass of functional analysis that every stationary process without deterministic components can be approximated by an ARMA process with any degree of accuracy. ARMA processes are processes which in the time domain can be represented by means of a linear stochastic difference equation, and in the frequency domain have rational spectral densities. See Brockwell and Davis (1992), pp. 130–133 and 187–190.

ARIMA models are nonstationary models such that after some degree of differencing they become stationary ARMA models. The I in the acronym stands for "integrated", because integration is the inverse operation of differencing. It turns out that ARIMA models can be successfully used in practice to represent many time series, specially in economics.

ARIMA models became very popular after the publication of the seminal book by Box and Jenkins (1976), whose first edition was published in 1970. In this book, a systematic procedure was for the first time proposed to handle the problem of time series modeling by means of iterative cycles consisting of

i) Model identification

ii) Model estimation

iii) Diagnostic checking

At the same time, in Box and Jenkins' book of 1976 some computer programs were given to implement this methodology with the aid of digital computers. In this way, given the big computing power of computers, the user was for the first time able to apply a new and powerful machinery which would prove very useful in the field of time series analysis. The arrival of personal computers contributed further to increase enormously the demand for these new tools.

However, the modeling procedure proposed by Box and Jenkins (1976) was by no means a process that could be fully automated with the help of computers. In particular, step i) of the procedure, which involves model identification, was rather an art which required an expert in time series analysis to carry it out. It was certainly the most difficult of the three steps.

We will describe later the most common tools used nowadays for ARIMA model identification but, in spite of the advances that have taken place in the last two decades, we can say that step i) continues to be the most difficult of the three steps. Besides, we have to be aware of the fact that the majority of time series encountered in practice usually have outliers, which makes even more difficult the modeling procedure.

We will also review later the most important methods for automatic detection and correction of outliers that are currently in use. As far as ARIMA model identification is concerned, the presence of outliers can make it very difficult due to the important biases induced by the outliers in the parameter estimates and in the sample autocorrelation and partial autocorrelation functions. For this reason, any good strategy for ARIMA model identification has to account for the presence of outliers.

There are many reasons to try to automate as much as possible the ARIMA model identification stage, but they can be basically reduced to two. The first one is that one should eliminate as much as possible all mundane and mechanical chores, which can be performed by the computer, thus increasing the analyst's productivity. If the user is an accomplished analyst, he may invest more of his precious time on troublesome data sets that he has to model. On the contrary, if he is not an expert in time series models, he can use a powerful methodology that he couldn't even dream of using before. The second reason has to do with the objectivity of the identification stage, since it is desirable that this stage be not subject to heuristic methods and ad-hoc procedures that vary with each time series expert. For example, if a National Statitical Office has to produce some statistical data which require the modeling of some time series datasets and an expert is involved in the production process who uses subjective techniques, it may be criticized for publishing data which are neither objective nor reproducible.

In this document, a method is proposed for automatic model identification of ARIMA models in the presence of outliers which tries to automate as much as possible the identification stage for ARIMA models.

Before describing the proposed methodology, we will start by considering the procedure proposed by Tsay (1986) for identification of an ARIMA model in the presence of outliers. This procedure consists of the following steps

1) *Identification.* A tentative ARMA(p, q) model is identified using the method of Tsay and Tiao (1984), which is based on the extended sample autocorrelation function (ESAF). The method produces consistent estimates of the AR part, which may include unit roots. That is, the autoregressive polynomial includes the differencing polynomial, which, typically contains regular and/or seasonal differences.

2) *Estimation of the MA part.* The data are first filtered using the autoregressive polynomial and the AR estimates given by step 1. Then, Durbin's method with bias correction is applied to the filtered data to obtain the MA estimates. See Tsay (1986).

3) *Outlier detection.* Using the AR estimates of step 1 and the MA estimates of step 2 as if they were the true model parameters, a search for an outlier is performed, given the critical value C chosen by the user. If there is no outlier, go to step 4. If, on the contrary, there is one, correct the series for the outlier effect and go to step 1 with the corrected series instead of the original series. The statistics and the tests used for outlier detection are those proposed by Chang and Tiao (1983).

4) *Summary and model checking.* The tentatively identified model is that given by the last iteration of step 1. The number of outliers is determined by the number of iterations. It is not recommended to specify more than one outlier per iteration due to possible masking effects. The analyst should investigate the possible causes of the identified outliers to see whether the model should be modified in consequence.

This procedure presents some very interesting and positive aspects. To the best of this author's knowledge, it is the only procedure that deals with the identification of ARIMA models in the presence of outliers in a systematic way and proposes an effective computational method. Since it is based on linear regression techniques only, it is possible to reduce considerably the computational burden so that even a personal computer can be used to implement the procedure. This is very important, because it is well known that computing time is one of the great problems in outlier treatment in time series.

However, the specification of an ARIMA model in each step of the procedure seems to be excessive. Once the series has been corrected for the most important outliers, it is possible to specify a model for the series which will not probably change in subsequent steps. Also, and this could be a serious deficiency, the use of the ESAF to identify ARIMA models presents some problems. It seems that it is not a very adequate tool, especially for seasonal models, and, in any case, to obtain the unit roots in the autoregressive part.

Another aspect that can be improved is that of the detection and correction of outliers. The detection of spurious outliers due to masking effects can be avoided by using multiple regressions. Other important improvements can be obtained by using a robust estimator of the residual standard deviation, computing the residuals by means of the Kalman filter and using an "exact" filter to estimate the parameters in the multiple regressions.

Extensions of the basic procedure, like performing some preliminary tests for the log–level specification and Trading Day and Easter effects, can also be incorporated.

Taking into account the procedure proposed by Tsay (1986) and the previous considerations on how it could be improved, we propose in this document an algorithmical procedure which, briefly described, is the following.

a) *Preliminary tests.* If desired by the user, the procedure can test for the log–level specification, Trading Day and Easter effects. These last two tests are performed using the default model (airline model).

b) *Initialization.* If the user wants the series to be corrected for outliers, accept the model specified by the user (the default model is the airline model) and go to step 3. Otherwise, go to step 1. The critical value C for outlier detection can be either entered by the user or specified by the procedure. In this last case, the value of C is chosen depending on the length of the series.

c) *Step* 1. If the user has specified the differencing orders and whether there should be a mean in the model, go to step 2. Otherwise, the series is first corrected for all regression effects, if any. Then, using the corrected series, the differencing orders for the ARIMA model are automatically obtained and, also automatically, it is decided whether to specify a mean for the series or not. Go to step 2.

e) *Step* 2. Perform automatic identification of an ARMA(p, q) model for the differenced series, corrected for all outliers and other regression effects, if any. If the user wants to test for Trading Day and Easter effects and any of these effects was specified in the preliminary tests, check whether the specified effects are significant for the new model. If the user wants to correct the series for outliers, go to step 3. Otherwise, stop.

d) *Step* 3. Assuming the model known, perform automatic detection and correction of outliers using C as critical value. If a stop condition is not satisfied, perhaps decrease the critical value C and go to step 1.

In this document, the preliminary tests and steps 1, 2 and 3 of the proposed procedure will be described in detail. The procedure will be extended later to cover the case of missing observations.

The previous algorithm doesn't use the method proposed by Tsay and Tiao (1984), which is based on the (ESAF), to specify an ARIMA model. Instead, it estimates first the unit roots in the initialization step and then, in step 2, it uses the BIC criterium, see Akaike (1978, 1979), to specify an ARMA model for the differenced series. This is due to two reasons. First, as mentioned above, the ESAF method of Tsay and Tiao (1984) apparently does not get very satisfactory results when it is used to obtain the differencing orders. Second, as reported by Liu (1989), the method is very informative as regards the identification of ARIMA models for time series with no seasonality, but for seasonal time series it is less successful. The estimation of both unit roots and the parameters of the ARMA models for which the BIC is computed is performed by means of the method of Hannan and Rissanen (1982), henceforth referred to as HR, although the exact maximum likelihood method can also be applied.

The proposed procedure uses a method for outlier treatment which incorporates substantial improvements with respect to the method proposed by Tsay (1986). Among them, we can mention

1) It allows for four types of outliers.

2) It uses the MAD estimator to estimate robustly the residual standard deviation.

3) It performs multiple regressions in order to eliminate possible masking effects.

4) Exact residuals are computed by means of a fast Kalman filter routine which is based on the algorithm of Morf, Sidhu and Kailath (1974).

5) It uses a method to incorporate or reject outliers which is similar to the stepwise regression procedure for selecting the "best" regression equation.

6) To estimate the parameters in the multiple regressions, it uses the algorithm of 4), applied to the data and to the columns of the design matrix, together with the QR algorithm. This last algorithm is described in, for example, Gill, Murray and Wright (1992), pp. 37–40.

7) It uses the HR method to estimate the parameters of ARMA models. This method produces estimators which have similar properties than those obtained using maximum likelihood estimation, but is much less computationally expensive.

Finally, the proposed procedure can be used with regression models with ARIMA errors, as opposed to the procedure proposed by Tsay (1986).

