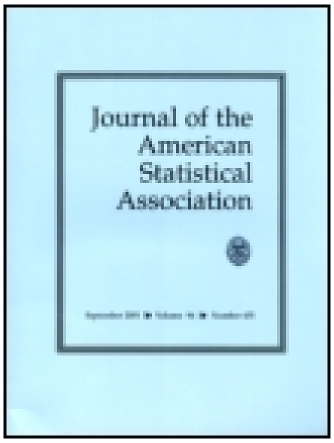

Journal of the American Statistical Association
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uasa20

Time Series Model Specification in the Presence of Outliers

Ruey S. Tsay [a]
[a] Department of Statistics , Carnegie-Mellon University , Pittsburgh , PA , 15213 , USA
Published online: 12 Mar 2012.

Time Series Model Specification in the Presence of Outliers

RUEY S. TSAY*

Outliers are commonplace in data analysis. Time series analysis is no exception. Noting that the effect of outliers on model identification statistics could be serious, this article is concerned with the problem of time series model specification in the presence of outliers. An iterative procedure is proposed to identify the outliers, to remove their effects, and to specify a tentative model for the underlying process. The procedure is essentially based on the iterative estimation procedure of Chang and Tiao (1983) and the extended sample autocorrelation function (ESACF) model identification method of Tsay and Tiao (1984). An example is given. Properties of the proposed procedure are discussed.

KEY WORDS: Autoregressive moving average model; Extended sample autocorrelation function; Model identification.

1. INTRODUCTION

Aberrant observations are often encountered in data analysis. Time series analysis is no exception. It is therefore logical to employ models that can reflect this fact and to use methods that are able to specify such models in practical time series modeling. In this article, I employ two commonly used models—additive and innovational—to characterize aberrant observations in a time series and propose an iterative procedure to detect them, to classify the type (additive or innovational) of each of them, and ultimately to specify a tentative model for a given time series in the presence of such aberrant observations.

For an outlier-free time series Z_t, assume that it follows an autoregressive moving average [ARMA(p, q)] model,

$$\Phi(B)Z_t = \theta(B)a_t, \tag{1.1}$$

where $\Phi(B) = 1 - \Phi_1 B - \cdots - \Phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ are polynomials in B, B is the backshift operator such that $BZ_t = Z_{t-1}$, and $\{a_t\}$ is a sequence of white noise random variables, iid with mean 0, variance σ^2, and finite fourth moment. For the model (1.1), assume that all of the zeros of $\Phi(B)$ are on or outside, and those of $\theta(B)$ are outside, the unit circle and that $\Phi(B)$ and $\theta(B)$ have no common factors. Furthermore, when some of the zeros of $\Phi(B)$ are on the unit circle, also assume that the process Z_t starts at a finite time point t_0.

Aberrant observations are usually referred to as *outliers* in statistics. Intuitively, outliers are discrepant observations that look discordant from most observations in a data set. For a time series in (1.1), these spurious observations may result from a gross error, for example, a recording or typing error, or from a nonrepetitive exogenous intervention, for example, a strike or an oil crisis in an economic series. Following Fox (1972), Abraham and Box (1979), Denby and Martin (1979), and Chang and Tiao (1983), I employ two modified models of (1.1) to describe the generating mechanism of an outlier. They are innovational outlier (IO) and additive outlier (AO) models. For the simple case of a single outlier, the models can be written, respectively, as

$$X_t = Z_t + [\theta(B)/\Phi(B)]\omega\xi_t^{(T)} \tag{1.2}$$

and

$$X_t = Z_t + \omega\xi_t^{(T)}, \tag{1.3}$$

where X_t is the observed series, Z_t is the unobservable outlier-free series in (1.1), ω represents the magnitude of the outlier, and $\xi_t^{(T)} = 1$ if $t = T$, and $=0$ otherwise, is a time indicator signifying the time occurrence of the outlier. Thus the AO model can be regarded as a "gross error" model, because only the level of the Tth observation is affected. On the other hand, the IO model represents a disturbance on the white noise series a_t at time point T, which influences the process Z_t on Z_T, $Z_{T+1}, \ldots$, through the dynamic structure $\theta(B)/\Phi(B)$. In practice, a time series might contain several outliers of various types, and the general outlier model is written as

$$X_t = \sum_{j=1}^{m} \omega_j v_j(B)\xi_t^{(T_j)} + Z_t$$

$$Z_t = [\theta(B)/\Phi(B)]a_t, \tag{1.4}$$

where $v_j(B) = 1$ for an AO, $v_j(B) = [\theta(B)/\Phi(B)]$ for an IO at time $t = T_j$, and m is the number of outliers.

In recent years, the estimation problem of model (1.4) has been widely investigated. When the order of Z_t is known, Guttman and Tiao (1978), Miller (1980), and Chang (1982) showed that the existence of outliers may cause serious bias in estimating the autoregressive (AR) and moving average (MA) parameters. If the generating mechanism of an outlier is also available, then these biases can often be reduced by using the intervention analysis technique of Box and Tiao (1975). When the timing and type of an outlier are unknown, Abraham and Box (1979) proposed a Bayesian approach, Martin (1980) a robust method, and Chang and Tiao (1983) an iterative procedure for resolving this estimation problem. However, the number, the timing and type of outliers, and the order of Z_t are seldom known a priori in practical modeling. It is therefore necessary to develop methods that can specify an appropriate model (or models) for a time series in the presence of outliers. Using the idea of M estimates, Martin (1980) modified Akaike's information criterion (Akaike 1974) for order selection in modeling AR processes. Here I adopt another approach. Though I agree fully with Martin that robust procedures are important and useful, I also believe that outliers sometimes provide useful

* Ruey S. Tsay is Assistant Professor, Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA 15213.

Journal of the American Statistical Association
March 1986, Vol. 81, No. 393, Theory and Methods

information that is too expensive to be ignored. Therefore, in this article I prefer to look into the outliers and to understand the causes so that appropriate handling methods can be used. In addition, I consider the general ARMA models instead of the purely autoregressive processes. Roughly speaking, I propose an iterative procedure consisting of identification–detection–removing cycles to solve the model specification problem for stationary and nonstationary ARMA series. The method essentially combines the iterative detection procedure of Chang and Tiao (1983) and the extended sample autocorrelation function (ESACF) identification method of Tsay and Tiao (1984). Some changes in estimation, however, are made to simplify the computation burden. The idea and the proposed procedure are given in Section 2. Section 3 applies the procedure to the annual spirits consumption data of the United Kingdom. Finally, some properties of the proposed procedure are discussed in Section 4.

