Joint Estimation of Model Parameters and Outlier Effects in Time Series

CHUNG CHEN and LON-MU LIU*

Time series data are often subject to uncontrolled or unexpected interventions, from which various types of outlying observations are produced. Outliers in time series, depending on their nature, may have a moderate to significant impact on the effectiveness of the standard methodology for time series analysis with respect to model identification, estimation, and forecasting. In this article we use an iterative outlier detection and adjustment procedure to obtain joint estimates of model parameters and outlier effects. Four types of outliers are considered, and the issues of spurious and masking effects are discussed. The major differences between this procedure and those proposed in earlier literature include (a) the types and effects of outliers are obtained based on less contaminated estimates of model parameters, (b) the outlier effects are estimated simultaneously using multiple regression, and (c) the model parameters and the outlier effects are estimated jointly. The sampling behavior of the test statistics for cases of small to large sample sizes is investigated through a simulation study. The performance of the procedure is examined over a representative set of outlier cases. We find that the proposed procedure performs well in terms of detecting outliers and obtaining unbiased parameter estimates. An example is used to illustrate the application of the proposed procedure. It is demonstrated that this procedure performs effectively in avoiding spurious outliers and masking effects. The model parameter estimates obtained from the proposed procedure are typically very close to those estimated by the exact maximum likelihood method using an intervention model to incorporate the outliers.

KEY WORDS: Estimation accuracy; Intervention analysis; Iterative estimation; Masking effect; Outlier detection; Power of detection; Spurious outlier.

Most time series data are observational in nature. In addition to possible gross errors, time series data are often subject to the influence of some nonrepetitive events; for example, implementation of a new regulation, major changes in political or economic policy, or occurrence of a disaster. Consequently, discordant observations and various types of structural changes occur frequently in time series data. Whereas the usual time series model is designed to grasp the homogeneous memory pattern of a time series, the presence of outlying observations or structural changes raises the question of efficiency and adequacy in fitting general autoregressive moving average (ARMA) models to time series data (see, for example, Abraham and Box 1979; Chen and Tiao 1990; Guttman and Tiao 1978; Hillmer 1984; Hillmer, Bell, and Tiao 1983; Ledolter 1988; and Tsay 1986).

A common approach to deal with outliers in a time series is to identify the locations and the types of outliers and then use intervention models discussed in Box and Tiao (1975) to accommodate the outlier effects. This approach requires iterations between stages of outlier detection and estimation of an intervention model. Procedures considered by Chang, Tiao, and Chen (1988), Hillmer et al. (1983), and Tsay (1988) are quite effective in detecting the locations and estimating the effects of large isolated outliers; however, a few issues still remain:

a. The presence of outliers may result in an inappropriate model.
b. Even if the model is appropriately specified, outliers in a time series may still produce bias in parameter estimates and hence may affect the efficiency of outlier detection. A typical difficulty found in this approach was that both the types and locations of outliers may change at different iterations of model estimation.
c. Some outliers may not be identified due to a masking effect.

Tsay (1985) attempted to resolve the issue of model identification in the presence of outliers; in this article we focus on solving the problems of b and c.

This study's main goal is to design a procedure that is less vulnerable to the spurious and masking effects during outlier detection and is able to jointly estimate the model parameters and outlier effects. In Section 1 four types of outliers are defined and issues of detecting outliers and adjusting their effects are investigated. To achieve outlier detection and parameter estimation jointly, the procedure proposed in this article consists of three major stages of iterations. The motivation and the detailed steps of the proposed procedure are discussed in Section 2. The behavior of the test statistics and the performance of the proposed procedure are investigated in Section 3. An illustrative example is given in Section 4, and conclusions are presented in Section 5.

1. OUTLIERS IN TIME SERIES

The proposed procedure may be applied to general seasonal and nonseasonal ARMA processes. To simplify the presentation, only the nonseasonal case without a constant term will be used to illustrate the procedure. Let $\{Y_t\}$ be a time series following a general ARMA process,

$$Y_t = \frac{\theta(B)}{\alpha(B)\phi(B)} a_t, \qquad t = 1, \ldots, n, \tag{1}$$

where n is the number of observations for the series; $\theta(B)$, $\phi(B)$, and $\alpha(B)$ are polynomials of B; all roots of $\theta(B)$ and $\phi(B)$ are outside the unit circle; and all roots of $\alpha(B)$ are on

* Chung Chen is Associate Professor, Department of Quantitative Methods, School of Management, Syracuse University, NY 13244-2130. Lon-Mu Liu is Professor, Department of Information and Decision Sciences, College of Business Administration, University of Illinois at Chicago, IL 60680. This work was supported in part by Summer research grants of Syracuse University and by Scientific Computing Associates. The authors thank George C. Tiao, Tuey S. Tsay, and Gregory B. Hudak for their helpful comments and suggestions, along with the associate editor and two referees, whose comments helped greatly in revising this article.

Journal of the American Statistical Association
March 1993, Vol. 88, No. 421, Theory and Methods

the unit circle. The model in (1) may include a constant term when the nonstationary operator $\alpha(B)$ is contained on the left side of the model equation. To describe a time series subject to the influence of a nonrepetitive event, the following model is considered:

$$Y_t^* = Y_t + \omega \frac{A(B)}{G(B)H(B)} I_t(t_1), \qquad (2)$$

where Y_t follows a general ARMA process described in (1), $I_t(t_1) = 1$ if $t = t_1$, and $I_t(t_1) = 0$ otherwise. Here $I_t(t_1)$ is an indicator function for the occurrence of the outlier impact, t_1 is the possibly unknown location of the outlier, and ω and $A(B)/\{G(B)H(B)\}$ denote the magnitude and the dynamic pattern of the outlier effect. The model for a time series that allows for multiple outliers is presented in (19). If the location and the dynamic pattern of an event is known, then model (2) is the intervention model studied by Box and Tiao (1975). In this article we consider the estimation problem when both the location and the dynamic pattern are not known a priori. The approach is to classify an outlier impact into four types by imposing a special structure on $A(B)/\{G(B)H(B)\}$. The types include an innovational outlier (IO), an additive outlier (AO), a level shift (LS), and a temporary change (TC). Their definitions are given below:

$$\text{IO:} \quad \frac{A(B)}{G(B)H(B)} = \frac{\theta(B)}{\alpha(B)\phi(B)}, \qquad (3)$$

$$\text{AO:} \quad \frac{A(B)}{G(B)H(B)} = 1, \qquad (4)$$

$$\text{TC:} \quad \frac{A(B)}{G(B)H(B)} = \frac{1}{(1-\delta B)}, \qquad (5)$$

and

$$\text{LS:} \quad \frac{A(B)}{G(B)H(B)} = \frac{1}{(1-B)}. \qquad (6)$$

For a more detailed discussion on the nature and the motivation of these outlier types, see Chen and Tiao (1990), Fox (1972), Hillmer et al. (1983), and Tsay (1988). The four outliers represent various types of simple outlier effects; more complicated responses usually can be approximated by combinations of the four types. In principle, the proposed procedure can handle any other specific form of outlier responses.

1.1 Effect of Outliers on the Observed Series

It is useful to note that, except for the case of an IO, the effects of outliers on the observed series are independent of the model. Also, the AO and LS are two boundary cases of a TC, where $\delta = 0$ and $\delta = 1$. For a TC, the outlier produces an initial effect ω at time t_1, and this effect dies out gradually with time. The parameter δ is designed to model the pace of the dynamic dampening effect. In practice, the value of δ can be specified by the analyst. We recommend that $\delta = .7$ be used to identify a TC. In the case of an AO, the outlier causes an immediate and one-shot effect on the observed series. A LS produces an abrupt and permanent step change in the series.

On a time series, the effect of an IO is more intricate than the effects of other types of outliers. Using the formulation of model (3), we see that when an IO occurs at $t = t_1$, the effect of this outlier on Y_{t_1+k}, for $k \geq 0$, is equal to $\omega\psi_k$, where ω is the initial effect and ψ_k is the kth coefficient of the $\psi(B)$ polynomial where

$$\psi(B) = \{\theta(B)\}/\{\alpha(B)\phi(B)\}$$
$$= (\psi_0 + \psi_1 B + \psi_2 B^2 + \cdots), \qquad \psi_0 = 1.$$

For a stationary series, an IO will produce a temporary effect because the ψ_j's decay to 0 exponentially. The pattern of ψ_j's for a nonstationary series can be quite different. Depending on the model of Y_t, an IO may produce (a) an initial effect at the time of the intervention and a level shift from the second period of the intervention, if Y_t follows an autoregressive integrated moving average (ARIMA)(0, 1, 1) model; (b) an initial effect at the time of intervention, gradually converging to a permanent level shift if Y_t follows an ARIMA(1, 1, 1) model; (c) a seasonal level shift if Y_t follows a pure seasonal ARIMA $(0, 1, 1)_s$ model (e.g., a level shift at every January of each year for monthly series), and (d) an annual trend changes if Y_t follows a multiplicative seasonal ARIMA$(0, 1, 1) \times (0, 1, 1)_s$ model.