As mentioned above, the differencing orders are obtained in the proposed procedure by estimating the unit roots. This is done in step 1 by first estimating autoregressive models of the form

$$(1 + \phi_1 B + \phi_2 B^2)(1 + \Phi B^s)(z(t) - \mu) = a(t), \tag{1}$$

where $\{z(t)\}$ is the observed series, s is the number of observations per year, μ in the mean of the process, B is the backshift operator, $Bz(t) = z(t-1)$, and $\{a(t)\}$ is a sequence of i.i.d. $N(0, \sigma^2)$ random variables. Then, the series is differenced using the differencing orders given by the unit roots obtained after estimating (1) and an ARMA$(1,1)\times(1,1)_s$ model with mean, that is, a model of the form

$$(1 + \phi B)(1 + \Phi B^s)(x(t) - \mu) = (1 + \theta B)(1 + \Theta B^s)a(t), \tag{2}$$

is fitted to the differenced series $\{x(t)\}$. If any new unit roots appear after estimating (2), the differencing orders are properly increased and a new model (2) is fitted. The process is continued until no more unit roots are found. Then, the residuals of the last estimated model are used to decide whether to specify a mean for the model or not. The choice of models (1) and (2) will be justified later.

The proposed procedure has two well delimited parts. The first one consists of the automatic model identification. In this part, the differencing polynomial and an ARMA(p, q) model for the differenced series are obtained. The second one implements the algorithm for automatic detection and correction of outliers. The four types of outliers considered will be described later. Both parts can be sequentially combined to handle the automatic model identification of time series in the presence of outliers. The document is structured as follows. In Section 2, the automatic model identification procedure is described. Section 3 deals with the algorithm for outlier detection and correction. Some computational aspects of the proposed procedure are described in Section 4. Also in this Section, a Fortran program written by the author to implement the proposed procedure is described. Finally, in Section 5, the proposed procedure is applied to some real and simulated series and the conclusions are summarized.

2 Automatic Model Identification

Traditionally, in order to identify an ARIMA model for a time series, a time series expert is needed who, with the help of some statistical software for time series modeling like, for example, SCA, examines the graphs of the original series, some differences of the series, the sample autocorrelation and partial autocorrelation functions, and the ESAF. These tools help the expert determine first the differencing degree and then an ARMA model for the differenced series, which should be chosen according to the principle of parsimony.

The process of determining the differencing orders can sometimes be very difficult, and the selection is always made with some subjectivity. In addition, for stationary models, the theoretical autocorrelation and partial autocorrelation functions present a clear pattern only for pure moving average or pure autoregressive models, whereas, as already mentioned, it seems that the

ESAF can be used with some guarantee of success only with non seasonal models. If to the previous considerations we further add that the functions used in practice are all sample functions, which may differ substantially from the theoretical ones, and that there may be outliers in the series which distort even more the patterns of the theoretical functions, we conclude that the identification stage is possibly the most difficult of the three stages of time series modeling.

As an alternative to the traditional methods, we can consider the penalty methods, which use criteria based on information theory, such as AIC or BIC. However, these methods can only be used with ARMA models, which are stationary. Therefore, the problem remains of obtaining the necessary differencing orders to render a time series stationary. The penalty methods, and in particular AIC, have been criticized for their tendency to overparametrize. Other methods, like the corner method or the minimal canonical correlation method (SCAN) don't seem to be very useful. A good reference to review all identification methods for ARMA models is the book by Choi (1992).

Recently, the SCA software package has incorporated a module for automatic ARIMA model identification, called "SCA–Expert". This module uses a procedure based on the filtering method proposed by Liu (1989) and certain heuristic rules. Briefly, this method consists of the following.

1. Examine first the sample autocorrelation functions, henceforth referred to as SAF, of $z(t)$, $(1-B)z(t)$, $(1-B^s)z(t)$ and $(1-B)(1-B^s)z(t)$ to assert the differencing orders and to see whether seasonality is present. After that, examine the SAF of the properly differenced series. If an obvious seasonal ARIMA model can be specified from the SAF, stop. Otherwise, go to the following step. Denote by $y(t)$ the differenced series.

2. If an obvious tentative model cannot be deduced from the SAF of $y(t)$, estimate an intermediate model of the type ARMA$(1,1) \times (1,1)_s$. If no one of the autoregressive parameters has is close to 1, generate the series $R(t)$ and $S(t)$, which are the result of filtering $y(t)$ with the ARMA$(1,1)_s$ and ARMA$(1,1)$ models that make up the intermediate model.

   If any of the autoregressive parameters is close to 1, then difference properly. After differencing, a new intermediate model of the same type is estimated and new $R(t)$ and $S(t)$ series are generated.

3. Use the sample autocorrelation and partial autocorrelation functions, as well as the ESAF, of $R(t)$ to identify an ARMA model adequate for the $R(t)$ series.

4. In order to identify a model for $S(t)$, the SAF of $S(t)$ can be used. If a model is not clear for $S(t)$, examine also the estimated parameters for the seasonal part in the intermediate model and use them to specify a model for $S(t)$.

This procedure can be criticized for using heuristic rules that are not documented.

ARIMA modeling has proved useful in many scientific fields and the reason because it is not used more often in practice lies in the barrier imposed by the need of a time series expert to model the time series data and the time consumed by the modeling process. As mentioned in Section 1, it is very important in practice to have an automatic time series modeling procedure.

The procedure proposed in this document for automatic model identification consists of two parts.

1. Identification of the differencing degrees for the series and specification of the mean if necessary.

2. Identification of an ARMA(p, q) for the differenced series, possibly corrected for outliers and other regression effects, by means of the BIC criterium, using a modification of HR's method.

Let the observed series $\{z(t)\}$ follow the ARIMA(p, d, q) model

$$\phi(B)\delta(B)(z(t) - \mu) = \theta(B)a(t),$$

where $\phi(B) = 1 + \phi_1 B + \ldots + \phi_p B^p$, $\delta(B) = 1 + \delta_1 B + \ldots + \delta_d B^d$ and $\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$ are polynomials in the backshift operator B of degrees p, d and q, $\{a(t)\}$ is a i.i.d. $N(0, \sigma^2)$ sequence of random variables and μ is the mean of the process. The roots of $\delta(B)$ are assumed to lie on and those of $\phi(B)$ outside the unit circle, so that the process $w(t) = \delta(B)z(t)$

follows a stationary ARMA(p,q) process. Most economic series follow so-called multiplicative seasonal models, where

$$\begin{aligned} \delta(B) &= \nabla^a \nabla_s^b, \\ \phi(B) &= \phi_r(B)\phi_S(B^s), \\ \theta(B) &= \theta_r(B)\theta_S(B^s), \end{aligned} \tag{3}$$

s is the number of observations per year, $\nabla^a = (1-B)^a$ and $\nabla_s^b = (1-B^s)^b$. In practice, for economic time series, the inequalities $0 \le a \le 2$ and $0 \le b \le 1$ hold. For simplicity, we will use in the rest of the Section the notation (2), even for multiplicative seasonal models. We will make specific reference to these models when necessary.

2.1 Procedure to Obtain the Differencing Polynomial δ(B) and to Specify a Mean for the Model, if Necessary

Suppose that the series $\{z(t)\}$ follows model (2), where it is assumed $\mu = 0$ to simplify matters. Then, by theorems 3.2 and 4.1 of Tiao and Tsay (1983), the ordinary least squares, henceforth OLS, estimators obtained from an AR(k) regression, where $k \ge d$, asymptotically verify

$$\hat{\Phi}_k(B) \doteq \hat{\delta}(B)\hat{\phi}_m(B),$$

where $\doteq$ denotes asymptotic equivalence in probability, $m = k-d$ and $\hat{\Phi}_k(B)$, $\hat{\delta}(B)$ and $\hat{\phi}_m(B)$ are, respectively, the polynomials estimated by OLS in the autoregressions

$$\begin{aligned} \Phi_k(B)z(t) &= a(t), \\ \delta(B)z(t) &= a(t), \\ \phi_m(B)w(t) &= a(t), \end{aligned}$$

where $w(t) = \delta(B)z(t)$ is a stationary process that follows the ARMA(p,q) model $\phi(B)w(t) = \theta(B)a(t)$ and the subindex in $\Phi_k(B)$ and $\phi_m(B)$ denotes the polynomial degree. In addition, the equality $\hat{\delta}(B) = \delta(B) + O_p(N^{-1})$ holds, where N is the series length.

The practical implication of this result is that if we perform an autoregression of order greater than or equal to the (unknown) degree of the polynomial $\delta(B)$, we obtain a consistent estimate of $\delta(B)$ as a component of $\hat{\Phi}_k(B)$. If we specify a model of the form AR(2)$\times(1)_s$ for $\Phi_k(B)$, we cover the cases $\delta(B) = 1$, $\delta(B) = \nabla$, $\delta(B) = \nabla_s$, $\delta(B) = \nabla\nabla_s$ and $\delta(B) = \nabla^2\nabla_s$, which are the ones of most applied interest.

In the case of non seasonal models, where $\delta(B) = \nabla^a$ and $0 \leq a \leq 2$ is assumed, if we specify an AR(2) model, all important cases are covered.

Based on the previous considerations, the proposed algorithm to identify the differencing polynomial is the following.

I) Specify a model of the form AR(2)$\times(1)_s$ with mean, given by equation (1) if the process is multiplicative seasonal, or an AR(2) model with mean , also given by (1) but without the second factor, if the process is regular. This autoregressive process is estimated using HR's method, which will be described later, unless the user decides to use exact maximum likelihood. If the roots estimated with HR's method lie outside the unit circle, the autoregression is estimated again using unconditional least squares. A root is considered to be a unit root if its modulus is greater than a specified value, which by default is .97. Go to II).