2. TENTATIVE MODEL SPECIFICATION

2.1 Effects of Outliers on Identification Statistics

Suppose that the process Z_t in (1.1) is stationary, that is, all of the zeros of $\Phi(B)$ are outside the unit circle. In this case, the sample autocorrelation function (SACF) and sample partial autocorrelation function (SPACF) of Z_t are often used to determine the order (p, q), especially for pure AR or pure MA models. This is the well-known Box–Jenkins approach for model identification. In the literature, properties of these identification statistics have been well derived (e.g., Anderson and Walker 1964, Bartlett 1946, Fuller 1976, and Kendall and Stuart 1961). Consider now the model (1.4), where the observed series X_t are subject to outlier influences. Since Z_t is unobservable, the SACF and SPACF of X_t are usually treated as those of Z_t and used for model identification. Obviously, these statistics are contaminated by the existing outliers. The problem then is, What are the effects of outliers on these statistics? A similar problem occurs to the ESACF of Tsay and Tiao (1984) for specifying nonstationary and stationary mixed ARMA models.

The effects of outliers on SACF and SPACF have been investigated by Chang (1982) and that on ESACF by Tsay (1984). The results in general are complicated, involving lengthy algebraic expressions, but the message is clear. It says that the existence of outliers may cause substantial biases in SACF, SPACF, and ESACF and hence can seriously jeopardize their function as model identification tools. In summary, besides the underlying model of Z_t, which can carry the effects through the serial dependence, the biases depend on the number, the type, the magnitude, and the relative position of outliers. For small- or moderate-sized data sets, the biases can cause underspecification as well as overspecification in using SACF, SPACF, and ESACF. Note that the effect of the relative position of outliers can be regarded as "interaction" between outliers and has important practical implications, especially in model checking. Interested readers are referred to Chang (1982) and Tsay (1984) for details.

2.2 Motivation and Outline

Since the effects of outliers on identification statistics such as SACF, SPACF, and ESACF could be serious, it is necessary to consider some methods that can specify an appropriate model for Z_t when outliers are present. It is intuitive that an outlier can always be properly handled if it can be identified. The key, therefore, is to detect the outlier and to understand its nature. Many techniques are available or can be derived for detecting a single outlier. Complication, however, may arise in practice from multiple outliers, and it is impractical to consider a detection procedure for every fixed number of outliers, because the number of outliers is unknown. A simple and direct approach, then, is to use an iterative procedure that identifies the "most" apparent outlier in each iteration. This consideration governs the research of this article. I discuss an iterative method consisting of identification–detection–removing cycles to reduce the outlier effects so that the ESACF can still be used to specify a suitable model for a time series in the presence of outliers. The goal here is to identify, by means of an iterative identification procedure, an appropriate model of (1.4) for an outlier-contaminated time series. Experience of analyzing real examples shows that the proposed method not only can identify the number of outliers but also specifies the timing and type of each outlier identified.

The proposed method is essentially based on the iterative estimation procedure of Chang and Tiao (1983; see also Hillmer, Bell, and Tiao 1983) and the identification technique ESACF of Tsay and Tiao (1984). Several changes in estimation, however, have been made so that the procedure can efficiently handle the data of outlier-contaminated time series. For example, instead of maximizing the complicated likelihood function as Chang and Tiao (1983) did, the proposed method only uses the least squares method, that is, the linear regression techniques, to obtain parameter estimates. More specifically, the proposed method makes use of some consistency properties of the iterated AR estimates (e.g., see Theorem 2.2 of Tsay and Tiao 1984) to obtain estimates for the AR parameters and employs a bias-corrected Durbin's method (Durbin 1959) to compute the MA estimates. These changes can substantially reduce the computation involved in the iterative identification process and greatly increase the practical value of the proposed procedure, because as clearly pointed out by Martin (1980) and Chang and Tiao (1983), computation is one of the major difficulties in handling outliers in time series data analysis. In addition, since it only uses linear regression techniques, the proposed method can be easily implemented in any existing linear regression package. In what follows, I first consider the outlier detection technique and the parameter estimation method to be used in the proposed identification procedure.

2.3 Outlier Detection

Let $\pi(B) = \Phi(B)/\theta(B) = 1 - \pi_1 B - \pi_2 B^2 - \cdots$ and $e_t = \pi(B)X_t$. The single outlier models (1.2) and (1.3) can be rewritten as

$$e_t = \omega\xi_t^{(T)} + a_t \tag{2.1}$$

and

$$e_t = \omega\pi(B)\xi_t^{(T)} + a_t, \tag{2.2}$$

respectively. From (2.1), (2.2), and the least squares theory, the magnitude ω of an outlier can be estimated by

$$\hat{\omega}_{I,T} = e_T \quad \text{(IO)} \tag{2.3}$$

and

$$\tilde{\omega}_{A,T} = \eta^2\pi(B)e_T$$
$$= \eta^2(1 - \pi_1 F - \pi_2 F^2 - \cdots - \pi_{n-T}F^{n-T})e_T \quad \text{(AO)}, \qquad (2.4)$$

where $\eta^2 = (1 + \pi_1^2 + \pi_2^2 + \cdots + \pi_{n-T}^2)^{-1}$ and F is the forward shift operator such that $Fe_t = e_{t+1}$. The variances of these estimates are $\text{var}(\tilde{\omega}_{I,T}) = \sigma^2$ and $\text{var}(\tilde{\omega}_{A,T}) = \eta^2\sigma^2$, respectively, where σ^2 is the variance of a_t. Based on these results, Chang and Tiao (1983) proposed the following testing statistics for outlier detection:

$$\lambda_{I,T} = \tilde{\omega}_{I,T}/\sigma \quad \text{(IO)} \qquad (2.5)$$

and

$$\lambda_{A,T} = \tilde{\omega}_{A,T}/\eta\sigma \quad \text{(AO)}. \qquad (2.6)$$

Under the null hypothesis that there is no outlier, both $\lambda_{I,T}$ and $\lambda_{A,T}$ have the standard normal distribution and hence can readily be used in practical modeling. Of course, the true parameters π and σ^2 are usually unknown in the modeling stage, but they can be estimated by any consistent estimators. Similarly, the timing T is seldom known a priori, but one may check every time point T for $T = 1, 2, \ldots, n$. In other words, one employs

$$\lambda_I = \max_{\{T:1\le T\le n\}}|\lambda_{I,T}| \quad \text{(IO)} \qquad (2.7)$$

and

$$\lambda_A = \max_{\{T:1\le T\le n\}}|\lambda_{A,T}| \quad \text{(AO)} \qquad (2.8)$$

as testing criteria for outlier detection. By comparing these testing statistics with some critical value C, one can determine the existence of outliers. The time points at which the above maxima occur are the timings of the corresponding outliers. The possible values of C suggested by Chang and Tiao are 3.0, 3.5, and 4.0.