1.2 Estimating and Adjusting the Effect of an Outlier

To examine the effects of outliers on the estimated residuals, we assume that the time series parameters are known and the series is observed from $t = -J$ to $t = n$, where J is an integer larger than $p + d + q$, and that $1 \leq t_1 \leq n$ where p, d, and q are orders of the polynomials $\phi(B)$, $\alpha(B)$, and $\theta(B)$. We define the $\pi(B)$ polynomial as

$$\pi(B) = \{\phi(B)\alpha(B)\}/\{\theta(B)\} = 1 - \pi_1 B - \pi_2 B^2 - \cdots,$$

where the π_j weights for j beyond a moderately large J become essentially equal to 0, because the roots of $\theta(B)$ are all outside the unit circle. The estimated residuals $\hat{e}_t$, which may be contaminated with outliers, can be expressed as

$$\hat{e}_t = \pi(B)Y_t^*, \quad \text{for} \quad t = 1, 2, \ldots. \qquad (7)$$

For our four types of outliers, we have

$$\text{IO:} \quad \hat{e}_t = \omega I_t(t_1) + a_t, \qquad (8)$$

$$\text{AO:} \quad \hat{e}_t = \omega\pi(B)I_t(t_1) + a_t, \qquad (9)$$

$$\text{TC:} \quad \hat{e}_t = \omega\{\pi(B)/(1-\delta B)\}I_t(t_1) + a_t, \qquad (10)$$

and

$$\text{LS:} \quad \hat{e}_t = \omega\{\pi(B)/(1-B)\}I_t(t_1) + a_t. \qquad (11)$$

Alternatively, we can rewrite equations (8)–(11) as

$$\hat{e}_t = \omega x_{it} + a_t,$$
$$t = t_1, t_1 + 1, \ldots, n \quad \text{and} \quad i = 1, 2, 3, 4, \qquad (12)$$

where $x_{it} = 0$ for all i and $t < t_1$, $x_{it_1} = 1$ for all i and $k \geq 1$, $x_{1(t_1+k)} = 0$, $x_{2(t_1+k)} = -\pi_k$, $x_{3(t_1+k)} = 1 - \sum_{j=1}^{k} \pi_j$, and $x_{4(t_1+k)} = \delta^k - \sum_{j=1}^{k-1} \delta^{k-j}\pi_j - \pi_k$. Hence the least squares estimate for the effect of a single outlier at $t = t_1$ may be expressed as

$$\hat{\omega}_{\rm IO}(t_1) = \hat{e}_{t_1}$$

$$\hat{\omega}_{\rm AO}(t_1) = \frac{\sum_{t=t_1}^{n} \hat{e}_t x_{2t}}{\sum_{t=t_1}^{n} x_{2t}^2}$$

$$\hat{\omega}_{\rm LS}(t_1) = \frac{\sum_{t=t_1}^{n} \hat{e}_t x_{3t}}{\sum_{t=t_1}^{n} x_{3t}^2}$$

$$\hat{\omega}_{\rm TC}(t_1) = \frac{\sum_{t=t_1}^{n} \hat{e}_t x_{4t}}{\sum_{t=t_1}^{n} x_{4t}^2}. \quad (13)$$

It is important to note that for the last observation (i.e., $t_1 = n$), $\hat{\omega}_{\rm IO}(n) = \hat{\omega}_{\rm AO}(n) = \hat{\omega}_{\rm LS}(n) = \hat{\omega}_{\rm TC}(n) = \hat{e}_n$. As a result, it is impossible to empirically distinguish the type of an outlier occurring at the very end of a series.

As discussed in Chang et al. (1988), a possible approach for detecting outliers is to examine the maximum value of the standardized statistics of the outlier effects

$$\hat{\tau}_{\rm IO}(t_1) = \hat{\omega}_{\rm IO}(t_1)/\hat{\sigma}_a$$

$$\hat{\tau}_{\rm AO}(t_1) = \{\hat{\omega}_{\rm AO}(t_1)/\hat{\sigma}_a\}\left(\sum_{t=t_1}^{n} x_{2t}^2\right)^{1/2}$$

$$\hat{\tau}_{\rm LS}(t_1) = \{\hat{\omega}_{\rm LS}(t_1)/\hat{\sigma}_a\}\left(\sum_{t=t_1}^{n} x_{3t}^2\right)^{1/2}$$

$$\hat{\tau}_{\rm TC}(t_1) = \{\hat{\omega}_{\rm TC}(t_1)/\hat{\sigma}_a\}\left(\sum_{t=t_1}^{n} x_{4t}^2\right)^{1/2}. \quad (14)$$

For a given location, these standardized statistics follow an approximately normal distribution. Knowing the type and the location of an outlier, one can adjust the outlier effects on the observations and the residuals using Equation (3) and Equations (8)–(11). In general, the adjusted observation at t_1, denoted $\tilde{Y}_{t_1}$, can be expressed as a weighted sum of the entire observed series. In the case of IO, it can be shown that the adjusted observation $\tilde{Y}_{t_1}$ is the conditional expectation of Y_{t_1} given the past observations. Under an AO, the adjusted observation is the interpolation based on both the past and the future Y's, but it does not involve Y_{t_1}. This suggests a possible approach to estimating missing values in a time series by treating the missing data as an AO.

1.3 The Issue of Multiple Outliers

When there are multiple outliers, the previously described estimate of ω at time period t_1 may not be an unbiased estimate for the outlier effect at t_1, due to the influence of neighboring outliers. Consider the following special case of two additive outliers in an ARMA model:

$$Y_t^* = \omega_1 I_t(t_1) + \omega_2 I_t(t_2) + \{\theta(B)/\phi(B)\} a_t. \quad (15)$$

Assuming that the series is fitted by an appropriate ARMA model and that $\hat{e}_t$ is the estimated residual, we have

$$\hat{e}_t = \pi(B) Y_t^* = \omega_1 \pi(B) I_t(t_1) + \omega_2 \pi(B) I_t(t_2) + a_t. \quad (16)$$

If we know the location of the outliers, say t_1 and t_2, the effects of outliers at t_1 and t_2 may be estimated jointly as

$$\hat{\omega}_1 = [1 - \alpha_{12}^2/\alpha_{11}\alpha_{22}]^{-1}\{\tilde{\omega}_1 - (\alpha_{21}/\alpha_{11})\tilde{\omega}_2\}$$

$$\hat{\omega}_2 = [1 - \alpha_{12}^2/\alpha_{11}\alpha_{22}]^{-1}\{-(\alpha_{21}/\alpha_{22})\tilde{\omega}_1 + \tilde{\omega}_2\}, \quad (17)$$

where $\pi(B) = (1 - \pi_1 B - \pi_2 B^2 - \cdots) = \phi(B)/\theta(B)$, $\alpha_{kk} = \sum_{i=0}^{n-t_k} \pi_i^2$, $k = 1, 2$, $\pi_0 = 1$, $\alpha_{21} = \alpha_{21} = \pi_{t_2-t_1} + \sum_{i=1}^{n-t_2} \pi_i \pi_{i+t_2-t_1}$ for $t_2 > t_1$, and $(\tilde{\omega}_1, \tilde{\omega}_2)$ are estimates of ω_1 and ω_2 obtained separately assuming that only a single outlier is present, as described in (13). In an iterative outlier detection procedure, as proposed by Chang et al. (1988) and Tsay (1988), one could adjust the effect of ω_1 on the residual $\hat{e}_t$'s and then use the adjusted residuals to estimate ω_2 or vice versa. Let $\tilde{\omega}_{i.j}$ denote the estimate of ω_i after the effect of ω_j has been adjusted. We can derive that

$$\tilde{\omega}_{2.1} = \tilde{\omega}_2 - (\alpha_{12}/\alpha_{22})\tilde{\omega}_1$$

$$\tilde{\omega}_{1.2} = \tilde{\omega}_1 - (\alpha_{12}/\alpha_{11})\tilde{\omega}_2. \quad (18)$$

Consequently, for the two-outlier case, the preceding approach results in the estimates of either $(\tilde{\omega}_1, \tilde{\omega}_{2.1})$ or $(\tilde{\omega}_{1.2}, \tilde{\omega}_2)$. Depending on the structure of the time series and the relative positions of t_1 and t_2, the estimates of (ω_1, ω_2) obtained from a sequentially adjusted procedure of Chang et al. (1988) can be quite different from the results of joint estimation, as illustrated in (17).

From a computational standpoint, the strategy of detecting outliers one by one may be the only feasible approach to dealing with multiple outliers. It seems more appropriate, however, to estimate the outlier effects jointly rather than sequentially. The preceding analysis also indicates that a procedure based solely on iteratively adjusted residuals often may produce biased estimates for adjacent outliers.

1.4 Estimation of Residual Standard Deviation σ_a

To compute the test statistics of outliers as given in (14), one needs to estimate σ_a. The determination of outliers can be sensitive to this estimate. In the presence of outliers, the residuals are contaminated; hence σ_a may be overestimated if the usual sample standard deviation is used. Three methods for obtaining a better estimate of σ_a are considered in this study: (1) the median absolute deviation (MAD) method, (2) the α% trimmed method, and (3) the omit-one method.

The MAD estimate of the residual standard deviation is defined as

$$\hat{\sigma}_a = 1.483 \times \text{median}\{|\hat{e}_t - \tilde{e}|\},$$

where $\tilde{e}$ is the median of the estimated residuals (Andrews et al. 1972, p. 239). To compute the α% trimmed standard deviation, we first remove the α% largest values (according to their asbolute values) and then compute the sample standard deviation based on the trimmed sample. When conducting an outlier test at time point t_1, the omit-one method calculates the usual residual standard deviation with the residual at time t_1 omitted. The MAD and α% trimmed methods require sorting of the residuals. In cases of large sample size, the required computing time can be much greater for these two methods than for the omit-one method. The motivation of these methods is to reduce the possibility of misdetection due to an inflated estimate of the residual standard deviation. Once the locations of outliers are identified and their effects are estimated, σ_a can be estimated based on the sample standard deviation of the adjusted residuals.

2. A JOINT ESTIMATION PROCEDURE IN THE PRESENCE OF MULTIPLE OUTLIERS

Suppose that the series Y_t is subject to m interventions at time points $t_1, t_2, \ldots, t_m$, resulting in various types of outliers. The model for Y_t^* can be expressed as

$$Y_t^* = \sum_{j=1}^{m} \omega_j L_j(B) I_t(t_j) + \frac{\theta(B)}{\phi(B)\alpha(B)} a_t, \qquad (19)$$

where $L_j(B) = \theta(B)/\{\phi(B)\alpha(B)\}$ for an IO, $L_j(B) = 1$ for an AO, $L_j(B) = 1/(1 - B)$ for an LS, and $L_j(B) = 1/(1 - \delta B)$ for a TC at $t = t_j$. Without distinguishing the notations of the estimated and the true parameters, the residuals $\{\hat{e}_t\}$ by fitting an ARMA model to Y_t^* may be expressed as

$$\hat{e}_t = \sum_{j=1}^{m} \omega_j \pi(B) L_j(B) I_t(t_j) + a_t \qquad (20)$$

when the underlying model is correctly specified but outlier effects are not taken into consideration. Equations (19) and (20) are the foundation of the proposed procedure. If the effects of outliers and their locations are available, then we can adjust the outlier effects based on Equation (19) and thereafter estimate the model parameters. On the other hand, when the model parameters are known, we can identify outliers and estimate their effects based on Equation (20). It is difficult, if not impossible, to achieve our stated goals in a single step. Thus we develop an iterative procedure that consists of three major stages. In Stage I all the potential outliers, t_j and $L_j(B)$, are identified, based on preliminary model parameter estimates. In Stage II joint estimates of the model parameters and outlier effects are obtained using the accumulated outlier information of Stage I. In Stage III outliers t_j and $L_j(B)$ are identified and their effects estimated again, based on the less-contaminated estimates of model parameters obtained in Stage II.