II) In addition to the differencing degrees identified in I) as a result of the estimated unit roots, a model of the form ARMA$(1,1)\times$ $(1,1)_s$ with mean for seasonal series, or a model ARMA$(1,1)$ with mean for non seasonal series, is specified. Letting $x(t)$ be the series that results from differencing $z(t)$ with the differencing polynomial obtained after the estimation of the initial autoregression, the equations for these models are given by (2) in the seasonal case, and by (2) without the factors involving B^s in the regular case. The model is estimated using HR's method or exact maximum likelihood, depending on the option chosen by the user, and if any of the estimated autoregressive parameters is close to 1, the degree of differencing is increased accordingly. A parameter is considered to be close to 1 if its modulus is greater than a specified value, which by default is .88. To avoid cancelation of terms in the model, the absolute value of the difference between each autoregressive parameter and its corresponding moving average parameter should be greater than .15. For multiplicative seasonal models, it is not possible to pass from 0 differencing to $\nabla\nabla_s$ directly. If this happens, the roots of the autoregressive polynomial obtained in I) are considered again, the one with greatest modulus is chosen, and the series is differenced accordingly. If the series has been differenced in this step, repeat II). Otherwise, go to III).

III) Using the residuals of the last estimated model, it is decided whether to specify a mean for the model of the series or not depending on the significance of the estimated residual mean. Stop.

The ARMA$(1,1)\times(1,1)_s$ model used in II) is very flexible and constitutes a generalization of the airline model of Box and Jenkins (1976). For stationary series, it approximates well many of the ARMA models encountered in practice. When it is used with nonstationary series, it can detect autoregressive unit roots which have not been detected by the autoregressive model used I). Imagine, for example, a model of the form $(1-B)(z(t)-\mu)=(1-.8B)a(t)$, where the autoregressive and the moving average part almost cancel out. In this case, an ARMA$(1,1)$ model would probably estimate the unit root better than an AR(2) model.

The proposed procedure also allows for the detection of the presence of a pair of complex unit roots. This is done by means of first specifying an AR(2) model to each of the series that result from applying $\delta(B)=\nabla^a\nabla_s^b$ to the original series, where $(a,b)=(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$, after having subtracted the estimated mean, and then estimating these autoregressive models. If there are complex unit roots in any of these models, the pair with a modulus closer to 1 is chosen and from that moment onwards these roots are considered fixed and the previous algorithm continues.

If there is a pair of complex unit roots, the corresponding degree two polynomial would be treated as part of the polynomial $\delta(B)$, as far as differencing the series is concerned.

The identification of the differencing polynomial $\delta(B)$ constitutes the first part of the proposed procedure for automatic model identification. Traditionally, $\delta(B)$ is specified by examining the graphs of the SAF for the different regular and seasonal differences considered. However, specifying $\delta(B)$ can be difficult on some occasions, especially if the series follows a multiplicative seasonal model. In fact, it is not unusual to find series in published articles which have been overdifferenced. See, for example, the monthly sales series of Chatfield and Prothero (1973) in appendix A.

Consider now the case of a regression model with ARIMA errors. The question naturally arises as to whether the previous analysis is still valid and if, in consequence, the procedure just described is also applicable in this case. In this respect, the results of Tsay (1984), pp. 119–120, allows one to state that, under very general conditions, it is possible to work with the original series in order to identify the differencing polynomial.

2.2 Procedure to Obtain the Arma(p, q) Model for the Differenced Series, Possibly After Having Corrected for Outliers and Other Regression Effects

As mentioned in Section 1, the step 2 of the proposed procedure consists of the identification of an ARMA(p,q) model for the differenced series, possibly corrected for outliers and other regression effects. We will start by considering that there are neither outliers nor other regression effects and we will extend the results later to the general case. The method used in step 2 of the proposed procedure is based on HR's method, which in turn uses the BIC criterium and computes the ARMA model estimators by means of linear regressions. Therefore, these estimators are computationally cheap, although it can be shown that they have similar properties to those obtained by maximum likelihood.

Let $z = (z(1), \ldots, z(N))'$ the observed series, which follows model (2), where we assume $\mu = 0$ for simplicity. Since $\delta(B)$ has been identified in step 1 of the proposed procedure, we can compute the differenced series $w(t) = \delta(B)z(t)$, $t = d+1, \ldots, N$, which follows the ARMA(p,q) model

$$\phi(B)w(t) = \theta(B)a(t), \tag{4}$$

where $\phi(B)$, $\theta(B)$ and $\{a(t)\}$ are like in (2). If the model is multiplicative seasonal, the decomposition (3) holds. In order to avoid notational problems, let the differenced series be $w = (w(1), \ldots, w(N-d))'$. If the orders of the fitted model (4) are (p,q), the BIC statistic is

$$\text{BIC}_{p,q} = \log(\hat{\sigma}^2_{p,q}) + (p+q)\log(N-d)/(N-d), \tag{5}$$

where $\hat{\sigma}^2_{p,q}$ is the maximum likelihood estimator of σ^2. The criterium estimates the orders (p,q) by choosing $(\hat{p}, \hat{q})$ which minimizes (5).

The method just described to choose the orders, which is based on the traditional BIC criterium, is computationally expensive because one has to perform a nonlinear optimization for each (p, q) to compute $\hat{\sigma}^2_{p,q}$. For this reason, Hannan and Rissanen (1982) propose to perform the estimation using linear regression techniques in three steps, although the third step is used to compute the estimators of the ARMA model chosen by the BIC criterium and, therefore, only the first two steps are used to select the orders (p, q).

2.2.1 Computation of $\mathbf{BIC_{p,q}}$

The procedure proposed to compute the BIC criterium performs the parameter estimation of each ARMA (p, q) model for which the $\mathrm{BIC}_{p,q}$ is computed in two steps and is slightly different from the method originally proposed by HR. In the first step, which takes place only if there is a moving average part $(q > 0)$, estimates $\hat{a}(t)$ of the innovations $a(t)$ in (4) are obtained by fitting a long autoregressive model to the series. That is, given a big n, the $\hat{a}(t)$ are computed using

$$\hat{a}(t) = \sum_{j=0}^{n} \hat{\phi}_n(j) w(t-j), \quad \hat{\phi}_n(0) = 1, \quad t \geq 1,$$

where $w(t) = 0$ if $t \leq 0$ and the $\hat{\phi}_n(j)$ are computed using Durbin-Levinson's algorithm. This last algorithm consists of first estimating the sample autocorrelations

$$c(t) = \frac{1}{N-d} \sum_{s=1}^{N-d-t} w(s) w(s+t)$$

and then recursively computing the $\hat{\phi}_n(j)$ using the equations

$$\hat{\phi}_n(n) = -\sum_{j=0}^{n-1} \hat{\phi}_{n-1}(j) c(n-j) / \hat{\sigma}^2_{n-1}, \qquad \hat{\phi}_n(j) = \hat{\phi}_{n-1}(j) + \hat{\phi}_n(n) \hat{\phi}_{n-1}(n-j),$$

$$\hat{\sigma}^2_n = \{1 - \hat{\phi}^2_n(n)\} \hat{\sigma}^2_{n-1}, \quad \hat{\phi}_1(1) = c(1)/c(0), \quad \hat{\sigma}^2_0 = c(0).$$

In the proposed procedure, the value of n is chosen to be $n = max\{[\log^2(N-d)], 2max\{p, q\}\}$, where (p, q) are the orders of the ARMA model for which the BIC is being computed and $[\log^2(N-d)]$ is the integer part of $\log^2(N-d)$. This choice is based on the fact that Hannan and Rissanen (1982), p. 88, assume that n is greater than $\log(N-d)$, but not greater than $\log^b(N-d)$, for some $b < \infty$.

In the second step, given the orders (p, q), the parameters of model (4) are obtained by minimizing

$$\sum_{t=m}^{N-d} \Big\{ \sum_{j=0}^{p} \phi_j w(t-j) - \sum_{j=1}^{q} \theta_j \hat{a}(t-j) \Big\}^2,) \tag{6}$$

where $m = max\{p+1, q+1\}$ and $\phi_0 = 1$.

If there is no moving average part ($q = 0$), the estimation of the parameters of the ARMA model finishes here. Note that, in this case, the estimates obtained for the autoregressive part coincide with the ones obtained by OLS.

If there is a moving average part ($q > 0$), the estimators $\tilde{\phi}_j$ and $\tilde{\theta}_j$, obtained by minimizing (6), are consistent but have a bias and, therefore, they are not asymptotically efficient. In order to obtain bias–corrected, consistent and asymptotically efficient estimators, see Zhao–Guo (1985), first form

$$\tilde{a}(t) = -\sum_{j=1}^{q} \tilde{\theta}_j \tilde{a}(t-j) + \sum_{j=0}^{p} \tilde{\phi}_j w(t-j), \quad t \geq 1,$$

where $\tilde{a}(t) = 0$ and $w(t) = 0$ if $t \leq 0$. Then put

$$\eta(t) = -\sum_{j=1}^{p} \tilde{\phi}_j \eta(t-j) + \tilde{a}(t), \quad \xi(t) = -\sum_{j=1}^{q} \tilde{\theta}_j \xi(t-j) + \tilde{a}(t), \quad t \geq 1,$$

where $\eta(t) = 0$ and $\xi(t) = 0$ if $t \leq 0$. Finally, regress $\tilde{a}(t)$ on $-\eta(t-j)$, $j = 1, \ldots, p$, and $\xi(t-j), j = 1, \ldots, q$. The estimated regression coefficients are added to the estimators $\tilde{\phi}_j$ and $\tilde{\theta}_j$ to obtain the desired estimators $\hat{\phi}_j$ and $\hat{\theta}_j$.