A likelihood ratio test was also derived by Chang and Tiao (1983) to distinguish an AO from an IO. But the authors further suggested that by comparing the testing statistics λ_I and λ_A of (2.7) and (2.8), respectively, one may specify the type of an outlier without requiring any additional computation. More precisely, if λ_I is greater than λ_A, an IO is found; otherwise, an AO is identified (see Chang and Tiao 1983 or Hillmer et al. 1983 for details).

2.4 Parameter Estimation

I now briefly discuss the estimation method used in the proposed iterative procedure. If the order (p, q) of model (1.1) is known, Theorem 2.2 of Tsay and Tiao (1984) shows that consistent estimates of the AR parameters can be obtained from the qth iterated AR(p) regression. Tsay [1984, Corollary 2.6, Theorem 3.3(b), and (2.15)] further showed that this property continues to hold for most of the outlier-contaminated series, although the number of iterations needed might be greater than q. In practice, this means that for the outlier-contaminated series, the iterated AR estimates of AR(p) regression would still demonstrate a constancy pattern starting from some ($j \ge q$) if p is the correct AR order. Without repeating the complicated formulas and definitions of the two papers, I illustrate the result by considering the temperature data, Series C, of Box and Jenkins (1976). Table 1 gives the iterated AR estimates of AR(2) regression for the original data and an artificially contaminated series. Here the contaminated data set is obtained from the original one by changing X_{71} from 25.8 to 28.5, resulting in an AO of $\omega = 2.7$ at $t = 71$. This change can simply be regarded as a typing error. For the original data, the iterated estimates of AR(2) regression show a constancy pattern starting at $j = 0$. The same pattern is also obtained for the contaminated data, although it appears only after $j = 2$. Based on this property, one may obtain consistent AR estimates by checking the iterated AR estimates once a tentative order has been specified.

Table 1. Least Squares Estimates of AR Parameters in AR(2) Regression of Series C With and Without X_{71} Being Changed From 25.8 to 28.5

Iteration j	With		Without	
0	1.12	−.15	1.81	−.82
1	3.28	−2.27	1.83	−.84
2	1.91	−.92	1.84	−.85
3	1.81	−.82	1.89	−.90
4	1.89	−.90	1.92	−.93
5	1.79	−.80	1.92	−.92

I now turn to the MA parameters. After the consistent AR estimates have been selected, X_t is transformed into W_t by treating the AR estimates as the true AR parameters. That is,

$$W_t = X_t - \sum_{i=1}^{p} \hat{\Phi}_i X_{t-i}, \qquad (2.9)$$

where $\hat{\Phi}_i$'s are the specified AR estimates. Since $\hat{\Phi}_i$'s are consistent, W_t would asymptotically follow a pure MA(q) model. One therefore may employ a bias-corrected Durbin's method (Durbin 1959) to derive estimates for the MA parameters. Roughly speaking, Durbin's method is based on the consideration that any invertible MA process can be approximated as accurately as possible by an AR process, say for instance,

$$W_t = \beta_1 W_{t-1} + \cdots + \beta_k W_{t-k} + f_t. \qquad (2.10)$$

Clearly, the true MA parameters are functions of the AR coefficients β_i in (2.10). This implies that in practice one can first fit an AR(k) regression to W_t with a sufficiently large k and then solve a system of linear equations to obtain estimates for the MA parameters. The linear equations, of course, are based on the relationship between the MA parameters and the AR coefficients β_i of W_t in (2.10). To improve the efficiency of the MA estimates, one may modify Durbin's estimates by adjusting the corresponding biases. Hannan and Rissanen (1982) proposed a regression approach to estimate the biases up to order $O_p(n^{-1})$. Readers are referred to Durbin (1959) and Hannan and Rissanen (1982) for detailed procedures of estimating the MA parameters.

2.5 Identification Procedure

Using the outlier detection technique of Section 2.3 and the parameter estimation method of Section 2.4, I now propose an iterative method to specify tentative models for time series when the data are subject to outlier influences.

Step 1. Identification. Use the ESACF to identify a tentative order (p, q) for the data and obtain the AR estimates from the iterated AR(p) regression. Let $\hat{\Phi}_i$, $i = 1, \ldots p$, be the AR estimates. Usually, these are the coefficients in the qth iterated AR(p) regression. But for the outlier-contaminated series, one may use the constancy pattern of the iterated AR estimates of AR(p) regression to specify these AR estimates.

Step 2. MA Estimates. Using the AR estimates of Step 1, transform the data X_t into W_t by $W_t = X_t - \sum_{i=1}^{p} \hat{\Phi}_i X_{t-i}$. Then employ the bias-corrected Durbin's method to compute the MA estimates.

Step 3. Detection. Perform the outlier detection tests of Section 2.2 by treating the AR estimates of Step 1 and the MA estimates of Step 2 as the true parameters. In other words, compute the tests λ_I and λ_A of (2.7) and (2.8). If both λ_I and λ_A are less than a critical value C, terminate the identification procedure and go to Step 4. Based on limited experience, I found that $C = 3.5$ works well in most of the analyses. When either λ_I or λ_A is greater than C, an outlier is identified. If λ_I is greater than λ_A, an IO is found. In this case, let T be the corresponding time and $\omega_{I,T}$ the magnitude of this IO. Then one can remove this IO effect by modifying the observed data set according to the model (1.2). More precisely, one adjusts the data by the following equation:

$$\begin{aligned} \tilde{X}_t &= X_t, && \text{for } t = 1, 2, \ldots, T-1 \\ &= X_t - \Psi_{t-T}\omega_{I,T}, && \text{for } t = T, T+1, \ldots, n, \end{aligned} \tag{2.11}$$

where Ψ_i's are the Ψ weight of X_t based on the AR and MA estimates of Steps 1 and 2, that is, Ψ_i is the coefficient of B^i in $\Psi(B) = \hat{\theta}(B)/\hat{\Phi}(B)$. On the other hand, if λ_A is greater than