2.1 The Detection and Estimation Procedure

In this section we provide a detailed summary of the proposed iterative procedure for the joint estimation of model parameters and outlier effects.

Stage I: Initial Parameter Estimation and Outlier Detection

I.1. Compute the maximum likelihood estimates of the model parameters based on the original or the adjusted series and obtain the residuals. For the very first iteration, the original series is used to initiate the procedure; after the first iteration, the adjusted series is used.

Inner Loop of Outlier Detection for Fixed Model Parameter Estimates

I.2. For $t = 1, \ldots, n$, compute $\hat{\tau}_{IO}(t)$, $\hat{\tau}_{AO}(t)$, $\hat{\tau}_{LS}(t)$, and $\hat{\tau}_{TC}(t)$ in (14) using the residuals obtained from I.1, and let $\eta_t = \max\{|\hat{\tau}_{IO}(t)|, |\hat{\tau}_{AO}(t)|, |\hat{\tau}_{LS}(t)|, |\hat{\tau}_{TC}(t)|\}$. If $\max_t \eta_t = |\hat{\tau}_{tp}(t_1)| > C$, where C is predetermined critical value, then there is a possibility of a type tp outlier at t_1; tp may be IO, AO, LS, or TC.

I.3. If there is no outlier found, then go to step I.4. Otherwise, remove the effect of this outlier from the residuals and the observations according to its type, then go back to step I.2 to check if an additional outlier can be found.

I.4. If no outliers are found in the very first iteration of this inner loop, then stop—the observed series is free from outlier effects. If outliers are found in the inner loop under the given parameter estimates, then go to step I.1 to revise the parameter estimates. If the total number of outliers in all of the inner loops is greater than 0 and no additional outliers are detected in the current inner loop, then go to step II.1.

Stage II: Joint Estimation of Outlier Effects and Model Parameters

II.1. Suppose that m time points $t_1, t_2, \ldots, t_m$ are identified as possible outliers of various types. The outlier effects ω_j's can be estimated jointly using the multiple regression model described in (20), where $\{\hat{e}_t\}$ is regarded as the output variable and $\{L_j(B)I_t(t_j)\}$ are the input variables.

II.2. Compute the $\hat{\tau}$ statistics of the estimated ω_j's, where $\hat{\tau}_j = \hat{\omega}_j/\text{std}(\hat{\omega}_j)$, $j = 1, \ldots, m$. If $\min_j |\hat{\tau}_j| = \hat{\tau}_\nu \le C$, where C is the same critical value used in step I.2, then delete the outlier at time point t_ν from the set of the identified outliers and go to step II.1 with the remaining $m - 1$ outliers. Otherwise, go to step II.3.

II.3. Obtain the adjusted series by removing the outlier effects, using the most recent estimates of ω_j's at step II.1. In other words, remove only the outlier effects that are significant based on the iterations of steps II.1 and II.2.

II.4. Compute the maximum likelihood estimates of the model parameters based on the adjusted series obtained at step II.3. If the relative change of the residual standard error from the previous estimate is greater than ε, go to step II.1 for further iterations; otherwise, go to step III.1. The tolerance ε is a predetermined constant chosen by the user as a means to control the accuracy of parameter estimates. An appropriate tolerance value, for example, could be .001.

Stage III: Detection of Outliers Based on the Final Parameter Estimates

III.1. Compute the residuals by filtering the original series based on the parameter estimates obtained at step II.4.

III.2. Use the residuals obtained at step III.1 and iterate through Stages I and II with the modifications that (a) the parameter estimates used in the inner loop of Stage I are fixed to those obtained at step II.4 and (b) steps II.3 and II.4 are omitted in Stage II. The estimated $\hat{\omega}_j$'s of the last iteration at step II.1 are the final estimates of the effects of the detected outliers.

2.2 Some Remarks on the Proposed Procedure

For a given set of parameter estimates, the inner-loop iteration (steps I.2 and I.3) detects outliers one by one in a descending order of magnitude in terms of the $\hat{\tau}$ statistics. Whenever an outlier is detected, its effects on the residuals and on the observations are adjusted accordingly at step I.3. The residual series is then used as step I.2 to detect another outlier. This is essentially the procedure proposed by Chang et al. (1988). The main reason for using this procedure, which detects only a single outlier in an inner-loop iteration, is to simplify the computation involved in joint detection of multiple outliers. But as discussed in Section 1.3, such a procedure may suffer from masking effects, because a later iteration of outlier detection is based on the adjusted residuals of the previous iterations. In addition, the parameter estimates may have bias due to the presence of the outliers. Hence is is important to adjust both the observed and the residual series based on the type of outlier detected. These series may be used in the latter iterations of parameter estimation and outlier detection. The iterations of estimation (step I.1) and detection (steps I.2 and I.3) in Stage I are designed to reduce the possibility of masking effects. It is helpful to point out that this practice in Stage I is essentially the procedure M proposed by Tsay (1988).

A potential problem for the preceding iterative detection procedure is that the identified outliers are not evaluated on the same basis in terms of $\hat{\sigma}_a$. For instance, the first identified outlier is detected based on the assumption that no outlier is present. Once we decide the presence of the first outlier, the residuals and the observations are adjusted accordingly and the process of detecting the second outlier begins. As a result, computations of the estimates of outlier effects and the $\hat{\tau}$ statistics given in (13) and (14) at different iterations are not based on the same residual series. Consequently, the joint effects of these outliers and their $\hat{\tau}$ statistics are not clearly exhibited. In Stage II, a procedure akin to the backward elimination procedure in multiple regression is used to jointly evaluate the outlier effects and to remove any spurious outliers. The multiple regression approach has been considered by P. Burman (1989), who modified the outlier detection programs developed by the U.S. Bureau of the Census.

Based on the notation of "robustness" adopted by Box and Andersen (1955), a method is considered to be "robust" when the inferential results are sensitive to the main concerns but are insensitive to variations of nuisance factors. The proposed procedure is robust in the sense that the model parameter estimates are sensitive to the overall memory pattern of a time series but are insensitive to occasional outliers. In the next section we conduct extensive simulation studies to investigate the power and properties of the proposed joint detection and estimation procedure under the assumption of Gaussian noise. We have also investigated the performance of the proposed procedure under certain non-Gaussian noise, such as noise with exponential distributions. In such situations, we find that the proposed procedure is effective in determining extreme values in a time series, but it cannot distinguish such extreme values as outliers or regular observations associated with the inherent nature of the distribution. Further study on outlier detection and adjustment is needed for such distributional assumptions.

3. A STUDY OF THE PERFORMANCE OF THE JOINT ESTIMATION PROCEDURE

The proposed procedure is iterative in nature and is designed to accomplish outlier detection and model estimation jointly. In Section 3.1 we first investigate the behavior of the test statistics used in the procedure. Such information is useful in providing guidelines for the selection of critical values in the detection stage. In Section 3.2 we study the performance of the proposed procedure. There are two aspects to the evaluation of procedure performance: (1) the power of outlier detection and (2) the accuracy of the parameter estimates. For the power study, we report the relative frequency of correct outlier detection and the average frequency of misidentified outliers. The estimation performance is examined, based on a representative sample of stationary and nonstationary models as well as on a selective set of outlier alternatives. The criteria of evaluation are the sample mean and sample root mean square errors (RMSE's) of the parameter estimates. Other studies on the power of related outlier detection procedures were discussed in Chang et al. (1988) and Tsay (1988).

3.1 The Sampling Behavior of the Test Statistics

The detection procedure essentially is based on the maxima of the test statistics considered in (14). The sampling behaviors of the maxima of these test statistics are associated with (a) the sample size, (b) the type of outlier, (c) the pattern of π weights of the fitted model, and (d) the estimates of the residual standard deviation. The simulation study in this section is designed to investigate the sampling properties of the maxima of the outlier test statistics. Table 1 lists the models considered in this simulation study, which represent a broad spectrum of π weight patterns.

Model 1 is the mean model; its result is considered a ref-

Table 1. Underlying Models for the Calculation of Test Statistics

Model	Calculation
1	$Y_t = 10 + a_t$
2	$Y_t = 10 + (1 - 0.8B)a_t$
3	$Y_t = 10 + (1 + 0.5B)a_t$
4	$(1 + .5B)Y_t = 10 + a_t$
5	$(1 - .8B)Y_t = 10 + a_t$
6	$(1 - B + .24B^2)Y_t = 10 + a_t$
7	$(1 - .9B + B^2)Y_t = a_t$
8	$(1 - .5B)\nabla Y_t = a_t$
9	$\nabla Y_t = (1 - .8B)a_t$
10	$\nabla Y_t = (1 + .5B)a_t$
11	$\nabla\nabla_{12}Y_t = (1 - .5B)(1 - .8B^{12})a_t$