When there are a moving average part ($q > 0$) and an autoregressive ($p > 0$), better estimates of the moving average part can be obtained by repeating the previous procedure with the series filtered with the autoregressive filter. That is, the series is first filtered with the autorregresive filter $\hat{\phi}(B)$, which has been estimated in the two previous steps, to obtain the series $x(t) = \hat{\phi}(B)w(t)$. Then, the series $x(t)$, which asymptotically follows the model $x(t) = \theta(B)a(t)$ and, therefore, does not have an autoregressive part, is subject to the two previous steps.

Once the parameter estimates of model (4) have been obtained for some orders (p, q), the estimator $\hat{\sigma}^2_{p,q}$ is needed to compute the $\text{BIC}_{p,q}$ statistic. In the proposed procedure, the residuals $r(t)$, $t = 1, \ldots, n = max\{p, q\}$ of the series $w(t)$ are first computed using a fast Kalman filter routine based on the algorithm of Morf, Sidhu and Kailath (1974). Then, the rest of the residuals $r(t), t = n+1, \ldots, N-d$ are recursively obtained using the difference equation (4). Finally, the estimator $\hat{\sigma}^2_{p,q}$ is computed by the formula

$$\hat{\sigma}^2_{p,q} = \frac{1}{N-d} \sum_{t=1}^{N-d} r^2(t),$$

and the $\text{BIC}_{p,q}$ statistic is computed using (5).

2.2.2 Optimization of $\mathbf{BIC_{p,q}}$

After having described the algorithm to compute $\text{BIC}_{p,q}$ for each (p, q), we now describe the algorithm used by the proposed procedure to obtain the optimal model of the form (4). To this end, suppose the general case, where the series follows a multiplicative seasonal model given by (3). In practice, it is assumed that the orders of the $\text{ARMA}(p_r, q_r)\times(p_s, q_s)_s$ model that the series follows verify $0 \leq p_r, q_r \leq 3$ and $0 \leq p_s, q_s \leq 2$, and the BIC statistic should be computed for all these combinations. Since the resulting number of combinations is high, in the proposed procedure the search is performed sequentially. The algorithm is

I) Specify first an $\text{ARMA}(3, 0)$ model for the regular part. Then, compute the BIC statistic for models where the seasonal part verifies $0 \leq p_s, q_s \leq m_s$, and choose the minimum. The number m_s is chosen by the user, the default value being 1.

II) Fix the seasonal part to that chosen in I), compute the BIC statistic for models where the regular part verifies $0 \leq p_r, q_r \leq m_r$, and choose the minimum. The number m_r is chosen by the user, the default value being 3.

III) Fix the regular part to that chosen in II), compute the BIC statistic for models where the seasonal part verifies $0 \leq p_s, q_s \leq m_s$, and choose the minimum. The number m_s is that of I).

The previous algorithm allows for a substantial reduction in computing time and, however, the results obtained with it are very satisfactory. Once the previous algorithm has finished, and in order to avoid the tendency of BIC to overparametrize, especially in the seasonal part, the smallest five BIC are first ordered in ascending order. Then, the first one is compared to the other four and if the difference in absolute value is less than a certain number and the biggest of the two BIC corresponds to a more parsimonious seasonal part, this last one is chosen. Among all the BIC that satisfy this condition, the one that corresponds to the more parsimonious part is chosen, provided that the seasonal part exists ($p_s > 0$ or $q_s > 0$). The procedure also favours balanced models (models where the degrees of the autoregressive and the moving average parts coincide).

In the previous algorithm, if the parameters estimated for an ARMA model using HR's method are such that the roots of the autoregressive or the moving average polynomials lie within the unit circle, this fact is considered as an indication of model inadequacy and the model is rejected.

The tentative model ARMA$(3, 0)$ specified in I) of the previous algorithm seems to be robust and the sequential search of the algorithm has given very satisfactory results in all performed tests of the proposed procedure, with real and simulated series.

If there is a mean or other regression effects in model (4), the proposed procedure obtaines first OLS estimators of the regression parameters. Then, these effects are subtracted from the differenced series before computing the parameter estimates of model (4) and also before computing the residuals $r(t)$ needed in the computation of $\hat{\sigma}^2_{p,q}$ and the BIC statistic.

The orders (p, q) obtained using the BIC criterium are consistent, see Hannan and Rissanen (1982), whereas those obtained using Akaike's AIC are not. However, the proof of the consistency in the BIC case is based on the existence of a true ARMA model for the differenced series, which has been doubted by some authors, included Hannan himself. For this reason, some authors justify the use of the AIC criterium, although it is not consistent, on the grounds that there is no such a thing as a true model.

3 Automatic Detection and Correction of Outliers

When analysing time series data, it is not unusual to find outlying observations due to uncontrolled or unexpected interventions, like strikes, major changes in political or economic policy, the occurrence of a disaster, gross errors, and so forth. Since ARIMA models, which are frequently used in time series modeling, are designed to grasp the information of processes with a homogeneous memory pattern, the presence of outlying observations or structural changes may influence the efficiency and goodness of fit of these models. See, for example, Abraham and Box (1979), Chen and Tiao (1990), Tsay (1986), and Guttman and Tiao (1978).

The traditional approach to handle the problem of outliers, once it is assumed that a proper ARIMA model has been correctly identified for the series, consists of first identifying the location and the type of outliers and then use the intervention analysis proposed by Box and Tiao (1975). This procedure requires that a time series expert first examines the data and then, with the help of some time series software, analyses the sample autocorrelation and partial autocorrelation functions of the residuals, graphs of the series and the residuals, etc. For this reason, it is important to try to find some procedure which automates as much as possible the process of detection and correction of outliers. Among the first steps in this direction, we can mention the procedures of Chang, Tiao and Chen (1988), Hillmer, Bell and Tiao (1983), and Tsay (1988). These procedures are quite effective in detecting the locations and estimating the effects of large isolated outliers. However, the problem is not solved because

1) The presence of outliers may result in an incorrectly specified model

2) Even if the model is appropiately specified, outliers may still produce important biases in parameter estimates

3) Some outliers may not be identified due to a masking effect

The method proposed by Tsay (1986), which has been described in Section 1 of this document, is an important contribution to solve the problem of model identification in the presence of outliers. Chen and Liu (1993) have proposed a method for outlier treatment that tries to solve problems 2) and 3). This method works rather satisfactorily, although it presents some deficiencies. For example, the method uses exact maximum likelihood estimation several times, thus making it computationally expensive. It does not use exact residuals. The algorithm is too complicated. Multiple regressions are performed without filtering the data and the columns of the design matrix by an "exact" filter, like the Kalman filter. It uses instead a conditional filter. Finally, the method uses a sort of backward elimination procedure to select the "best regression equation", instead of a stepwise procedure which is more robust.

The method proposed in this document for the detection and correction of outliers attempts to solve problems 2) and 3) in such a way that the deficiencies in the procedure of Chen and Liu (1993) are eliminated, as will be described later. In addition, if it is used in the algorithm proposed in Section 1, together with the algorithm for automatic model identification of Section 2, the proposed algorithm constitutes an alternative procedure to that of Tsay (1986) to solve the problem of automatic model identification in the presence of outliers that complements and may improve considerably Tsay's procedure.

Like in Section 2, suppose first that there are no regression effects. We will extend the results to the general case later. Let the series $\{z(t)\}$ follow the ARIMA(p,d,q) model given by (2), where it is understood that if the model is multiplicative seasonal, (3) holds. We will assume for simplicity that $\mu = 0$ and we will use the notation (2), referring to (3) when necessary. To model the effect of an outlier, consider the model

$$z^*(t) = z(t) + \omega\nu(B)I_t(T), \tag{7}$$

where $\nu(B)$ is a quotient of polynomials in B, $\{z(t)\}$ is the outlier free series, $I_t(T) = 1$ if $t = T$ and $I_t(T) = 0$ otherwise, is an indicator function to refer to the time in which the outlier takes place and ω represents the magnitude of the outlier. Proposed procedure considers four types of outliers

IO: $\nu(B) = \theta(B)/(\delta(B)\phi(B))$,

AO: $\nu(B) = 1$,

TC: $\nu(B) = 1/(1-\delta B)$,

LS: $\nu(B) = 1/(1-B)$.

The acronyms stand for innovational outlier (IO), additive outlier (AO), temporary change (TC) and level shift (LS). The value of δ is considered fixed and is made equal to 0.7. For a more detailed discussion on the nature and motivation for these outliers, see Chen and Tiao (1990), Fox (1972), Hillmer, Bell and Tiao (1983), and Tsay (1988). These four outliers correspond to some simple types of outliers. More complicated outliers can be usually approximated by combinations of these four types.