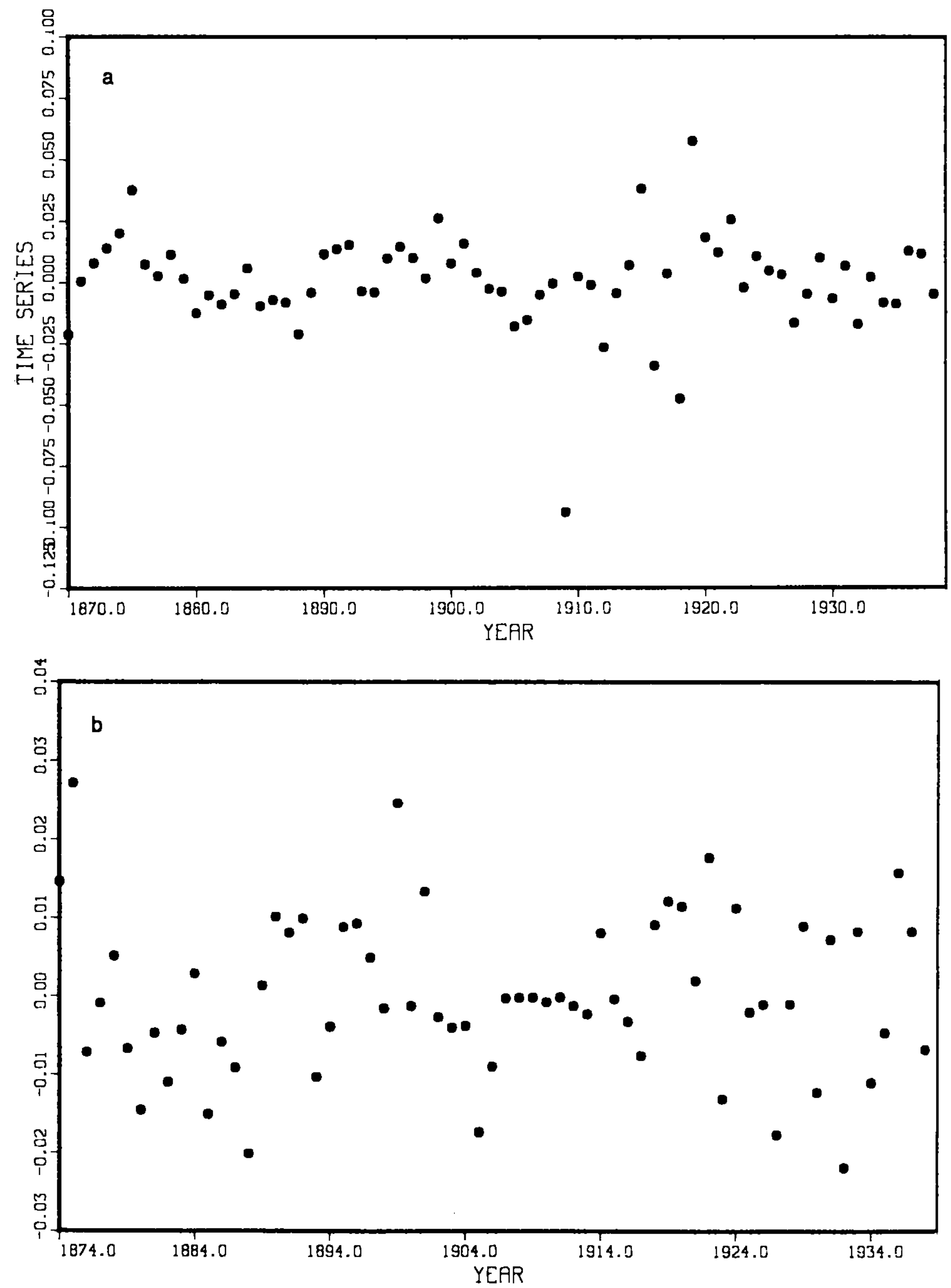

Figure 1. a, The Residual Plot of Model (3.2); b, the Residual Plot of Model (3.3).

λ_I, an AO is found, and one can modify the data according to the following equation:

$$\tilde{X}_t = X_t, \qquad \text{for } t \neq T$$
$$= X_T - \omega_{A,T},$$

where T is the time at which λ_A of (2.8) occurs and $\omega_{A,T}$ is the magnitude of the corresponding outlier.

After removing the outlier effect, one can iterate the procedure and go to Step 1 with the modified data.

Step 4. Summary and Checking. Summarize the tentative model as follows: (a) The order (p, q) in the last iteration of Step 1 is the tentative order for the outlier-free series Z_t; (b) each iteration except the last one identifies an outlier and also specifies the timing and type of the outlier. The number of outliers, therefore, is determined by the number of iterations. Consequently, a model of (1.4) is identified by this iterative method. After a tentative model is specified, one should examine the original data and search for the possible causes of the identified outliers. Experience shows that for a data analyst, understanding the cause and nature of an outlier is almost as important as being able to detect it. Very often understanding the source and reason of an outlier can provide insight into the series and hence suggest an appropriate method for handling the outlier. Thus at the final stage of the iterative identification procedure, one should always ask the following questions and modify the model in accordance with the answers to these questions: Why are those observations different from the others? Are there sufficient reasons to suspect those observations as "outliers"? Is the peculiar behavior of those observations caused by some exogenous disturbances? What is the nature of those disturbances?

Although the data are modified according to the outlier identified in the identification process, I do not suggest using the modified data as observed series in any further data analysis. On the contrary, once a tentative model is specified, one should use the original data, unless other information suggests the opposite, in the process of estimation, model checking, and forecasting. Note that it is possible that many λ_I and λ_A of Step 3 are greater than C in an iteration. I, however, do not suggest specifying more than one outlier in an iteration mainly because of the potential masking effects among the outliers.