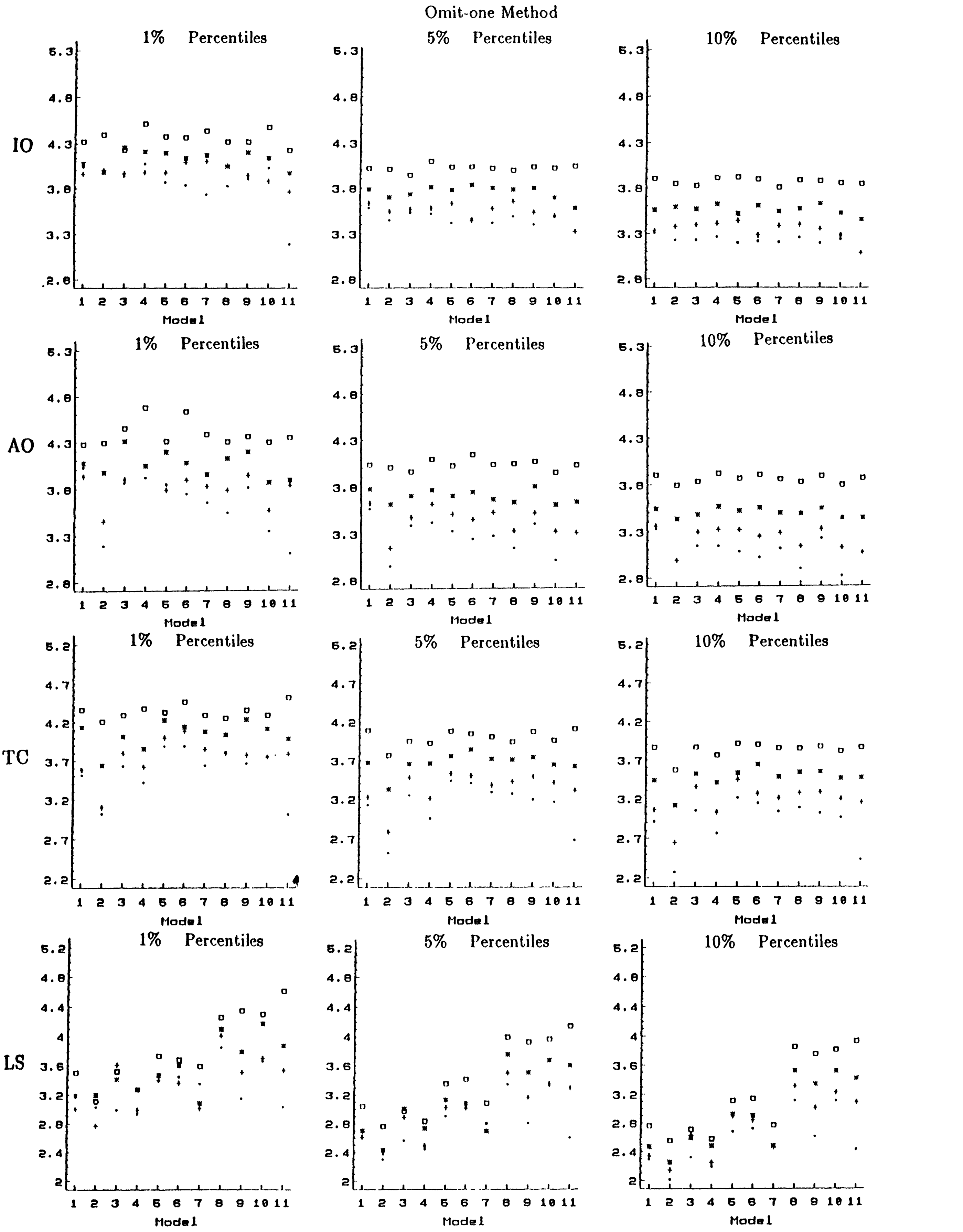

Figure 1. Estimated Percentiles of Test Statistics, Omit-One Method. Here, ● represents n = 50, + represents n = 100, ∗ represents n = 250, and □ represents n = 1,000.

erence base for the simulation study. For each model, four cases of sample size are examined: n = 50, 100, 250, and 1,000. The random errors, a_t's, follow iid normal distribution with mean 0 and variance 1. For each model and each sample size, 500 series are generated. The test statistics for IO, AO, TC, and LS outliers in each series are calculated separately based on (14). In this study the focus is on examining the sampling behavior of $\eta_{tp} = \max_{t_1} |\hat{\tau}_{tp}(t_1)|$, where tp = IO, AO, TC, and LS. In particular, we wish to estimate the percentiles of those statistics at the 1%, 5%, and 10% levels when no outlier is present in the series.

In the simulation study three methods for obtaining a robust estimate of the residual standard deviation, $\hat{\sigma}_a$, are considered. The estimate is computed using (a) the omit-one, (b) the MAD, and (c) the 5% trimmed methods discussed previously. Due to space limitations, in Figure 1 we only provide a graphical summary of the estimates of those percentiles at three significance levels (1%, 5%, and 10%) for IO, AO, TC, and LS test statistics based on the omit-one method. (Numerical results for the MAD and the 5% trimmed methods may be obtained from the authors.) Of the three methods of computing the residual standard deviation, the 5% trimmed method produces the highest values of the test statistics, whereas the omit-one method and the MAD method give similar results. There are 12 plots in Figure 1, consisting of the combination of four types of test statistics and three significance levels. In each plot percentiles for each of the 11 models and four sample sizes are displayed. Except for a few cases of 1% percentiles, the estimated percentiles are an increasing function of sample size n. The IO test statistic is quite homogeneous with respect to various time series models. This is not surprising, because the IO test statistic is simply the maximum of the standardized residuals. In general, the sample size is an important factor affecting the behavior of the test statistics. Under the omit-one method, the estimated 5% percentiles for IO range from 3.4 to 4.0 for sample sizes n = 50 to n = 1,000. The AO test statistic has slightly higher variations, depending on model pattern. For the estimates of 5% percentiles, they range from 3.0 to 4.0 for sample sizes n = 50 to n = 1,000. There is no clear model effect for the TC statistic either, and the estimated 5% percentiles range from 2.5 to 4.1. The LS test statistic has two distinct characteristics. First, its estimated percentiles are relatively smaller (in absolute value) than those of other outlier types. Second, for models that do not include a regular differencing operator, it tends to be substantially lower than the others. The estimates of 5% percentiles using the omit-one method range from 2.3 to 3.4 in this case.

Based on these simulation results, the following guidelines for choosing the critical value C are recommended. For a series with moderate length (say between 100 and 200 observations), a critical value C = 3.0 seems to be appropriate. For a series of shorter length, a critical value between 2.5 and 2.9 is recommended. We may consider a critical value greater than 3.0 for series of longer length (e.g., over 200 observations). In practice, it is recommended that more than one critical value be used in the analysis, to allow examination of the sensitivity of the results to the choice of the critical values. Other considerations in the choice of critical values are discussed in the next subsection.

3.2 Performance of the Proposed Procedure

The proposed procedure is designed to handle multiple outliers of various types in a time series. We design simulations to study the performance of the procedure applied to cases characterized by a combination of the following factors: (a) four outlier types; (b) four underlying models consisting of an AR(1), an MA(1), an IMA(0, 1, 1), and a multiplicative seasonal model IMA(0, 1, 0) × (0, 1, 1)$_4$; (c) three outlier sizes; (d) a single outlier and two adjacent outliers, and (e) outliers occurring at the beginning, in the middle, and at the end of the series. Due to space limitations, we report only the results of the power of correct detection of a single outlier and the accuracy of model parameter estimates. (Results of other cases may be obtained from the authors.)

Table 2 lists the cases considered in this study. Cases 1–9 are combinations of three underlying models and three outlier sizes. For these cases, series are generated to contain one of the four outlier types discussed in Section 1. Cases 10–18 cover situations of two neighboring outliers occurring at the beginning, in the middle, and at the end of the series using three different models. For series generated in these cases, the first outlier may be one of the four types, and the second outlier is fixed as an AO. The standard deviation of the noise process for each model is set to 1. The true value of model parameters in all cases is set to .6. For a given underlying model and a specification of the sizes, locations, and types of the outliers, 500 series of length 100 are generated using the SCA Statistical System (Liu, Hudak, Box, Muller, and Tiao 1986). The procedure using the omit-one method for estimating the residual standard deviation is applied to each of the 500 series using six different critical values: C = 2.25, 2.5, 2.75, 3.0, 3.25, and 3.5.

The relative frequency of correct detection and the frequency of Type I errors are reported in Table 3. Because the

Table 2. List of Cases in the Performance Study

Case	Model	Location & size of AO	Case	Model	Location & size of AO
1	AR(1)	$t_1 = 40$ $\omega_1 = 3$	10	AR(1)	$t_1 = 40$ $\omega_1 = 3$ $t_2 = 41$ $\omega_2 = 4$
2	MA(1)	$t_1 = 40$ $\omega_1 = 3$	11	AR(1)	$t_1 = 10$ $\omega_1 = 4$ $t_2 = 11$ $\omega_2 = -3$
3	IMA(0, 1, 1)	$t_1 = 40$ $\omega_1 = 3$	12	AR(1)	$t_1 = 99$ $\omega_1 = 4$ $t_2 = 100$ $\omega_2 = -3$
4	AR(1)	$t_1 = 40$ $\omega_1 = 4$	13	IMA(0, 1, 1)	$t_1 = 40$ $\omega_1 = 3$ $t_2 = 41$ $\omega_2 = 4$
5	MA(1)	$t_1 = 40$ $\omega_1 = 4$	14	IMA(0, 1, 1)	$t_1 = 10$ $\omega_1 = 4$ $t_2 = 15$ $\omega_2 = 3$
6	IMA(0, 1, 1)	$t_1 = 40$ $\omega_1 = 4$	15	IMA(0, 1, 1)	$t_1 = 96$ $\omega_1 = 4$ $t_2 = 98$ $\omega_2 = 3$
7	AR(1)	$t_1 = 40$ $\omega_1 = 5$	16	Seasonal IMA	$t_1 = 40$ $\omega_1 = 4$ $t_2 = 44$ $\omega_2 = 4$
8	MA(1)	$t_1 = 40$ $\omega_1 = 5$	17	Seasonal IMA	$t_1 = 10$ $\omega_1 = 3$ $t_2 = 12$ $\omega_2 = 4$
9	IMA(0, 1, 1)	$t_1 = 40$ $\omega_1 = 5$	18	Seasonal IMA	$t_1 = 96$ $\omega_1 = 3$ $t_2 = 97$ $\omega_2 = 4$