3.1 Estimation and Adjustment for the Effect of an Outlier

Suppose that the parameters in model (2) are known, the observed series is $z^* = (z^*(1), \ldots, z^*(N))'$, the outlier free series is $z = (z(1), \ldots, z(N))'$ and put $Y = (\nu(B)I_1(T), \ldots, \nu(B)I_N(T))'$. Then, (7) can be written as the regression model with ARIMA errors

$$z^* = Yw + z. \tag{8}$$

To simplify the exposition, we will assume that z in (8) follows an ARMA model or, what amounts to the same thing, $\delta(B) = 1$ in (2). If this is not the case, we would work with the series obtained by differencing z^*, z and Y in (8). Let $Var(z) = \sigma^2\Omega$ and $\Omega = LL'$, with L lower triangular, the Cholesky decomposition of Ω. Premultiplying (8) by L^{-1}, the following ordinary least squares model is obtained

$$L^{-1}z^* = L^{-1}Yw + L^{-1}z. \tag{9}$$

Letting $r = L^{-1}z$, the equality $Var(r) = \sigma^2 I_N$ holds and vector r is the residual vector of the series (not observed). If we let the estimated residuals be $r^* = L^{-1}z^*$ and write $X = L^{-1}Y$, (9) can be written as

$$r^* = Xw + r. \tag{10}$$

If Y is 0 in (8), the model would be $z^* = z$ and if we applied the Kalman filter to this model, we would obtain $L^{-1}z^*$. This result, which a standard result of control theory, allows us to see the Kalman filter as an algorithm that, applied to any vector v instead of z^*, yields $L^{-1}v$. Therefore, if we apply the Kalman filter to the vector of observations z^* and to the vector Y, we can move from (8) to (9) or, what amounts to the same thing, from (8) to (10).

We can estimate ω by OLS in (10 to obtain

$$\tilde{\omega} = (X'X)^{-1}X'r^*, \tag{11}$$

where the estimator variance is $Var(\tilde{\omega}) = (X'X)^{-1}\sigma^2$. To test the null hypothesis that there is no outlier at $t = T$, we can use the statistic

$$\tau = (X'X)^{1/2}\tilde{\omega}/\sigma, \tag{12}$$

which is distributed $N(0,1)$ under the null.

In practice, the parameters of model (2) will not be known and they will have to be estimated. Under these circumstances, the usual procedure consists of estimating first the parameters of model (2) by exact maximum likelihood, as if there were no outliers, and then using instead of (11) and (12) their sample counterparts

$$\hat{\omega} = (\hat{X}'\hat{X})^{-1}X'\hat{r}^*, \qquad \hat{\tau} = (\hat{X}'\hat{X})^{1/2}\hat{\omega}/\hat{\sigma},$$

which are obtained by replacing in (11) and (12) the unknown parameters with their estimates. It can be shown that $\hat{\tau}$ is asymptotically equivalent to τ. See Chang, Tiao and Chen (1988), p. 196. Each matrix X and, therefore, $X'X$ depends on the type of the outlier.

To see whether there is an outlier at $t = T$, the four estimators $\hat{\omega}_{IO}(T)$, $\hat{\omega}_{AO}(T)$, $\hat{\omega}_{TC}(T)$ and $\hat{\omega}_{LS}(T)$ are first computed, along with the statistics $\hat{\tau}_{IO}(T)$, $\hat{\tau}_{AO}(T)$, $\hat{\tau}_{TC}(T)$ and $\hat{\tau}_{LS}(T)$, where the subindex refers to the outlier type. Then, as proposed by Chang, Tiao and Chen (1988), the statistic $\lambda_T = max\{|\hat{\tau}_{IO}(T)|, |\hat{\tau}_{AO}(T)|, |\hat{\tau}_{TC}(T)|, |\hat{\tau}_{LS}(T)|\}$ is used. If $\lambda_T > C$, where C is a predetermined critical value, then there is the possibility of an outlier of the type given by the subindex of the statistic $\hat{\tau}$ for which the maximum is obtained.

Since the time $t = T$ at which the outlier occurs is unknown in practice, the criterium based on the likelihood quotient of Chang y Tiao (1983), leads to repeat the previous operation for $t = 1, \ldots, N$ and compute $\lambda = max_t \lambda_t = |\hat{\tau}_{tp}(T)|$, where tp can be IO, AO, TC or LS. If $\lambda > C$, then there is an outlier of type tp at T.

Once the type of an outlier at $t = T$ is known, the series and the residuals can be corrected for its effect using (7) and (10).

Up to now, we have assumed that r^* and X were computed by means of an "exact" filter, which was the Kalman filter. This is the correct thing to do, since the number of observations in a time series is always finite and we cannot apply the semi–infinite filter, given by the inverse of the series model $\pi(B) = 1 + \pi_1 B + \pi_2 B^2 + \cdots = \phi(B)\delta(B)/\theta(B)$, to (7) to obtain

$$\pi(B)z_t^* = \omega[\pi(B)\nu(B)I_t(T)] + a(t), \quad t = 1, \ldots, N,$$

instead of (9). In practice, the usual procedure consists of truncating the filter $\pi(B)$ and disregarding some observations at the beginning of the series. See Chen and Liu (1993), p. 285. In the procedure proposed in this document, the residuals are filtered with an exact filter to obtain r^* and the filter $\pi(B)$ is used to filter the vector Y in (8).

3.2 The Case of Multiple Outliers

When multiple outliers are present, we should use instead of (7) the model

$$z^*(t) = z(t) + \sum_{i=1}^{k} \omega_i \nu_i(B) I_t(t_i). \tag{13}$$

As shown in Chen and Liu (1993), the estimators of the ω_i obtained simultaneously using (13), can be very different from the ones obtained by an iterative process using the results of the previous Section. That is, by obtaining first $\hat{\omega}_1$, then $\hat{\omega}_2$, etc. For this reason, it is important that every algorithm for outlier detection performs at some point multiple regressions to detect spurious outliers and correct the bias produced in the estimators sequentially obtained.

In order to estimate the parameters in the multiple regressions, when the parameters of the ARIMA model (2) are assumed to be known, the algorithm proposed in this document uses first the Kalman filter like when we moved from (8) to (9). Then, the estimators of the ω_i and the corresponding statistics are computed using (11) and (12). This is done in an efficient manner, using the QR algorithm and Housholder transformations. A more detailed description will be given in Section 5.

3.3 Estimation of the Standard Deviation σ of the Residuals

When outliers are present in the series, the usual sample estimator can overestimate δ. For this reason, it is advisable to use a robust estimator. In the proposed procedure the estimator used is the MAD estimator, defined by

$$\hat{\sigma} = 1.483 \times median\{|r^*(t) - \tilde{r}^*|\},$$

where $\tilde{r}^*$ is the median of the estimated residuals $r^* = L^{-1}z^*$. The parameters of the model were assumed to be known in the previous formula. If they were unknown, they would be replaced with their estimates, as usual.

3.4 Description of the Proposed Procedure

For outlier treatment, the proposed procedure assumes that the orders (p, d, q) of model (2) are known and it proceeds iteratively. In the first stage, outliers are detected one by one and the model parameters are modified after each outlier has been detected. When no more outliers have been detected, the procedure goes to the second stage, where a multiple regression is performed. The outliers with the lowest t–value is discarded and the procedure goes back to the first stage to iterate.

The procedure used to incorporate or reject outliers is similar to the stepwise regression procedure for selecting the "best" regression equation. This results in a more robust procedure than that of Chen and Liu (1993), which uses "backward elimination" and may therefore detect too many outliers in the first stage of the procedure.

Up to now, we have supposed that there were no regression effects, but it is easy to incorporate these effects into the proposed procedure. Let the series follow the regression model with ARIMA errors

$$z(t) = y'(t)\beta + \nu(t), \quad t = 1, \ldots, N, \tag{14}$$

where $\beta = (\beta_1, \ldots, \beta_k)'$ is the vector containing the regression parameters, which may include the mean as the first component, $\{z(t)\}$ is the observed series, $\{y(t)\}$ are the vectors containing the regression variables and $\{\nu(t)\}$ follows the ARIMA model (2) with $\mu = 0$. Then, the algorithm proposed for automatic detection and correction of outliers, described in detail, is the following

Initialization

If there are any regression variables in the model, included the mean, the regression coefficients are estimated by OLS and the series is corrected for their effects.

Stage I: Detection and estimation of outliers one by one

I.1) The ARIMA parameters are estimated, using Hannan–Rissanen's method and the series corrected for all regression effects present at the time, included the outliers so far detected.

I.2) Considering the estimates of the ARIMA parameters obtained in I.1 as fixed, the regression coefficients are estimated by GLS and their t statistics are computed. To this end, the fast algorithm of Morf, Sidhu and Kailath (1974) is used, followed by the QR algorithm. New estimated residuals are obtained.

I.3) With the estimated residuals obtained in I.2, the robust MAD estimator of the standard deviation of the residuals is computed.