3. AN ILLUSTRATIVE EXAMPLE

I now illustrate the iterative identification method by considering the data of annual consumption of spirits in the United Kingdom from 1870 to 1938. This data set has been analyzed by Prest (1949), Durbin and Watson (1951), and Fuller (1976). Strictly speaking, this data set belongs to the so-called time series regression models. It consists of a dependent variable and four explanatory variables. The dependent variable Y_t is the annual per capita consumption of spirits, and two of the explanatory variables, say X_{1t} and X_{2t}, are per capita income and price of spirits, respectively, both deflated by a general price index. These variables are in logarithms. The other two explanatory variables are the linear and squared terms of time trend. The model fitted by Prest is

$$Y_t = 2.14 + .69X_{1t} - .63X_{2t} - .0095t$$
$$- .00011(t - 35)^2 + e_t, \quad \sigma_e^2 = .000983, \quad (3.1)$$

where σ_e^2 is the residual mean square and t is an index of time defined by t = actual year − 1869. Based on the Durbin–Watson statistic, Fuller (1976) found that autocorrelation exists in the residual series. Moreover, using the SPACF, Fuller concluded that an AR(1) model is suitable for the residual and suggested the following model for the data:

$$Y_t = 2.36 + .72X_{1t} - .80X_{2t} - .0081t$$
$$- .000092(t - 35)^2 + Z_t$$
$$Z_t = .7633Z_{t-1} + h_t, \qquad \text{with } \sigma_h^2 = .000417. \quad (3.2)$$

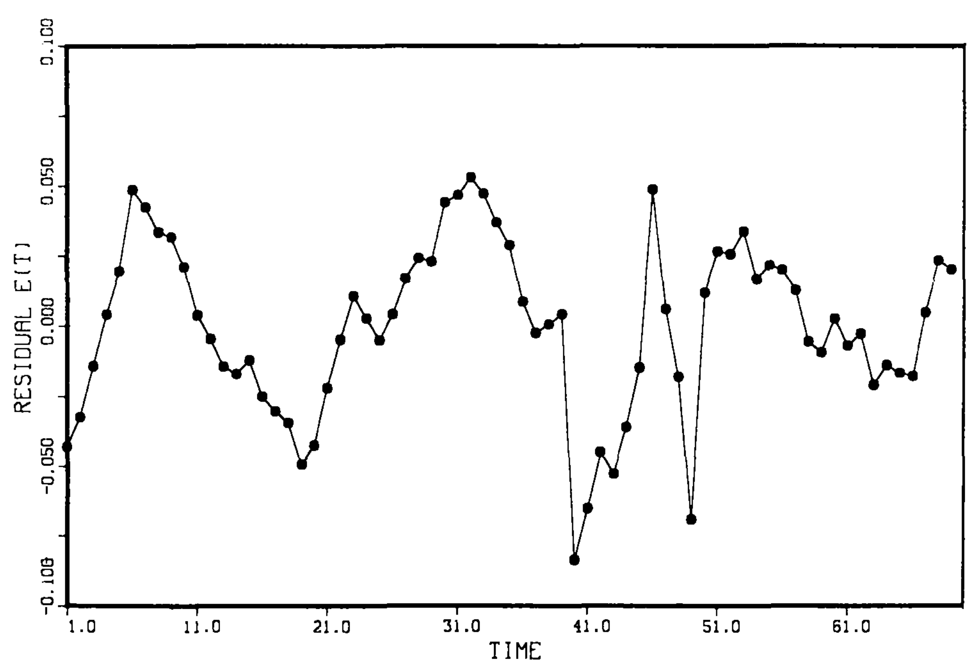

Figure 2. The Residual Plot of Model (3.1).

Table 2. ESACF of the Spirits Data

p	q = 0	1	2	3	4	5	6	7	8	9
					a. ESACF					
0	.72	.46	.25	.15	.00	−.13	−.18	−.27	−.34	−.51
1	.17	.14	.00	.16	.01	−.17	.00	−.07	.04	−.32
2	−.44	.13	.00	.17	.09	−.16	.01	−.04	−.01	−.34
3	.39	−.11	−.06	.15	.11	−.16	−.06	.05	.01	−.33
4	.17	−.23	−.42	.04	.11	−.17	−.03	.03	.04	−.23
5	−.50	.39	−.45	.18	.06	−.03	.16	−.07	.13	−.26
6	.27	−.24	−.47	−.22	−.04	.01	.14	.10	.08	−.25
					b. Indicator Symbols					
0	X	X	O	O	O	O	O	X	X	X
1	O	O	O	O	O	O	O	O	O	X
2	X	O	O	O	O	O	O	O	O	X
3	X	O	O	O	O	O	O	O	O	X
4	O	O	X	O	O	O	O	O	O	O
5	X	X	X	O	O	O	O	O	O	O
6	X	O	X	O	O	O	O	O	O	O

Compared with (3.1), the AR(1) modification does provide considerable improvement in terms of residual mean square. Some simple diagnostic checkings reveal, however, that further investigation is needed for this data set. Figure 1a shows the plot of the residual h_t of (3.2), which clearly does not behave as a white noise series. In fact, it not only displays several spurious values but also exhibits some serial correlation pattern. Thus in spite of its improvement over (3.1), the adequacy of model (3.2) is questionable.

Assume that the four explanatory variables in (3.1) are adequate for the spirits data. Pursue the investigation by treating the residual e_t of (3.1) as an observed time series. These observations are plotted in Figure 2. The proposed iterative identification procedure was applied to specify a tentative model for e_t. First, the ESACF of e_t were computed (see Table 2). Clearly, an AR(1) is suggested by the ESACF table. After an AR(1) is specified, a simple examination of the iterated AR estimate of AR(1) regression shows that $\hat{\Phi}_1 = .72$. We then proceed to test the existence of outliers. For simplicity, $C = 3.5$ is used as the ciritical value and $k = q + 5$ as the AR order in Durbin's method in the investigation. Based on this convention, an IO was found at $t = 40$ with $\omega_I = -.08663$ and $\lambda_{I,40} = -4.22$. The data were then modified according to this IO [see (2.11)] and the procedure was iterated. The identification process was terminated at iteration 7, with the detailed results given in Table 3. The ESACF of the modified data after removing the six outliers identified in Table 3 is shown in Table 4, which, except for a marginal X in column 4, clearly suggests an ARMA(1, 1) model. The initial parameter values of this ARMA(1, 1) model are given in the last row of Table 3. Thus the iterative identification procedure identifies an ARMA(1, 1) model with six outliers of various types for the residual series e_t of (3.1). This result of course is different from the AR(1) model suggested by the SPACF and by the original ESACF of e_t.

Next, check the six outliers identified by the proposed iterative procedure. Is it reasonable to suspect those observations as outliers? Why? What are the possible causes if they indeed are outliers? Why are there four consecutive IO's at $t = 40$, 41, 42, and 43? Why do the magnitudes of the first three IO's seem to decay rather exponentially? How can one specify a suitable model for e_t after understanding the reasons for the presence of these outliers?

A close examination of Figure 2 shows that (a) e_{40} and e_{49} are indeed the two lowest points of e_t, (b) the big drop at e_{40} has an impact on several points whereas that of e_{49} does not, and (c) the structure of e_t might have changed after $t = 40$. These facts seem to support the finding of the proposed iterative procedure.