Table 3. Summary of the Detection Performance

Case	*C* =	2.25	2.50	2.75	3.00	3.25	3.50	Case	*C* =	2.25	2.50	2.75	3.00	3.25	3.50
				AO								*TC*			
1	*P*	.89	.83	.75	.64	.54	.40	1	*P*	.81	.70	.59	.49	.37	.28
2	*P*	.50	.60	.62	.54	.46	.37	2	*P*	.64	.74	.78	.79	.79	.78
3	*P*	.88	.80	.70	.62	.52	.40	3	*P*	.87	.80	.70	.61	.50	.41
4	*P*	.99	.97	.96	.93	.88	.81	4	*P*	.96	.93	.89	.83	.74	.64
5	*P*	.62	.81	.86	.88	.86	.80	5	*P*	.68	.76	.81	.83	.83	.83
6	*P*	.99	.98	.95	.92	.87	.80	6	*P*	.99	.98	.95	.92	.85	.79
7	*P*	1.00	1.00	1.00	.99	.99	.98	7	*P*	1.00	1.00	.99	.98	.96	.91
8	*P*	.69	.87	.94	.97	.97	.96	8	*P*	.71	.77	.80	.80	.80	.80
9	*P*	1.00	.99	.99	.99	.98	.97	9	*P*	1.00	1.00	.99	.99	.98	.97
1	*E*	4.2	1.8	.7	.3	.1	.1	1	*E*	5.1	2.2	.9	.4	.1	.1
2	*E*	2.1	1.2	.5	.2	.1	.0	2	*E*	.8	.4	.2	.1	.0	.0
3	*E*	4.8	1.9	.8	.3	.1	.0	3	*E*	4.0	1.8	.8	.3	.1	.0
4	*E*	4.3	1.8	.7	.3	.1	.0	4	*E*	5.4	2.4	1.0	.5	.2	.1
5	*E*	2.1	1.3	.6	.3	.1	.0	5	*E*	.9	.4	.2	.1	.0	.0
6	*E*	5.0	1.9	.8	.3	.1	.0	6	*E*	4.1	1.8	.8	.4	.2	.1
7	*E*	4.3	1.8	.8	.3	.1	.1	7	*E*	5.3	2.3	1.0	.5	.2	.1
8	*E*	2.3	1.3	.6	.3	.1	.0	8	*E*	.8	.4	.1	.1	.0	.0
9	*E*	5.0	2.1	.8	.3	.1	.1	9	*E*	4.2	1.8	.8	.4	.1	.1
				IO								*LS*			
1	*P*	.83	.72	.60	.49	.39	.29	1	*P*	.69	.55	.36	.22	.10	.04
2	*P*	.83	.71	.60	.49	.37	.28	2	*P*	.54	.49	.51	.56	.52	.51
3	*P*	.85	.76	.66	.55	.46	.32	3	*P*	.70	.78	.75	.66	.55	.47
4	*P*	.98	.95	.89	.83	.76	.66	4	*P*	.93	.87	.77	.62	.48	.33
5	*P*	.97	.93	.90	.84	.77	.65	5	*P*	.69	.66	.65	.63	.67	.69
6	*P*	.97	.94	.91	.85	.74	.66	6	*P*	.85	.93	.93	.92	.87	.83
7	*P*	1.00	.99	.99	.97	.95	.93	7	*P*	.99	.97	.94	.89	.82	.71
8	*P*	1.00	1.00	.98	.97	.95	.91	8	*P*	.76	.75	.75	.74	.69	.72
9	*P*	1.00	.99	.98	.96	.94	.91	9	*P*	.90	.98	.99	.99	.98	.98
1	*E*	5.3	2.2	.8	.4	.1	.0	1	*E*	.7	.4	.1	.1	.0	.0
2	*E*	4.9	2.1	.8	.3	.1	.0	2	*E*	.7	.6	.5	.5	.3	.2
3	*E*	4.9	2.0	.8	.3	.1	.0	3	*E*	2.8	1.4	.5	.2	.1	.0
4	*E*	5.0	2.1	.9	.4	.2	.0	4	*E*	.8	.4	.2	.1	.0	.0
5	*E*	4.9	2.2	1.0	.4	.1	.1	5	*E*	.7	.7	.6	.4	.4	.3
6	*E*	4.9	2.1	.9	.4	.1	.1	6	*E*	3.0	1.5	.6	.3	.1	.0
7	*E*	5.2	2.3	1.0	.4	.2	.1	7	*E*	1.0	.6	.3	.2	.1	.0
8	*E*	5.2	2.2	.9	.4	.2	.1	8	*E*	.7	.6	.6	.5	.4	.4
9	*E*	4.9	2.0	.9	.4	.2	.1	9	*E*	3.0	1.5	.7	.3	.1	.0

NOTE: *P* refers to the relative frequency of correct detection, and *E* refers to the average number of misidentified outliers in a series of length 100.

effects of neighboring outliers may be approximated by a combination of various consecutive outliers, the power study for such situations is more complicated. Here we report only the results of the power study for cases with a single outlier. The rows labelled "P" in this table summarize the relative frequency of correct detection, defined as a correct identification of both type and location of an outlier. The relative frequency of correct detection can be interpreted as the power of the procedure in terms of outlier detection. For most cases, the power is a decreasing function of the critical value. There are exceptions, however, when the critical value $C = 2.25$ is used. Examining the detailed detection record (not reported here), we found that when the critical value is too low, there is a higher frequency to misidentify the location of an outlier by one time period. This is due to the high correlation between neighboring test statistics. For critical values 2.75 and 3.0, the power of the procedure for detecting outliers of size 3 standard deviation (σ_a) ranges between 50% and 60%, and that for detecting outliers of sizes 4 and 5 standard deviation ranges between 85% and 99%.

The rows labeled "E" in Table 3 report the average number of misidentified outliers (i.e., the number of observations in a series that are identified as outliers while they are not outliers) in a series of length 100. These results indicate the frequency of Type I errors in the detection procedure. This is different from the Type I errors of the test statistics studied in the previous section. In the latter simulation, because the procedure allows for checking the significance of the estimated outlier effects jointly, it generates a lower frequency of Type I errors. We find that the frequency of misidentification is also a decreasing function of the specified critical value. For critical values 2.75 and 3.0, the average number of misidentified outliers in a series of length 100 is less than 1. This finding further validates the guidelines for critical value selection provided in Section 3.1.

To investigate the impact of outlier adjustment on parameter estimation, we compute the estimates using the standard ARIMA model and the intervention model, which incorporates the information of outlier type and location. Table 4 summarizes the estimation results of this simulation study, including the sample mean and sample RMSE of the model parameter estimates and the residual standard deviation es-