I.4) If $u = (u_{d+1}, \ldots, u_N)'$, where d is the degree of the differencing operator, denotes the differenced series, the statistics $\hat{\tau}_{IO}(t)$, $\hat{\tau}_{AO}(t)$, $\hat{\tau}_{LS}(t)$ and $\hat{\tau}_{TC}(t)$ are computed for $t = d+1, \ldots, N$. To this end, the residuals computed in I.2 and the MAD obtained in I.3 are used. Let, for each $t = d+1, \ldots, N$, $\lambda_t = max\{|\hat{\tau}_{IO}(t)|, |\hat{\tau}_{AO}(t)|, |\hat{\tau}_{TC}(t)|, |\hat{\tau}_{LS}(t)|\}$. If $\lambda = max_t \lambda_t = |\hat{\tau}_{tp}(T)| > C$, where C is a pre–selected critical value, then there is a possible outlier of type tp at T. The subindex tp can be IO, AO, TC or LS. If no outlier has been found the first time the algorithm passes through this point, then stop. The series is free from outlier effects. If no outlier has been found, but it is not the first time that the algorithm passes through this point, then go to II.1. If, on the contrary, an outlier has been found, then correct the series for all regression effects, using the estimates obtained in I.2 and the last outlier coefficient estimate obtained while computing λ, and go back to I.1 to iterate.

Stage II: Multiple Regression

Using the estimates of the multiple regression and their t statistics obtained the last time the algorithm passed through I.2, check whether there are any outliers with a t statistic $< C$, where C is the same critical value than in I.4. If there aren't any, stop. If, on the contrary, there are some, then remove the one with the lowest absolute t–value and go back to I.2 to iterate.

4 Preliminary Tests and the Missing Observations Case

In this section, the algorithms of the previous sections are extended to the case in which there are missing observations. Also, tests for the log–level specification and for the presence of Trading Day and Easter effects are given.

4.1 Missing Observations

The proposed procedure can be extended easily to the case of missing observations. Missing observations are treated as additive outliers. This implies that we can work with a complete series, because the missing values are first assigned tentative values. Then, after the model has been estimated, the difference between the tentative value and the estimated regression coefficient is the interpolated value. See Gómez, Maravall and Peña (1997) for details.

Since we work with a complete series (there are no holes in it), we can use same algorithms described previously for automatic model identification and for automatic detection and correction of outliers. The tentative values assigned to the missing observations are the semisum of the two adjacent values.

4.2 Tests for the Log–Level Specification

The first test for the log–level specification is based on the maximum likelihood estimation of the parameter λ in the Box–Cox transformations. We fit an airline model with mean to the data, first in logs ($\lambda = 0$) and then without logs ($\lambda = 1$). Let $z = (z_1, \ldots, z_N)'$ be the differenced series and let T be a transformation of the data, which can be any of the Box–Cox transformations. It is assumed that $T(z)$ is normally distributed with mean zero and $\mathrm{Var}(T(z)) = \sigma^2\Sigma$. Then, the logarithm of the density function $f(z)$ of z is

$$\ln(f(z)) = k - \frac{1}{2}\left\{N\ln(\sigma^2) + \ln|\Sigma| + T(z)'\Sigma^{-1}T(z)/\sigma^2 + \ln(1/J(T))^2\right\},$$

where k is a constant and $J(T)$ is the jacobian of the transformation. Considering the parameter λ in the T transformation fixed, the previous density function is maximized first with respect to the other model parameters. It is easy to see that σ^2 can be concentrated out of this function by replacing it with the maximum likelihood estimator $\hat{\sigma}^2 = T(z)'\Sigma^{-1}T(z)/N$. The concentrated function is

$$l(z) = -\frac{1}{2}\left\{N\ln(T(z)'\Sigma^{-1}T(z)) + \ln(1/J(T))^2\right\} + \cdots,$$

where the dots indicate terms that do not depend on the value of λ associated with the transformation T. After having maximized with respect to all model parameters different from λ, we maximize with respect to λ. Denoting the sum of squares $T(z)'\Sigma^{-1}T(z)$ by $S(z,T)$, the maximum likelihood principle leads to the minimization of the quantity $S(z,T)(1/J(T))^{2/N}$. It is easy to see that $(1/J(T))^{1/N}$ is the geometric mean in the case of the logarithmic transformation, and unity in the case of no transformation. Therefore, the test compares the sum of squares of the model without logs with the sum of squares multiplied by the square of the geometric mean in the case of the model in logs. Logs are taken in case this last function is the minimum.

The other test is the following. First, the data is divided into groups of succesive observations of length l equal to a multiple of the number of observations per year s. That is, the first group is formed with the first l observations, the second group is formed with observations $l+1$ to $2l$, etc. Then, for each group, the observations are sorted and the smallest and largest observations are rejected. This is done for protection against outliers. With the other observations, the range and mean are computed. Finally, a range–mean regression of the form $r_t = \alpha + \beta m_t + u_t$ is performed. The criterion is the slope β. If β is greater than a specified value, logs are taken.

4.3 Trading Day and Easter Effects

Traditionally, six variables have been used to model the trading day effect. These are: (no. of *Mondays*) - (no. of *Sundays*), ... (no. of *Saturdays*) - (no. of *Sundays*).

The motivation for using these variables is that it is desirable that the sum of the effects of each day of the week cancel out. Mathematically, this can be expressed by the requirement that the trading day coefficients β_j, $j = 1,\ldots,7$, verify $\sum_{j=1}^{7}\beta_j = 0$, which implies $\beta_7 = -\sum_{j=1}^{6}\beta_j$.

Sometimes, a seventh variable, called the length–of–month variable, is also included. This variable is defined as $m_t - \bar{m}$, where m_t is the length of the month (in days) and $\bar{m} = 30.4375$ is the average month length.

There is the possibility of considering a more parsimonious modeling of the trading day effect by using one variable instead of six. In this case, the days of the week are first divided into two categories: working days and non–working days. Then, the variable is defined as (no. of (M, T, W, Th, F)) - (no. of $(Sat, Sun) \times 5/2$).

Again, the motivation is that it is desirable that the trading day coefficients β_j, $j = 1, \ldots, 7$ verify $\sum_{j=1}^{7} \beta_j = 0$. Since $\beta_1 = \beta_2 = \ldots = \beta_5$ and $\beta_6 = \beta_7$, we have $5\beta_1 = -2\beta_6$.

The Easter variable models a constant change in the level of daily activity during the d days before Easter. The value of d is usually supplied by the user.

The variable has zeros for all months different from March and April. The value assigned to March is equal to $p_M - m_M$, where p_M is the proportion of the d days that fall on that month and m_M is the mean value of the proportions of the d days that fall on March over a long period of time. The value assigned to April is $p_A - m_A$, where p_A and m_A are defined analogously. Usually, a value of $m_M = m_A = 1/2$ is a good approximation.

Since $p_A - m_A = 1 - p_M - (1 - m_M) = -(p_M - m_M)$, the sum of the effects of both months, March and April, cancel out, a desirable feature.

Since Trading Day and Easter effects are modeled by means of regression variables, a possible test for these effects is the following. If no model has been identified, specify an airline model with mean. Otherwise, use the identified model. Then, using the differenced series y, apply first the Kalman filter to move from model (16) to model (17), where β is the vector of regression parameters, that includes the Trading Day and/or Easter parameters. Since model (17) is an OLS model, we can use an ordinary F–test to test if all Trading Day parameters are zero or not. A t–test can be used to test if the Easter parameter is zero.

5 Computational Aspects

A Fortran program has been written by the author for personal computers, workstations and mainframes, that implements the proposed procedure. This program is part of a larger program called TRAMO, which the author has developed together with Agustín Maravall. The program is freely available at the internet address

To estimate the regression parameters of a regression model with ARIMA errors, when the autoregressive and moving average parameters of the ARIMA model are assumed to be known, the proposed procedure uses the following algorithm. Let the observed series $z = (z(1), \ldots, z(N))'$ follow the regression model with ARIMA errors

$$z = Y\beta + u, \tag{15}$$

where $\beta = (\beta_1, \ldots, \beta_k)'$ is the vector containing the regression parameters, which may include the mean as the first component, Y is an $N \times k$ matrix of full column rank and u follows the ARIMA model (2) with $\mu = 0$, which is supposed to be known. After differencing z, the columns of Y and u in (15), it is obtained that

$$y = X\beta + v, \tag{16}$$

where $y = (y(d+1), \ldots, y(N))'$, X is an $(N-d) \times k$ matrix, the components of $v = (v(d+1), \ldots, v(N))'$ follow the ARMA model $\phi(B)v(t) = \theta(B)a(t)$ and it is assumed that the degree of the differencing polynomial $\delta(B)$ is d.

If $Var(v) = \sigma^2\Omega$ and $\Omega = LL'$, with L lower triangular, is the Cholesky decomposition of Ω, then, premultiplying (16) by L^{-1}, it is obtained that

$$L^{-1}y = L^{-1}X\beta + L^{-1}v, \tag{17}$$

which is an OLS model. As described in Section 3.1, the Kalman filter can be applied to y and the columns of the X matrix to move from (16) to (17). β can now efficiently estimated in (17) by means of the QR algorithm. This last algorithm produces an orthogonal matrix Q such that $Q'L^{-1}X = (R', 0)'$, where R is a nonsingular upper triangular matrix. Partitioning $Q' = (Q_1, Q_2)'$ conforming to $(R', 0)'$, one can move from (17) to

$$\begin{aligned} Q_1'L^{-1}y &= R\beta + Q_1'L^{-1}v \\ Q_2'L^{-1}y &= + Q_2'L^{-1}v, \end{aligned}$$

from which $\hat{\beta} = R^{-1}Q_1'L^{-1}y$ and $\hat{\sigma}^2 = y'(L^{-1})'Q_2Q_2'L^{-1}y/(N-d-k)$ are easily obtained. The Q matrix is obtained by means of Housholder transformations.