In searching for the possible causes of the identified outliers, it is of interest to consider some historical facts of the United Kingdom. First, the actual years of e_{46} and e_{49}, both identified as AO, are 1915 and 1918, respectively. Since World War I started in August 1914 and we are dealing with annual data, these two AO's might well mark the beginning and the end of World War I. Second, e_{40} corresponds to 1909, which is the year Lloyd George, Prime Minister of U.K., marched his social reform program. It is not unreasonable to believe that the instituted heavy tax on liquor had a substantial impact on the consumption of spirits. Furthermore, it is logical to assume that the impact of this social reform program damped out gradually. This may explain, at least partially, why there are four consecutive IO's starting at e_{40} and damping out exponentially. These historical facts provide sufficient reasons to believe that those outliers identified by the proposed iterative procedure are indeed different from the other observations. Thus the proposed iterative procedure in this example correctly pinpoints the tim-

Table 3. Iterative Identification Results of the Spirits Data

Iteration	Suggested Model	Parameter Estimates: AR		MA	Outlier: Type	Time	Magnitude
1	AR(1)	.72			IO	40	−.08663
2	AR(1)	.72			AO	49	−.06628
3	AR(1)	.78			IO	41	−.06772
4	AR(2)	1.03	−.30		AO	46	.04662
5	AR(2)	1.19	−.40		IO	42	−.04653
6	ARMA(1, 1)	.72		−.30	IO	43	−.05379
7	ARMA(1, 1)	.74		−.28			

Table 4. ESACF Table of e_t in (3.1) After Six Iterations

p	q: 0	1	2	3	4	5	6	7	8	9
					a. ESACF					
0	.82	.59	.38	.20	.02	−.10	−.20	−.29	−.37	−.43
1	.33	.23	.20	.23	.05	−.11	−.14	−.10	−.15	−.12
2	−.35	−.06	−.08	.26	−.13	.02	−.04	.07	−.06	.01
3	−.49	.07	−.08	.20	−.12	−.03	.00	.08	−.06	.02
4	−.30	−.25	−.47	.14	−.13	−.05	−.01	.10	.06	.04
5	−.03	−.33	−.33	−.27	.00	.07	.09	.10	−.03	.04
6	−.05	−.15	−.09	−.22	−.01	.06	−.04	.02	.03	.05
					b. Indicator Symbols					
0	X	X	X	O	O	O	O	X	X	X
1	X	O	O	O	O	O	O	O	O	O
2	X	O	O	X	O	O	O	O	O	O
3	X	O	O	O	O	O	O	O	O	O
4	X	O	X	O	O	O	O	O	O	O
5	O	X	X	X	O	O	O	O	O	O
6	O	O	O	O	O	O	O	O	O	O

ings of those observations that require further consideration, even though it cannot tell us what went wrong with the data.

Finally, based on the possible causes and nature of the six outliers, I suggest the following model for the spirits data:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 t + \beta_4 (t - 35)^2 + e_t$$

$$e_t = [(\omega_1 + \omega_2 B^3)/(1 - \delta B)]\xi_t^{(40)} + \omega_3 \xi_t^{(46)} + \omega_4 \xi_t^{(49)} + [(1 - \theta B)/(1 - \Phi B)]a_t. \quad (3.3)$$

The intervention model of $\xi_t^{(40)}$ in (3.3) is based on the pattern of the impact of this intervention. As estimated in Table 3, the impacts at t = 40, 41, 42, and 43 are −.08663, −.06772, −.04653, and −.05379, respectively. Clearly, the first three decay exponentially but the last one drops a little bit farther.

The estimation results of model (3.3) along with the corresponding residual SACF are given in Table 5. This estimation is based on the exact likelihood method of the SCA system (Liu and Hudak 1983). It is clear that the residual mean square is further reduced to .000104, which is about one-quarter of that of (3.2). The residual a_t of (3.3) is plotted in Figure 1b, which fails to show any major model inadequacy. The Box–Pierce test of residual autocorrelation also fails to suggest any model discrepancy. The testing statistic is $Q_{12}^2 = 8.4$ and $\chi^2_{.05,10} = 18.3$. From Figure 1b, a somewhat large value appears at the beginning of the residual series, but it is only marginal compared with the corresponding standard deviation. A modified model was further entertained, but the fitted result is hardly changed from that of Table 5. As a matter of fact, the residual mean square only reduces to .000096. Thus (3.3) is adopted as the final model. Note that the final estimates of θ and Φ in (3.3) are in close agreement with the initial estimates given in Table 3. In addition, the effects of exogenous interventions coincide with the expectation.

4. DISCUSSION

In this section, I briefly discuss some properties of the proposed identification procedure and use simulation to illustrate further its efficacy. There are three major advantages of the proposed identification method. First, it is simple both conceptually and in computation. Under the model assumption (1.4), the procedure basically uses linear regression techniques to deal with a rather complicated problem: it uses autoregressions to obtain consistent estimates for both AR and MA parameters when the process is free from outlier contamination; it also uses simple regressions to identify the outliers when they are present. The proposed procedure, therefore, is conceptually simple and can easily be implemented in any existing linear regression package. Second, it provides a vehicle for data analysts to gain insight into a time series, because the procedure can point to those discordant observations that deserve further considerations. This integrated work is essential to data analysis. Finally, the procedure does not sacrifice any efficiency

Table 5. Estimation Results of Model (3.3)

	Parameter: β_0	β_1	β_2	β_3	β_4	ω_1	ω_2	δ	ω_3	ω_4	θ	Φ	σ_a^2
Estimate	2.3303	.6907	−.7381	−.00865	−.000097	−.0923	−.0237	.7592	.0418	−.0587	−.3085	.8003	.000104
S.D.	.1857	.0876	.0438	.00083	.000030	.0105	.0097	.1034	.0071	.0071	.1354	.0872	
						Lag 1–Lag 12							
Residual SACF	.04	.14	.03	−.00	−.02	−.08	−.01	−.11	−.03	−.23	−.08	−.10	
S.D.	.12	.12	.13	.13	.13	.13	.13	.13	.13	.13	.14	.14	

NOTE: S.D. = standard deviation.