Table 4. Summary of the Estimation Performance

AO	ITV C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA	TC	ITV C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA
			Case #1								Case #1				
PHI	.577	.636	.603	.586	.575	.568	.562	.540	PHI	.569	.489	.531	.554	.563	.568	.569	.576
RMSE1	.090	.104	.094	.092	.094	.095	.096	.109	RMSE1	.093	.203	.147	.117	.103	.096	.094	.087
SIGMA	.964	.739	.856	.923	.966	.990	1.014	1.097	SIGMA	.964	.700	.833	.918	.963	.994	1.013	1.071
RMSE2	.137	.305	.210	.170	.151	.148	.151	.179	RMSE2	.138	.349	.245	.184	.164	.153	.152	.162
			Case #2								Case #2				
THETA	.639	.966	.825	.702	.652	.631	.613	.556	THETA	.637	.790	.698	.653	.644	.634	.631	.304
RMSE1	.108	.498	.378	.238	.171	.154	.134	.102	RMSE1	.104	.382	.235	.151	.142	.126	.131	.304
SIGMA	.965	.732	.815	.912	.966	.995	1.016	1.112	SIGMA	.963	.911	.932	.956	.965	.972	.975	1.359
RMSE2	.148	.355	.269	.201	.177	.172	.172	.200	RMSE2	.145	.205	.164	.153	.148	.146	.147	.401
			Case #3								Case #3				
THETA	.618	.537	.578	.598	.607	.612	.616	.624	THETA	.622	.649	.634	.619	.612	.604	.601	.580
RMSE1	.095	.144	.112	.100	.094	.092	.093	.088	RMSE1	.095	.124	.112	.103	.096	.092	.092	.092
SIGMA	.972	.715	.848	.923	.964	.992	1.012	1.116	SIGMA	.965	.736	.841	.909	.957	.988	1.009	1.108
RMSE2	.139	.332	.224	.178	.157	.155	.157	.195	RMSE2	.144	.313	.231	.187	.163	.161	.163	.189
			Case #4								Case #4				
PHI	.573	.635	.602	.585	.576	.571	.566	.514	PHI	.578	.497	.540	.561	.568	.574	.576	.589
RMSE1	.091	.098	.090	.089	.091	.092	.096	.124	RMSE1	.088	.193	.138	.111	.100	.093	.091	.082
SIGMA	.953	.729	.843	.905	.939	.959	.975	1.175	SIGMA	.975	.687	.820	.903	.950	.985	1.009	1.155
RMSE2	.143	.317	.223	.174	.158	.153	.150	.239	RMSE2	.146	.357	.250	.196	.172	.164	.171	.229
			Case #5								Case #5				
THETA	.639	.963	.802	.722	.678	.659	.650	.512	THETA	.643	.755	.678	.651	.642	.637	.636	.206
RMSE1	.112	.506	.351	.259	.197	.175	.174	.126	RMSE1	.106	.382	.253	.185	.145	.131	.125	.399
SIGMA	.965	.739	.819	.891	.933	.955	.972	1.211	SIGMA	.956	1.028	.969	.970	.977	.981	.982	1.560
RMSE2	.142	.333	.252	.202	.172	.164	.169	.269	RMSE2	.140	.966	.164	.157	.149	.147	.146	.597
			Case #6								Case #6				
THETA	.616	.525	.573	.591	.600	.606	.609	.638	THETA	.618	.639	.628	.617	.610	.607	.603	.563
RMSE1	.094	.153	.116	.100	.095	.092	.092	.091	RMSE1	.100	.127	.120	.108	.104	.101	.102	.103
SIGMA	.975	.707	.849	.913	.951	.973	.989	1.208	SIGMA	.967	.729	.836	.900	.939	.964	.982	1.195
RMSE2	.143	.342	.230	.185	.159	.158	.157	.270	RMSE2	.150	.324	.239	.192	.170	.165	.162	.261
			Case #7								Case #7				
PHI	.581	.643	.614	.596	.587	.582	.580	.489	PHI	.570	.488	.523	.548	.558	.563	.567	.599
RMSE1	.086	.099	.090	.085	.085	.086	.087	.145	RMSE1	.091	.200	.154	.124	.108	.101	.095	.082
SIGMA	.947	.724	.834	.894	.928	.945	.953	1.289	SIGMA	.957	.677	.805	.879	.922	.946	.963	1.224
RMSE2	.144	.321	.233	.185	.164	.153	.147	.339	RMSE2	.141	.358	.253	.193	.165	.150	.149	.280
			Case #8								Case #8				
THETA	.626	.925	.768	.695	.659	.646	.639	.464	THETA	.637	.692	.646	.631	.630	.627	.625	.117
RMSE1	.100	.481	.317	.224	.165	.148	.143	.162	RMSE1	.111	.365	.234	.175	.146	.146	.141	.488
SIGMA	.961	.756	.844	.903	.931	.946	.956	1.317	SIGMA	.958	1.048	1.002	1.003	1.002	1.009	1.011	1.769
RMSE2	.148	.332	.256	.214	.187	.163	.159	.368	RMSE2	.149	.290	.218	.192	.182	.182	.182	.802
			Case #9								Case #9				
THETA	.620	.529	.572	.593	.601	.605	.606	.655	THETA	.624	.645	.632	.625	.618	.615	.612	.546
RMSE1	.096	.146	.115	.104	.095	.094	.090	.098	RMSE1	.098	.125	.113	.103	.097	.097	.095	.114
SIGMA	.969	.703	.827	.903	.939	.956	.966	1.302	SIGMA	.964	.727	.838	.895	.928	.951	.962	1.311
RMSE2	.146	.343	.241	.182	.161	.154	.152	.355	RMSE2	.140	.321	.231	.185	.164	.151	.149	.361
			Case #10								Case #10				
PHI	.579	.533	.536	.553	.560	.563	.562	.562	PHI	.571	.543	.546	.556	.566	.568	.573	.576
RMSE1	.087	.153	.155	.128	.112	.101	.096	.087	RMSE1	.088	.135	.137	.122	.104	.097	.091	.080
SIGMA	.970	.559	.688	.859	.951	1.014	1.061	1.180	SIGMA	.954	.557	.676	.820	.910	.957	.995	1.309
RMSE2	.137	.455	.361	.241	.192	.177	.181	.234	RMSE2	.148	.459	.374	.261	.196	.179	.173	.361
			Case #11								Case #11				
PHI	.574	.452	.477	.518	.537	.548	.552	.416	PHI	.566	.499	.522	.541	.552	.557	.559	.504
RMSE1	.092	.211	.195	.157	.131	.117	.112	.212	RMSE1	.091	.171	.148	.124	.111	.105	.103	.133
SIGMA	.959	.556	.670	.808	.904	.960	.995	1.415	SIGMA	.953	.547	.672	.805	.904	.957	.991	1.231
RMSE2	.151	.460	.380	.273	.213	.181	.168	.459	RMSE2	.142	.468	.376	.268	.193	.165	.162	.288
			Case #12								Case #12				
PHI	.569	.462	.491	.526	.546	.559	.566	.439	PHI	.577	.517	.539	.560	.568	.572	.574	.516
RMSE1	.089	.198	.179	.143	.120	.107	.100	.191	RMSE1	.089	.162	.149	.117	.104	.098	.095	.123
SIGMA	.957	.551	.669	.816	.904	.950	.988	1.393	SIGMA	.967	.543	.684	.829	.908	.944	.968	1.205
RMSE2	.144	.463	.375	.262	.199	.167	.155	.437	RMSE2	.145	.473	.365	.250	.194	.176	.164	.270
			Case #13								Case #13				
THETA	.624	.711	.710	.673	.638	.626	.619	.603	THETA	.617	.679	.692	.655	.627	.616	.610	.566
RMSE1	.098	.183	.189	.155	.115	.104	.097	.095	RMSE1	.095	.165	.176	.140	.114	.104	.097	.104
SIGMA	.968	.500	.614	.788	.912	.971	1.021	1.251	SIGMA	.957	.508	.595	.748	.865	.933	.983	1.400
RMSE2	.147	.510	.431	.303	.204	.185	.183	.310	RMSE2	.143	.503	.456	.316	.230	.187	.166	.444

(continued)

Table 4. (continued)

AO	ITV	C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA	TC	ITV	C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA
Case #14									*Case #14*								
THETA	.623	.757	.760	.704	.667	.649	.642	.666	THETA	.623	.701	.706	.670	.637	.626	.620	.585
RMSE1	.100	.207	.225	.189	.157	.132	.118	.105	RMSE1	.096	.166	.182	.156	.120	.108	.103	.085
SIGMA	.966	.501	.575	.740	.864	.938	1.003	1.298	SIGMA	.960	.519	.599	.763	.880	.946	.996	1.289
RMSE2	.140	.511	.455	.323	.233	.193	.188	.344	RMSE2	.143	.500	.437	.307	.209	.182	.183	.342
Case #15									*Case #15*								
THETA	.612	.737	.732	.679	.643	.629	.627	.646	THETA	.619	.691	.701	.656	.633	.620	.616	.579
RMSE1	.093	.193	.200	.161	.119	.101	.096	.093	RMSE1	.095	.162	.186	.144	.114	.105	.100	.099
SIGMA	.960	.497	.592	.756	.887	.968	1.025	1.260	SIGMA	.974	.522	.604	.783	.895	.967	1.017	1.314
RMSE2	.150	.512	.446	.320	.229	.199	.194	.314	RMSE2	.149	.492	.431	.286	.211	.180	.189	.365
Case #16									*Case #16*								
THETA	.615	.651	.644	.630	.619	.619	.621	.652	THETA	.625	.613	.616	.615	.607	.603	.599	.573
RMSE1	.101	.115	.115	.106	.101	.100	.104	.095	RMSE1	.101	.107	.109	.105	.101	.099	.098	.094
SIGMA	.971	.531	.634	.801	.915	1.000	1.059	1.348	SIGMA	.967	.543	.639	.799	.909	.982	1.032	1.462
RMSE2	.148	.482	.409	.286	.224	.193	.206	.397	RMSE2	.153	.470	.405	.278	.209	.189	.202	.504
Case #17									*Case #17*								
THETA	.623	.660	.650	.635	.627	.624	.623	.676	THETA	.622	.590	.601	.604	.604	.604	.605	.506
RMSE1	.095	.128	.122	.115	.106	.102	.102	.115	RMSE1	.102	.132	.131	.116	.114	.109	.108	.159
SIGMA	.963	.525	.615	.755	.863	.931	.971	1.619	SIGMA	.963	.576	.659	.785	.859	.909	.942	1.840
RMSE2	.145	.489	.422	.305	.222	.182	.181	.653	RMSE2	.145	.467	.404	.284	.216	.184	.172	.874
Case #18									*Case #18*								
THETA	.623	.617	.622	.622	.622	.622	.624	.579	THETA	.612	.622	.622	.615	.612	.611	.611	.576
RMSE1	.098	.127	.121	.112	.105	.099	.098	.132	RMSE1	.097	.133	.130	.126	.118	.109	.104	.140
SIGMA	.966	.524	.638	.790	.899	.962	1.003	1.398	SIGMA	.963	.518	.612	.761	.863	.925	.969	1.570
RMSE2	.148	.489	.405	.285	.208	.177	.168	.442	RMSE2	.149	.494	.423	.303	.224	.190	.175	.608

IO	ITV	C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA	LS	ITV	C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA
Case #1									*Case #1*								
PHI	.577	.565	.569	.570	.572	.575	.575	.574	PHI	.566	.654	.692	.732	.763	.788	.799	.808
RMSE1	.081	.111	.098	.091	.086	.082	.082	.079	RMSE1	.095	.129	.147	.171	.189	.202	.208	.213
SIGMA	.962	.696	.835	.921	.966	.996	1.016	1.071	SIGMA	.951	.969	1.019	1.080	1.124	1.161	1.180	1.196
RMSE2	.148	.345	.242	.187	.168	.161	.160	.169	RMSE2	.151	.166	.183	.213	.234	.251	.257	.263
Case #2									*Case #2*								
THETA	.631	.644	.634	.631	.631	.631	.630	.625	THETA	.657	.296	.295	.315	.322	.289	.264	−.248
RMSE1	.105	.162	.134	.121	.108	.105	.104	.096	RMSE1	.117	.429	.441	.443	.406	.443	.460	.850
SIGMA	.958	.697	.828	.907	.955	.986	1.003	1.064	SIGMA	.956	1.750	1.735	1.610	1.504	1.521	1.553	3.139
RMSE2	.147	.348	.245	.191	.173	.164	.165	.169	RMSE2	.147	1.373	1.254	.993	.708	.818	.775	2.165
Case #3									*Case #3*								
THETA	.625	.614	.608	.612	.612	.612	.612	.606	THETA	.632	.820	.707	.645	.618	.602	.588	.534
RMSE1	.101	.118	.100	.101	.093	.091	.090	.085	RMSE1	.107	.297	.201	.144	.122	.117	.109	.113
SIGMA	.958	.704	.834	.910	.955	.979	1.008	1.083	SIGMA	.954	.947	.884	.921	.953	.978	.997	1.114
RMSE2	.145	.335	.230	.184	.164	.156	.159	.176	RMSE2	.142	.698	.302	.195	.160	.158	.158	.191
Case #4									*Case #4*								
PHI	.568	.561	.562	.562	.564	.564	.565	.569	PHI	.560	.676	.686	.698	.723	.746	.776	.862
RMSE1	.086	.109	.098	.094	.089	.089	.089	.084	RMSE1	.096	.145	.151	.167	.184	.199	.218	.264
SIGMA	.964	.712	.837	.909	.950	.976	.999	1.139	SIGMA	.964	.975	1.010	1.046	1.094	1.133	1.183	1.319
RMSE2	.142	.335	.235	.189	.162	.160	.161	.208	RMSE2	.151	.176	.185	.214	.248	.279	.312	.372
Case #5									*Case #5*								
THETA	.624	.638	.629	.626	.622	.622	.622	.617	THETA	.660	.299	.278	.268	.264	.246	.217	−.360
RMSE1	.099	.166	.136	.115	.105	.104	.102	.098	RMSE1	.122	.426	.440	.467	.511	.497	.483	.962
SIGMA	.962	.695	.819	.888	.937	.968	.991	1.136	SIGMA	.956	1.823	1.834	1.829	1.818	1.645	1.677	4.130
RMSE2	.149	.348	.253	.199	.169	.161	.166	.212	RMSE2	.137	1.599	1.427	1.287	1.238	.869	.892	3.155
Case #6									*Case #6*								
THETA	.620	.608	.604	.604	.604	.605	.604	.603	THETA	.627	.795	.697	.646	.623	.607	.598	.485
RMSE1	.097	.110	.104	.102	.095	.092	.091	.086	RMSE1	.103	.274	.186	.128	.101	.101	.102	.143
SIGMA	.952	.699	.820	.890	.932	.963	.985	1.141	SIGMA	.963	.929	.885	.921	.940	.964	.979	1.216
RMSE2	.144	.342	.244	.189	.167	.159	.163	.210	RMSE2	.142	.717	.360	.234	.160	.155	.158	.277
Case #7									*Case #7*								
PHI	.571	.561	.565	.567	.567	.568	.569	.569	PHI	.567	.710	.708	.695	.692	.697	.707	.895
RMSE1	.088	.108	.101	.093	.090	.089	.086	.083	RMSE1	.094	.166	.175	.168	.166	.174	.183	.296
SIGMA	.957	.694	.814	.891	.929	.952	.964	1.226	SIGMA	.964	.968	1.000	1.015	1.036	1.055	1.087	1.419
RMSE2	.151	.348	.253	.197	.173	.163	.165	.289	RMSE2	.144	.163	.162	.167	.196	.217	.250	.466