6 Examples and Conclusions

The proposed procedure has been tested using real and simulated series and the results have been very satisfactory. The automatic model identification procedure, without performing detection and correction of outliers, has been applied to 36 series which follow models covering a very broad spectrum. The results are reported in appendix A. Of the 36 series, 13 are series which have appeared in published articles and the rest are simulated series. For the simulated series, the proposed procedure has identified the same model without any problem, whereas for the other series, the identified model has been the same or a better model. The results are similar to the ones obtained with the "SCA–Expert" modulus of the SCA software package, although the results of TRAMO are sometimes better.

In addition to these 36 series, 252 series corresponding to several components of the Spanish industrial production index have also been modeled, combining the automatic model identification procedure with the procedure for automatic detection and correction of outliers, since these series present a lot of outliers. Again, the results have been rather good in terms of the goodness of fit.

In order to illustrate the use of the proposed algorithm for automatic detection and correction of outliers, we will consider the example of the ozone (O_3) mean levels in Los Angeles city during the period of January 1955 to December 1972. This series was analyzed by Box and Tiao (1975) as an example of a series for which intervention analysis could be applied.

Box and Tiao (1975) identified three intervention variables and a multiplicative moving average model for the series differenced with seasonal difference. More specifically, the model is

$$\begin{aligned} z(t) = {} & \omega_1 \ \ INT1(t) + \frac{\omega_2}{1 - B^{12}} INT2S(t) \\ + {} & \frac{\omega_3}{1 - B^{12}} INT2W(t) + \frac{(1 + \theta_1 B)(1 + \theta_{12} B^{12})}{1 - B^{12}} a(t), \end{aligned}$$

where $INT1$ is 1 in January 1960 and the following months and 0 otherwise, $INT2S$ is 1 in the summer months, starting in June 1966, and 0 otherwise, and $INT2W$ is 1 in the winter months, starting in 1966, and 0 otherwise.

The series, together with its intervention variables, was subject to the proposed procedure for automatic detection and correction of outliers. Using

Table 1: Outliers Identified for the Ozone Series

Outlier	Estimate	t–value	Type
$t = 21$	2.359	3.37	AO
$t = 39$	-1.983	-3.40	TC
$t = 43$	-1.930	-3.30	TC

program TRAMO and a critical value $C = 3$, the following outliers were identified

The results are practically identical to the ones obtained with the SCA–Expert modulus of the SCA package for automatic detection and correction of outliers.

To illustrate the combined use of the automatic model identification procedure and the procedure for automatic detection and correction of outliers, we consider the example of the monthly retail sales in Department Stores of Hillmer, Bell and Tiao (1983). For the logged series, these authors identified the ARIMA model

$$\nabla\nabla_{12}z(t) = \frac{1+\theta_{12}B^{12}}{1+\phi_1 B+\phi_2 B^2}a(t). \tag{18}$$

When the model followed by the series is unknown and the series is believed to present outliers, the proposed procedure uses the algorithm of Section 1, which sequentially combines the automatic model identification procedure and the procedure for automatic detection and correction of outliers. Following the instructions given by the user, program TRAMO can go through up to three rounds. In the first round, it uses the default model and default critical value C, or the model and critical value entered by the user, and detects and corrects the series for outliers. The default model is the airline model of Box and Jenkins (1972) and the default critical value C depends on the series length. In the second round, using the outlier corrected series, the program automatically identifies a model and, with that model, it performs a second automatic detection and correction of outliers. For this second round, the program decreases the critical level C used for outlier detection and the user can control the amount of the decrease. Usually, these two rounds are enough to identify a model with a good fit. If this is not

Table 2: Outliers Identified for the Retail Sales Series

Outlier	Estimate	t–value	Type
$t = 45$	.096	5.23	TC
$t = 96$	.084	-4.38	AO
$t = 112$	-.176	-10.18	LS

the case, the program iterates. After the third round, it the fit is still not acceptable, the program specifies a general model. This general model is an ARMA$(3,1)(0,1)_s$ or an ARMA$(4,0)(0,1)_s$ for the differenced series, where the differencing orders are the same of the last round. At some point of the procedure, the identified model is compared to the airline model and the model with the best fit is chosen. This is done because the airline model is a robust model and departures from this model can be unstable.

Using program TRAMO in the manner just described, the following results were obtained for the retail sales series. In the first round, using the default model (the airline model) and a critical value $C = 3.5$, the program detected outliers at $t = 45$, of type TC, at $t = 96$, of type AO and at $t = 112$, of type LS. After correcting the series for the outlier effects, the program identified first the differencing polynomial $\delta(B) = \nabla\nabla_{12}$, without mean. Then, the program identified model (18). For this model,using a critical value of 3, the program detected the same outliers than before.

Again, the results are practically identical to the ones obtained with the SCA package, when first the SCA–Expert modulus and then the "Extended UTS" modulus for automatic detection and correction of outliers are used.

As a conclusion, we can say that the proposed procedure has proved robust and easy to implement and, at the same time, it has a solid theoretical background. In addition, since it uses linear regression techniques only, it is not computationally expensive. By combining both the modulus of automatic model identification and the modulus for automatic detection and correction of outliers, the productivity of the user can be greatly increased. If he is a time series expert, he is relieved of mundane and repetitive chores and if he is not an expert, he can use a methodology that previously could only be used by the experts.

Appendix A: Summary of the Automatic Model Identification for 36 Real or Simulated Series

Series	Simulated models or manually identified models	Model obtained by TRAMO	Comments
(1) Maddala (1972) Grunfeld's inversion series (N=20)	$(1+\phi_1 B+\phi_2 B^2)z(t) = C + a(t)$	same as left	
(2) Hillmer, Bell and Tiao (1983) Clothing sales (N=153)	$\nabla\nabla_{12}z(t) = (1+\theta_1 B+\theta_2 B^2) \times(1+\theta_{12}B^{12})a(t)$	same as left	
(3) Hillmer, Bell and Tiao (1983) Hardware sales (N=155)	$\nabla\nabla_{12}z(t) = (1+\theta_1 B) \times(1+\theta_{12}B^{12})a(t)$	same as left	
(4) Hillmer, Bell and Tiao (1983) Variety stores sales (N=153)	$(1+\phi_1 B+\phi_2 B^2)\nabla\nabla_{12}z(t) = (1+\theta_1 B)a(t)$	same as left	
(5) Box and Tiao (1983) Ozone sales (N=216)	$\nabla_{12}z(t) = C + (1+\theta_1 B) \times(1+\theta_{12}B^{12})a(t)$	same as left	
(6) Box and Tiao (1983) CPI series (N=234)	$\nabla z(t) = C+ (1+\theta_1 B)a(t)$	$(1+\phi_{12}B^{12}\nabla z(t) = (1+\theta_1 B)a(t)$	(b)
(7) Chatfield and Prothero (1973) Monthly sales series (N=77)	$(1+\phi_1 B)\nabla\nabla_{12}z(t) = (1+\theta_{12}B)a(t)$	$\nabla_{12}z(t) = C +(1+\theta_1 B+\theta_2 B^2)a(t)$	(c)
(8) Hamilton and Watts (1978) Weekday coffee data (N=178)	$(1+\phi_1 B)\nabla z(t) = (1+\theta_5 B^5)a(t)$	$\nabla z(t) = C+ (1+\theta_1 B) \times(1+\theta_5 B^5)a(t)$	(a)
(9) Box and Jenkins (1976) Series A (N=197)	$\nabla z(t) = (1+\theta_1 B)a(t)$, or $(1+\phi_1 B)z(t) = C+(1+\theta_1 B^1)a(t)$	$\nabla z(t) = (1+\theta_1 B)a(t)$	(d)
(10) Box and Jenkins (1976) Series C (N=226)	$(1+\phi_1 B)\nabla z(t) = a(t)$, or $\nabla^2 z(t) = (1+\theta_1 B^1+\theta_2 B^2)a(t)$	$(1+\phi_1 B)\nabla z(t) = a(t)$	(d)
(11) Box and Jenkins (1976) Series E (N=100)	$(1+\phi_1 B+\phi_2 B^2)z(t) = C+a(t)$, or $(1+\phi_1 B+\phi_2 B^2+\phi_3 B^3)z(t) = C+(1+\theta_1 B^1)a(t)$	$(1+\phi_1 B+\phi_2 B^2)z(t) = C+(1+\theta_1)a(t)$	(f)