Table 6. Summary of the Z_t Process of (4.1b): 400 Repetitions, Each With 100 Observations

p	q = 0	1	2	3	4	5	6	7
	a. Average ESACF							
0	.47	.35	.26	.18	.14	.09	.06	.04
1	−.30	.02	.02	.02	.02	.00	.01	.00
2	−.18	−.05	.02	.01	.01	.00	.00	.00
3	.00	.02	−.01	.00	.01	.00	.00	.00
4	.01	.06	.00	.00	.01	.00	.01	.00
5	.00	.06	.01	.04	.01	.00	.01	.00
6	.00	.07	−.01	.02	.00	.01	.01	.00
7	−.02	.07	−.02	.02	.01	.03	.00	.00
	b. Sample Standard Errors of ESACF							
0	.11	.12	.13	.13	.13	.13	.13	.13
1	.16	.13	.12	.11	.10	.09	.09	.08
2	.30	.18	.11	.10	.10	.09	.08	.08
3	.35	.25	.15	.09	.09	.08	.08	.07
4	.35	.27	.21	.14	.09	.08	.08	.07
5	.35	.27	.22	.18	.13	.08	.08	.07
6	.35	.27	.23	.20	.17	.11	.07	.07
7	.35	.26	.23	.19	.18	.15	.11	.07
	c. Percentage: ESACF Greater Than Its Two Standard Errors in Modulus							
0	.99	.85	.49	.23	.08	.04	.02	.02
1	.75	.09	.08	.05	.02	.01	.01	.01
2	.74	.25	.07	.02	.02	.01	.01	.01
3	.74	.41	.14	.02	.02	.01	.00	.01
4	.73	.48	.29	.10	.03	.00	.01	.01
5	.71	.47	.33	.22	.09	.01	.01	.01
6	.73	.48	.33	.21	.14	.02	.01	.01
7	.73	.47	.31	.19	.15	.11	.04	.00

in model specification for the outlier-free series. In contrast, most of the robust methods proposed in the literature tend to pay a certain "premium" against heavy-tailed deviations.

Like other methods, the proposed procedure has its disadvantages too. For instance, when there are k outliers in a time series, the method requires k iterations to identify them. This presumably could become cumbersome when the number of data sets is large.

Besides the example given in Section 3, many other real data sets have been used to test the proposed procedure, including Series B and C of Box and Jenkins (1976) and the RESEX series of Martin, Samarov, and Vandaele (1983). The results are very encouraging. Of course, only through more applications can the contribution of the proposed procedure be judged. Here I briefly state some justifications that support the applicability of the procedure, especially when the number of outliers is small. First, the ESACF approach can often specify a reasonable starting model for the iterated procedure. This assertion is based on two grounds: (a) since the biases discussed in Section 2.1 are of order $O(n^{-1})$, the distortion on ESACF due to outliers is usually not dramatic when the outliers are moderate and the sample size is relatively large; (b) the biases on iterated AR estimates used in ESACF become negligible when the number of iterations is large enough (see Table 1 and Tsay 1984, Corollary 2.6). Second, the outlier detection method of Section 2 is based on sound theoretical considerations when the parameters are known. Third, if an outlier is extremely large, then one should be able to recognize the corresponding observation and take appropriate actions accordingly before applying the proposed procedure. Finally, the iterative nature of the procedure reveals the outliers one by one, providing opportunities for interaction between data analysts and the data. A better understanding of the data can always help obtain an appropriate model.

Table 7. Iterated Estimates of AR(1) Regression for Table 6

Iteration	Expected Value	Sample Average	Standard Error of Average
0	.5231	.4762	.0053
1	.8	.7582	.0097
2	.8	.7344	.0144
3	.8	.6338	.0759

To further illustrate the efficacy of the proposed procedure, I conduct a simulation study. The simulation has two purposes: to compare the behavior of ESACF for outlier-free and outlier-contaminated data and to show that when the outlier effect is moderate, the ESACF can still specify the underlying model for Z_t and the employed outlier detection procedure can identify the existing outlier. The model used is

$$X_t = Z_t + 6.0\xi_t^{(50)}, \qquad t = 1, 2, \ldots, 100, \quad (4.1a)$$

$$(1 - .8B)Z_t = (1 - .4B)a_t, \qquad \sigma^2 = 1.0, \quad (4.1b)$$

Four hundred repetitions each with 100 observations were generated from model (4.1). Here the random normal variates a_t

Table 8. Summary of the X_t Process of (4.1a): 400 Repetitions, Each With 100 Observations

p	q = 0	1	2	3	4	5	6	7
	a. Average ESACF							
0	.37	.27	.20	.14	.11	.07	.05	.03
1	−.32	.02	.03	.02	.01	.00	.01	.00
2	−.21	−.08	.02	.01	.01	.00	.00	.00
3	−.05	.00	−.03	.01	.01	.00	.00	.00
4	−.02	.04	.00	.00	.01	.00	.00	.00
5	−.01	.06	.01	.03	.01	.00	.00	.00
6	−.01	.03	.01	.02	.00	.01	.01	.00
7	−.03	.04	.01	.03	.00	.03	.00	.00
	b. Sample Standard Errors of ESACF							
0	.11	.12	.12	.13	.12	.12	.12	.12
1	.17	.12	.11	.10	.09	.09	.09	.09
2	.31	.21	.10	.10	.09	.08	.08	.08
3	.36	.26	.18	.09	.09	.08	.07	.08
4	.36	.26	.22	.15	.08	.08	.07	.08
5	.35	.27	.23	.19	.15	.07	.07	.08
6	.35	.27	.23	.20	.17	.13	.07	.07
7	.35	.26	.22	.21	.18	.15	.12	.07
	c. Percentage: ESACF Greater Than Its Two Standard Errors in Modulus							
0	.93	.64	.39	.19	.07	.04	.03	.03
1	.80	.07	.06	.04	.01	.02	.01	.01
2	.77	.35	.04	.03	.01	.01	.01	.01
3	.75	.43	.23	.03	.01	.01	.00	.01
4	.76	.46	.31	.13	.01	.01	.00	.01
5	.74	.48	.32	.22	.11	.00	.00	.01
6	.75	.46	.32	.27	.15	.06	.00	.00
7	.73	.45	.31	.26	.16	.11	.06	.00

Table 9. The Iterated AR Estimates of AR(1) Regression for Table 8

Iteration	Expected Value	Sample Average	Standard Error of Average
0	.3927	.3745	.0056
1	.8	.7559	.0155
2	.8	.7439	.0359
3	.8	.6187	.1347

were obtained from the GGNML subroutine of the IMSL package. For each repetition, 400 Z_t and X_t were computed with zero initial values. Only the last 100 observations, however, were used in the analysis. This procedure was taken as a precaution to reduce the effect of initial values.