(continued)

Table 4. (continued)

IO	ITV	C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA	LS	ITV	C = 2.25	2.50	2.75	3.00	3.25	3.50	ARIMA
				Case #8									Case #8				
THETA	.636	.643	.633	.635	.634	.633	.634	.622	THETA	.662	.250	.279	.257	.238	.237	.214	−.438
RMSE1	.101	.145	.113	.119	.110	.110	.110	.086	RMSE1	.132	.486	.467	.463	.494	.555	.552	1.040
SIGMA	.965	.691	.823	.893	.932	.955	.970	1.228	SIGMA	.967	2.016	1.924	1.940	1.970	1.939	1.837	5.277
RMSE2	.145	.350	.248	.195	.168	.158	.159	.279	RMSE2	.148	1.960	1.817	1.497	1.562	1.602	1.151	4.308
				Case #9									Case #9				
THETA	.614	.602	.601	.602	.601	.601	.600	.600	THETA	.642	.804	.706	.662	.641	.632	.626	.447
RMSE1	.099	.121	.102	.100	.093	.092	.092	.090	RMSE1	.111	.284	.184	.132	.110	.103	.096	.172
SIGMA	.957	.701	.823	.888	.926	.944	.960	1.236	SIGMA	.962	.953	.882	.911	.943	.965	.972	1.347
RMSE2	.154	.341	.243	.195	.174	.167	.168	.300	RMSE2	.142	.866	.356	.196	.178	.270	.270	.389
				Case #10									Case #10				
PHI	.579	.533	.536	.553	.560	.563	.562	.562	PHI	.551	.740	.745	.770	.796	.810	.808	.775
RMSE1	.087	.153	.155	.128	.112	.101	.096	.087	RMSE1	.101	.171	.178	.197	.212	.217	.214	.183
SIGMA	.970	.559	.688	.859	.951	1.014	1.061	1.180	SIGMA	.944	.696	.815	.945	1.046	1.112	1.153	1.462
RMSE2	.137	.455	.361	.241	.192	.177	.181	.234	RMSE2	.149	.385	.333	.241	.216	.226	.249	.502
				Case #11									Case #11				
PHI	.574	.452	.477	.518	.537	.548	.552	.416	PHI	.564	.782	.770	.757	.746	.738	.733	.692
RMSE1	.092	.211	.195	.157	.131	.117	.112	.212	RMSE1	.093	.196	.186	.171	.159	.152	.147	.111
SIGMA	.959	.556	.670	.808	.904	.960	.995	1.415	SIGMA	.962	.626	.775	.939	1.046	1.103	1.138	1.332
RMSE2	.151	.460	.380	.273	.213	.181	.168	.459	RMSE2	.142	.400	.315	.222	.201	.213	.235	.379
				Case #12									Case #12				
PHI	.569	.462	.491	.526	.546	.559	.566	.439	PHI	.569	.525	.535	.552	.563	.565	.564	.541
RMSE1	.089	.198	.179	.143	.120	.107	.100	.191	RMSE1	.092	.149	.144	.117	.104	.097	.098	.106
SIGMA	.957	.551	.669	.816	.904	.950	.988	1.393	SIGMA	.967	.540	.671	.812	.900	.945	.971	1.164
RMSE2	.144	.463	.375	.262	.199	.167	.155	.437	RMSE2	.146	.477	.381	.264	.203	.177	.166	.238
				Case #13									Case #13				
THETA	.618	.701	.715	.669	.634	.619	.615	.589	THETA	.637	.658	.677	.650	.628	.617	.613	.542
RMSE1	.100	.173	.197	.155	.125	.102	.100	.098	RMSE1	.115	.155	.186	.151	.120	.103	.101	.119
SIGMA	.959	.510	.612	.781	.884	.953	.990	1.285	SIGMA	.958	.518	.609	.771	.879	.949	.985	1.426
RMSE2	.153	.501	.481	.315	.225	.188	.176	.339	RMSE2	.146	.494	.436	.305	.212	.171	.159	.471
				Case #14									Case #14				
THETA	.618	.716	.726	.679	.642	.627	.622	.617	THETA	.632	.638	.666	.644	.627	.614	.607	.498
RMSE1	.101	.171	.201	.158	.131	.109	.103	.083	RMSE1	.109	.135	.171	.144	.121	.105	.108	.128
SIGMA	.960	.507	.590	.746	.865	.938	.985	1.267	SIGMA	.969	.537	.627	.780	.886	.952	.996	1.358
RMSE2	.139	.501	.448	.314	.225	.182	.182	.316	RMSE2	.146	.477	.428	.290	.211	.181	.188	.404
				Case #15									Case #15				
THETA	.621	.714	.723	.679	.641	.627	.626	.599	THETA	.633	.629	.659	.645	.628	.621	.618	.471
RMSE1	.098	.183	.200	.159	.119	.103	.099	.085	RMSE1	.103	.137	.162	.139	.119	.107	.106	.156
SIGMA	.960	.514	.595	.766	.889	.956	.998	1.265	SIGMA	.953	.524	.612	.768	.884	.940	.978	1.364
RMSE2	.147	.497	.436	.304	.209	.184	.178	.323	RMSE2	.149	.488	.424	.298	.211	.179	.178	.407
				Case #16									Case #16				
THETA	.628	.652	.644	.632	.624	.619	.616	.652	THETA	.627	.593	.600	.597	.594	.587	.578	.511
RMSE1	.105	.119	.114	.109	.101	.095	.092	.098	RMSE1	.108	.114	.116	.110	.110	.109	.115	.130
SIGMA	.963	.523	.624	.778	.884	.942	.978	1.524	SIGMA	.961	.547	.629	.778	.893	.960	1.015	1.566
RMSE2	.150	.492	.416	.292	.210	.181	.167	.565	RMSE2	.145	.468	.406	.292	.209	.184	.193	.606
				Case #17									Case #17				
THETA	.620	.643	.637	.623	.616	.610	.609	.638	THETA	.631	.590	.606	.613	.614	.612	.610	.486
RMSE1	.103	.119	.115	.102	.097	.094	.093	.099	RMSE1	.100	.123	.121	.110	.104	.102	.097	.172
SIGMA	.967	.524	.623	.774	.885	.954	.988	1.531	SIGMA	.969	.563	.660	.784	.877	.933	.964	1.888
RMSE2	.149	.492	.423	.306	.218	.181	.171	.569	RMSE2	.147	.470	.398	.288	.205	.175	.168	.919
				Case #18									Case #18				
THETA	.628	.636	.634	.623	.619	.616	.615	.600	THETA	.617	.611	.618	.617	.613	.611	.612	.544
RMSE1	.100	.119	.117	.106	.102	.093	.092	.095	RMSE1	.098	.120	.120	.108	.100	.091	.089	.133
SIGMA	.958	.517	.624	.792	.902	.962	1.004	1.285	SIGMA	.973	.522	.626	.776	.882	.950	.991	1.559
RMSE2	.144	.495	.414	.278	.200	.178	.171	.333	RMSE2	.152	.490	.413	.298	.225	.187	.178	.599

ITV: intervention model fitting; ARIMA: traditional ARIMA fitting; RMSE1: RMSE of the estimated model parameters; RMSE2: RMSE of the estimated residual standard deviations.

timates over the 500 series. For each case and each type of outliers, there are eight sets of statistics, six of them associated with six different critical values, one with the intervention model (labeled as ITV), and the other with the usual ARIMA model without outlier adjustment (labelled ARIMA). The results associated with the intervention model may be regarded as the best estimate in the sense that the correct information of outliers has been used. The usual ARIMA estimates are obtained without allowing for any outliers in the series. These estimates correspond to applying the proposed detection procedure with a critical value of infinity for outlier detection. The estimates associated with finite critical values can be interpreted as outcomes of the procedure allowing for various degrees of model perturbation in the form of model (19). Examining the results presented in Table 4, we find that the AO, TC, and LS outliers may cause substantial bias in the estimation of model parameters, but the IO effect seems to be less serious on the model parameter estimates. It is not surprising to find that estimates of residual standard deviation are sensitive to all types of outliers, and the proposed procedure successfully obtains a more unbiased estimate of the residual standard deviation. It is found that AO and TC outliers occurring at the beginning and at the end of the series (cases 11, 12, 14, 15, 17, and 18) produce higher RMSE in the ARIMA estimates than do outliers in the middle of the series. This finding suggests that outliers at the beginning and at the end of the series are more important in terms of their impact. When the critical values are too small (e.g., $C = 2.25$ and $C = 2.5$ in this study), the procedure tends to overadjust the series and produce unsatisfactory results. When the critical values are gradually increased to a proper level, we observe that the estimates become more stable and more consistent with those obtained from the intervention models. But if the critical value is too large, fewer or no outliers will be detected and the results will be very close to those of usual ARIMA estimates. We find that in most cases there is a range of critical values with which the procedure produces comparable improvements on the parameter estimates and RMSE's with respect to the intervention model. In practice it is useful to use the proposed procedure on data subsets with a range of critical values. This may help reveal structural changes other than those discussed in this article. To obtain better understanding of the outlier effects, it is very informative to plot the adjusted series as well as the actual observed series.