Series	Simulated models or manually identified models	Model obtained by TRAMO	Comments
(12) Box and Jenkins (1976) Series F (N=70)	$(1+\phi_1 B)z(t) = C + a(t)$,	same as left	
(13) Box and Jenkins (1976) Series G (N=144)	$\nabla\nabla_{12}z(t) = (1+\theta_1 B)$ $\times(1+\theta_{12}B^1 2)a(t)$,	same as left	
(14) Ljung and Box (1979) Simulated series (N=75)	$z(t) = (1+\theta_1 B)a(t)$	same as left	
(15) Tsay and Tiao (1984) Simulated series with AR complex unit roots (N=100)	$(1+\phi_1 B+\phi_2 B^2)\nabla^2 z(t) =$ $(1+\theta_1 B)a(t)$	same as left	
(16) Simulated series with AR complex unit roots (N=80)	$(1+\phi_1 B+\phi_2 B^2)\nabla\nabla_4 z(t) =$ $(1+\theta_1 B+\theta_2 B^2)a(t)$	same as left	
(17) Box and Tiao Simulated series R1 (N=150)	$z(t) = C+(1+\theta_1+\theta_2 B^2)a(t)$	same as left	
(18) Box and Tiao Simulated series R2 (N=162)	$(1+\phi_1+\phi_2 B^2)z(t) =$ $C+a(t)$	same as left	
(19) Box and Tiao Simulated series R3 (N=147)	$\nabla z(t) = C+(1+\theta_1 B)a(t)$	same as left	
(20) Box and Tiao Simulated series R4 (N=161)	$\nabla z(t) = (1+\theta_1 B+\theta_2 B^6)a(t)$	$(1+\phi_6 B^6)\nabla z(t) =$ $(1+\theta_1 B)a(t)$	(e)
(21) Box and Tiao Simulated series R5 (N=155)	$\nabla^2 z(t) =$ $(1+\theta_1 B+\theta_2 B^2)a(t)$	same as left	
(22) Box and Tiao Simulated series R6 (N=178)	$(1+\phi_1 B+\phi_2 B^2)\nabla z(t) =$ $a(t)$	same as left	
(23) Box and Tiao Simulated series R7 (N=149)	$(1+\phi_1 B)z(t) = C + a(t)$	same as left	

Series	Simulated models or manually identified models	Model obtained by TRAMO	Comments
(24) Box and Tiao Simulated series R8 (N=148)	$(1+\phi_1 B)\nabla z(t) = C + a(t)$	same as left	
(25) Box and Tiao Simulated series R9 (N=151)	$\nabla z(t) = (1+\theta_1 B)a(t)$	same as left	
(26) Box and Tiao Simulated series R10 (N=146)	$\nabla z(t) = C$ $+(1+\theta_1 B+\theta_2 B^2)a(t)$	same as left	
(27) Box and Tiao Simulated series S1 (N=150)	$\nabla\nabla_{12} z(t) = (1+\theta_1 B)$ $\times(1+\theta_2 B^{12})a(t)$	same as left	
(28) Box and Tiao Simulated series S2 (N=162)	$(1+\phi_1 B)\nabla_{12} z(t) =$ $(1+\theta_1 B^{12})a(t)$	same as left	
(29) Box and Tiao Simulated series S3 (N=147)	$\nabla z(t) = (1+\theta_1 B^{12})a(t)$	same as left	
(30) Box and Tiao Simulated series S4 (N=161)	$(1+\phi_1 B^{12})z(t) =$ $C+(1+\theta_1 B)a(t)$	same as left	
(31) Box and Tiao Simulated series S5 (N=155)	$(1+\phi_1 B)\nabla_{12} z(t) = C + a(t)$	same as left	
(32) Box and Tiao Simulated series S6 (N=178)	$\nabla\nabla_4 z(t) = (1+\theta_1 B)a(t)$	same as left	
(33) Box and Tiao Simulated series S7 (N=149)	$\nabla^2(1+\phi_1 B^6)z(t) = a(t)$	same as left	
(34) Box and Tiao Simulated series S8 (N=148)	$(1+\phi_1 B+\phi_2 B^2)z(t) =$ $C+(1+\theta_1 B^6)a(t)$	same as left	
(35) Box and Tiao Simulated series S9 (N=151)	$\nabla_{12} z(t) = C+(1+\theta_1 B)a(t)$	same as left	
(36) Box and Tiao Simulated series S10 (N=146)	$\nabla\nabla_{12} z(t) =$ $(1+\theta_1 B+\theta_2 B^{12})a(t)$	$\nabla\nabla_{12} z(t) = (1+\theta_1 B)$ $\times(1+\theta_2 B^{12})a(t)$	(e)

(a) The model obtained by TRAMO is better, although the original model is also acceptable

(b) The model obtained by TRAMO is better. However, the seasonality is small ($\phi_1 = -.17$) and may be neglected.

(c) The model obtained by TRAMO is better. The model used in the original article is overdifferenced.

(d) In the original book, two alternative models were considered. TRAMO obtains the best one.

(e) TRAMO uses, in its automatic option, multiplicative models due to their simplicity and that they have less problems with nonstationarity and noninvertibility. However, by selecting a non–automatic option, the analyst may use non–multiplicative models if he prefers to do so.

(f) In the original book, two alternative models were considered. TRAMO obtains a model better than any of them.

References

1. Abraham, B. and Box, G. E. P., (1979), "Bayesian Analysis of Some Outlier Problems in Time Series", *Biometrika*, **66**, 229-236.

2. Akaike, H. (1978), "A Bayesian Analysis of the Minimum AIC Procedure", *Ann. Inst. Statis. Math.*, **30 A**, 9-14.

3. Akaike, H. (1979), "A Bayesian Extension of the Minimum AIC Procedure of Autoregressive Model Fitting", *Biometrika*, **66**, 237-242.

4. Box, G.E.P., and Jenkins, G.M., (1976), *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco.

5. Box, G.E.P., and Tiao, G.C., (1975), "Intervention Analysis with Applications to Economic and Environmental Problems", *Journal of the American Statistical Association*, **70**, 70-79.

6. Brockwell, P., and Davis, R., (1992), *Time Series: Theory and Methods, Second Edition*, Springer-Verlag, Berlin.

7. Chang, I., Tiao, G.C. (1983), "Estimation of Time Series Parameters in the Presence of Outliers", Technical Report 8, University of Chicago, Statistics Research Center.

8. Chang, I., Tiao, G.C. and Chen, C., (1988), "Estimation of Time Series Parameters in the Presence of Outliers", *Technometrics*, **30**, 193-204.

9. Chen, C. and Liu, L., (1993), "Joint Estimation of Model Parameters and Outlier Effects in Time Series", *Journal of the American Statistical Association*, **88**, 284-297.

10. Chen, C. and Tiao, G. C., (1990), "Random Level Shift Time Series Models, ARIMA Approximation, and Level Shift Detection", *Journal of Business and Economic Statistics*, **8**, 170-186.

11. Choi, B., (1992), ARMA *Model Identification*, Springer Verlag, New York.

12. Fox, A. J., (1972), "Outliers in Time Series", *Journal of the Royal Statistical Society*, Ser. B, **34**, 350-363.

13. Gill, P. E., Murray, W. and Wright M. H., (1992) *Practical Optimization*, Academic Press, London.

14. Gómez, V., Maravall, A., and Peña, D. (1996), "Missing Observations in ARIMA Models: Skipping Strategy Versus Additive Outlier Approach", *Journal of Econometrics* (Forthcoming).

15. Guttman, I. and Tiao, G. C., (1978), "Effect of Correlation on the Estimation of a Mean in the Presence of Spurious Observations", *The Canadian Journal of Statistics*, **6**, 229-247.

16. Hamilton, D. C. and Watts, D. G., (1978), "Interpreting Partial Autocorrelation Function of Seasonal Time Series Models", *Biometrika*, **65**, 135-140.

17. Hannan, E. J. and Rissanen, J., "Recursive Estimation of Mixed Autoregressive-Moving Average Order", *Biometrika*, **69**, 81-94.

18. Ljung, G. M. and Box, G. E. P., (1979), "The Likelihood Function of Stationary Autoregressive-Moving Average Models", *Biometrika*, **66**, 265-270.

19. Liu, L. M., (1989), "Identification of Seasonal ARIMA Models Using a Filtering Method", *Communications in Statistics*, **18**, 2279-2288.

20. Maddala, G.S., (1977), *Econometrics*, New York: McGraw-Hill, Capítulo 12, 280-281.

21. Mélard, G., (1984), "A Fast Algorithm for the Exact Likelihood of Autoregressive-Moving Average Models", *Applied Statistics*, **35**, 104-114.

22. Morf, M., Sidhu, G.S., and Kailath, T., (1974), "Some New Algorithms for Recursive Estimation on Constant, Linear, Discrete- Time Systems", *IEEE Transactions on Automatic Control*, **AC-19**, 315-323.

23. Pearlman, J.G., (1980), "An Algorithm for the Exact Likelihood of a High-Order-Autore- gressive-Moving Average Process", *Biometrika*, **67**, 232-233.

24. Tiao, G. C. and Tsay, R. S., (1983), "Consistency Properties of Least Squares Estimates of Autoregressive Parameters in ARMA Models", *The Annals of Statistics*, **11**, 856-871.

25. Tsay, R. S., (1984), "Regression Models With Times Series Errors",*Journal of the American Statistical Association*, **79**, 118-124.

26. Tsay, R. S., (1986), "Time Series Model Specification in the Presence of Outliers",*Journal of the American Statistical Association*, **81**, 132-141.

27. Tsay, R. S., (1988), "Outliers, Level Shifts, and Variance Changes in Time Series",*Journal of Forecasting*, **7**, 1-20.

28. Tsay, R. S. and Tiao, G. C., (1984), "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Functions for Stationary and Nonstationary ARMA models", *Journal of the American Statistical Association*, **79**, 84-96.

29. Zhao-Guo, C., (1985), "The Asymptotic Efficiency of a Linear Procedure of Estimation for ARMA Models", *Journal of Time Series Analysis*, **6**, 53-62.