Tables 6–9 summarize the simulation results for Z_t and X_t, respectively. Table 6 gives the ESACF statistics, and Table 7, the iterated AR estimates of AR(1) regression for the outlier-free series Z_t. The percentage table (Table 6c) gives the percentage that the corresponding ESACF is greater than its two standard errors in modulus. Here the standard error is computed by using the Bartlett's formula of SACF. From these tables: (a) as expected, the average ESACF clearly suggests an ARMA(1, 1) model for Z_t. (b) The sample standard error table (Table 6b) indicates that the standard error of ESACF is in good agreement with the result of Bartlett's formula. For instance, for the first ESACF at lag 1, Table 6b shows its standard error as .13 and Bartlett's formula gives .11. (c) The percentage table indicates that the ESACF works well in identifying the underlying ARMA model. (d) The consistent property of iterated AR(1) estimates for ARMA(1, 1) models is clear (see Table 7).

Now, consider the outlier-contaminated series X_t. Tables 8 and 9 give the corresponding summary statistics for this series. Comparing Tables 6–9, one can see that except for the AR(1) estimate at the 0 iteration in Tables 7 and 9, the effect of the outlier is not severe. This is expected in the sense that the magnitude of the AO, $\omega = 6.0$, is only 4.15 var(Z_t) or 6 var(a_t). The ESACF, therefore, still suggests the generating model ARMA(1, 1) for X_t. Based on the specified order (p, q) = (1, 1), the AR(1) estimate at the 1 iteration is consistent. We may use it as the AR parameter and proceed to Step 2 of the proposed procedure. Table 10 summarizes the MA estimate and the results of outlier detection. Clearly, the location and the magnitude of the outlier are rightly identified. Out of 400 repetitions, the detection procedure only fails twice. For these two repetitions, the testing statistics λ_A of (2.8) are 3.468 and 3.476, respectively, which are only slightly less than the critical value used, 3.5. However, some cases were identified as IO. This is primarily due to the relatively small π weights of model (4.1b), that is, $\pi(B) = 1 - .4B - .16B^2 - .064B^3 - \cdots$. These small π weights make the testing statistics λ_I and λ_A close to each other occasionally [see (2.3) and (2.4)]. Finally, after

Table 10. Summary of Outlier Detection for Model (4.1)

Model		Outlier		
AR	MA	Location	Frequency	Average Magnitude
.7559	.3952	50	AO (347)	5.9603 (.0014)
(.0155)	(.0115)		IO (51)	6.2625 (.1361)

NOTE: Numbers in parentheses are standard errors or frequencies.

the outlier was identified and removed, the ESACF table was recalculated, with results very close to those in Table 6.

5. CONCLUSION

An iterative identification method was proposed in this article to specify tentative models for time series in the presence of outliers. The method was illustrated by analyzing the U.K. spirits data. Although the proposed method cannot reveal what caused the series to behave inconsistently, it can often pinpoint those observations that deserve special treatments. For the spirits data, it not only identified two additive outliers but also led to the discovery of an intervention. Consequently, a substantial (75%) reduction in residual mean square was obtained.

Finally, besides the parametric approach adopted in this article, there are many other methods for handling outliers in time series. For example, Martin (1981) and Martin et al. (1983) considered some robust alternatives. At this moment, no research has been done to compare these methods thoroughly. On the outset, each approach appears to have its own advantages and disadvantages. Further comparisons in both theory and applications are in order.

[*Received February 1984. Revised July 1985.*]

REFERENCES

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

Akaike, H. (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, AC-19, 716–723.

Anderson, T. W., and Walker, A. M. (1964), "On the Asymptotic Distribution of the Autocorrelations of a Sample From a Linear Stochastic Process," *Annals of Mathematical Statistics*, 35, 1296–1303.

Bartlett, M. S. (1946), "On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series," *Journal of the Royal Statistical Society*, Ser. B, 8 (Suppl.), 27–41.

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

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

Chang, I. (1982), "Outliers in Time Series," unpublished Ph.D. dissertation, University of Wisconsin, Dept. of Statistics.

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

Denby, L., and Martin, R. D. (1979), "Robust Estimation of the First Order Autoregressive Parameter," *Journal of the American Statistical Association*, 74, 140–146.

Durbin, J. (1959), "Efficient Estimation of Parameters in Moving Average Models," *Biometrika*, 46, 306–316.

Durbin, J., and Watson, G. S. (1951), "Exact Tests of Serial Correlation Using Non-Circular Statistics," *Annals of Mathematical Statistics*, 22, 446–451.

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

Fuller, W. A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley.

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

Hannan, E. J., and Rissanen, J. (1982), "Recursive Estimation of Mixed Autoregressive Moving Average Order," *Biometrika*, 69, 81–94.

Hillmer, S. C., Bell, W. R., and Tiao, G. C. (1983), "Modeling Considerations in the Seasonal Adjustment of Economic Time Series," in *Applied Time Series Analysis of Economic Data*, ed. A. Zellner, Washington, DC: U.S. Bureau of the Census, pp. 74–100.

Kendall, M. G., and Stuart, A. (1961), *The Advanced Theory of Statistics* (Vol. 3), New York: Hafner Press.

Liu, L. M., and Hudak, G. (1983), *The SCA System*, DeKalb, IL: Scientific Computing Associates.

Martin, R. D. (1980), "Robust Estimation of Autoregressive Models," in *Directions in Time Series*, eds. D. R. Brillinger and G. C. Tiao, Hayward, CA: Institute of Mathematical Statistics, pp. 228–254.

——— (1981), "Robust Methods for Time Series," in *Applied Time Series Analysis II*, ed. D. F. Findley, New York: John Wiley, pp. 683–759.

Martin, R. D., Samarov, A., and Vandaele, W. (1983), "Robust Methods for ARIMA Models," in *Applied Time Series Analysis of Economic Data*, ed. A. Zellner, Washington, DC: U.S. Bureau of the Census, pp. 153–169.

Miller, R. B. (1980), "Comments on Robust Estimation on Autoregressive Models by Martin," in *Directions in Time Series*, eds. D. R. Brillinger and G. C. Tiao, Hayward, CA: Institute of Mathematical Statistics, pp. 255–262.

Prest, A. R. (1949), "Some Experiments in Demand Analysis," *Review of Economics and Statistics*, 31, 33–49.

Tsay, R. S. (1984), "Time Series Model Specification in the Presence of Outliers," Technical Report 312, Carnegie-Mellon University, Dept. of Statistics.

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