The cases of an LS outlier in an MA(1) (cases 2, 5, and 8) require some explanation. The empirical distribution of the parameter estimates in this case is bimodal, with one mode near $-.4$ and the other mode near .5. The former is associated with situations in which the procedure fails to detect the LS outlier, and the latter corresponds to situations in which the LS outlier is successfully detected and adjusted. This is the reason that the sample mean of the parameter estimates is biased toward 0. These cases also indicate that the procedure may be less effective when the initial fitted model departs substantially from the true model.

4. AN ILLUSTRATIVE EXAMPLE

Here we consider the analysis of the variety store series discussed in Hillmer et al. (1983). In this example we demonstrate (a) the detailed steps of the new procedure, (b) the improvement provided by the new procedure, and (c) the comparison between estimates obtained from the intervention model incorporating the outliers and those from the proposed joint estimation procedure.

The data analyzed here are the time series of the log-transformed monthly retail sales of variety stores after the adjustment for trading day and holiday effects. The series begins in January 1967 and ends in September 1979. The plot of the series was given in Hillmer et al. (1983). A strong seasonal pattern and a level drop during 1976 can be found in the time series plot. Hillmer et al. (1983) used this series to illustrate the application of an iterative outlier detection procedure developed in Chang (1982). The following ARIMA model is found to be appropriate for the observed series:

$$\nabla\nabla_{12}Y_t = \frac{(1 - \theta_{12}B^{12})}{(1 - \phi_1 B - \phi_2 B^2)}\, a_t. \tag{21}$$

To contrast the results between the procedure described in this article and that used by Hillmer et al. (1983), only the AO and IO types are considered in the first part of the analysis. Table 5 summarizes the results using the 5% trimmed method to estimate the residual standard deviation. The results obtained using the other two methods (the omit-one and the MAD methods) are quite similar and are not reported here. The critical value $C = 3.0$ is used to detect outliers.

Table 5 is organized in three panels to provide the main results from each stage of the iterative procedure. The top panel summarizes the results of estimation and detection from iteration 1 to iteration 3 in Stage I. The procedure in this stage detects a total of six outliers. The middle panel covers Stage II, at which outlier effects are jointly estimated and the insignificant ones removed. In this case, outliers detected at $t = 103$ and $t = 73$ are not significant for the critical value $C = 3.0$ and are hence removed. The final estimates, reported in the row labelled II.4, are obtained based on the series adjusted for the effects of the four outliers. The bottom panel reports on Stage III, during which the detection procedure is run again without reestimating the model parameters. Nine outliers are detected in the intermediate steps. However, after joint estimation of the outlier effects, only six outliers are considered significant, and the final results are reported at the last step of III.2. Similar results are obtained using the MAD and the omit-one methods. The only major difference is the outcome in the intermediate steps of outlier detection. The 5% trimmed method tends to identify more outliers in the intermediate steps, but the step involving joint estimation of outlier effects successfully removes the spurious outliers.

Applying the procedure considered in Hillmer et al. (1983), nine outliers are identified during six iterations of model estimation. Their outlier detection results are closely compatible with those obtained in the intermediate iterations of stage I and III in the new procedure, but their final results are different from those obtained from the new procedure—particularly the estimate for θ_{12} ($\hat{\theta}_{12} = .89$ in Hillmer et al. versus $\hat{\theta}_{12} = .62$ using the new procedure).

To overcome potential misidentification of outliers, we now consider all four types of outliers, IO, AO, LS, and TC. The joint estimates of model parameters and outlier effects using the new procedure are listed below (numbers in parentheses are the t values of the estimates):

$$\hat{\theta}_{12} = .7128, \quad \hat{\phi}_1 = -.6871, \quad \hat{\phi}_2 = -.4617, \quad \hat{\sigma}_a = .02404.$$
$$(12.84) \qquad (-9.08) \qquad (-6.11)$$

Outlier	*Estimate*	*t Value*	*Type*
$t = 45$	.094	5.19	TC
$t = 96$	$-.083$	-4.36	AO
$t = 112$	$-.176$	-10.20	LS

These estimates are obtained using the 5% trimmed method for estimating the residual standard deviation. When the

Table 5. Parameter Estimation: The 5% Trimmed Method With Outlier Types AO and IO

III.1

I.1 (Iteration 1)

θ_{12}	ϕ_1	ϕ_2	σ_a
.8407	−.3969	−.2673	.02501
(17.12)	(−4.83)	(−3.26)	

Inner Loop (I.2 and I.3)

Time	*Estimate*	*t Value*	*Type*
112	−.15	−5.98	IO
96	−.08	−3.83	AO
113	−.09	−3.60	IO
45	.08	3.39	IO
103	−.07	−3.30	AO

I.1 (Iteration 2)

.7040	−.5616	−.3361	.024653
(12.28)	(−6.97)	(−4.18)	

Inner Loop (I.2 and I.3)

Time	*Estimate*	*t Value*	*Type*
73	.07	2.67	IO

I.1 (Iteration 3)

.6864	−.5409	−.3161
(11.82)	(−6.65)	(−3.90)

Inner Loop (I.2 and I.3)

No outlier detected

II.1–II.3 (Joint Estimation of Outlier Effect)

Time	*Estimate*	*t Value*	*Type*
45	.091	3.30	IO
96	−.079	−3.63	AO
112	−.150	−5.46	IO
113	−.121	−4.39	IO

II.4 (Final Estimation)

θ_{12}	ϕ_1	ϕ_2	σ_a
.6202	−.6062	−.3828	.02656
(9.65)	(−7.67)	(−4.86)	

III.2, Step I (Outlier Detection Based on Final Parameter Estimates)

Iteration 1 (Inner Loop)

Time	*Estimate*	*t Value*	*Type*
112	−.15	−5.58	IO
113	−.13	−5.08	IO
96	−.08	−3.83	AO
45	.09	3.54	IO
124	.08	3.19	IO
114	−.08	−3.16	IO

Iteration 2 (Inner Loop)

Time	*Estimate*	*t Value*	*Type*
103	−.05	−2.83	AO
73	.07	2.67	IO
136	.07	2.65	IO

Iteration 3 (Inner Loop)

No outlier detected

III.2, Step II (Final Results of Outlier Detection)

Time	*Estimate*	*t Value*	*Type*
45	.094	3.54	IO
96	−.079	−3.83	AO
112	−.148	−5.58	IO
113	−.135	−5.08	IO
114	−.084	−3.16	IO
124	.085	3.19	IO

omit-one and the MAD methods are used, the results are similar except for some minor differences on the parameter estimates. It is useful to note that by allowing for a more complete set of outliers, we obtain fewer but more meaningful outliers. In addition, the estimated standard deviation is reduced to $\hat{\sigma}_a = .02404$ (compared with $\hat{\sigma}_a = .02667$ when only AO and IO are considered).

Based on the preceding results, we can explicitly incorporate the outlier effects in model (21) and estimate the following intervention model:

$$\nabla\nabla_{12}Y_t = \frac{\omega_1}{1-.7B}\nabla\nabla_{12}I_t(45) + \omega_2\nabla\nabla_{12}I_t(96) + \frac{\omega_3}{1-B}\nabla\nabla_{12}I_t(112) + \frac{(1-\theta_{12}B^{12})}{(1-\phi_1 B-\phi_2 B^2)}a_t. \quad (22)$$

Using the exact maximum likelihood method, the following estimates of model (22) are obtained

$$\hat{\theta}_{12} = .7123, \quad \hat{\phi}_1 = -.6877,$$
$$(12.84) \quad (-9.05)$$
$$\hat{\phi}_2 = -.4622, \quad \hat{\sigma}_a = .02352,$$
$$(-6.10)$$
$$\hat{\omega}_1 = .0956, \quad \hat{\omega}_2 = -.0837, \quad \hat{\omega}_3 = 7.166.$$
$$(5.35) \quad (-4.44) \quad (-10.42)$$

It is found that the estimates obtained from the new joint estimation procedure are very close to those obtained from the intervention model with outlier information incorporated. If we consider the results of the intervention model to be the accurate ones, then the maximum difference in parameter estimates produced by the new procedure is less than 1% in this case.

5. SUMMARY AND CONCLUSION

In this article we study the issue of multiple outliers (AO, IO, TC, and LS) in time series modeling. We discuss potential masking and spurious effects using the traditional detection procedure and develop an iterative procedure for joint estimation of model parameters and outlier effects. Applying the proposed procedure, we obtain estimates of the model parameters with the consideration of a potential departure from the usual ARMA models. In this sense, the procedure provides a tool to bridge the gap between the reality and the traditional ARMA models in time series analysis.

Sampling behavior of the associated test statistics is investigated through a simulation study. It is found that both the length of the series and the method of estimating the residual standard deviation have impact on the choice of the critical value. The memory pattern of the underlying model does not seem to have major influence on the behavior of the test statistics except for the level shift test statistic. In the performance study we find that the proposed procedure is quite effective in outlier detection and parameter estimation when proper critical values are used.

An example is used to illustrate the application of the procedure. Based on this example and other studies, the procedure seems to be effective in reducing spurious and masking effects. The estimates of the model parameters using the new procedure are essentially identical to those obtained by explicitly incorporating the outliers in the model. Further applications of this joint estimation procedure to intervention analysis and transfer function modeling can be found in Liu and Chen (1991). It is shown that outlier adjustment is an indispensable part of intervention analysis.

[*Received September 1990. Revised November 1991.*]

REFERENCES

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

Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H., and Tukey, J. W. (1972), *Robust Estimates of Locations: Survey and Advances,* Princeton, NJ: Princeton University Press.

Box, G. E. P., and Andersen, S. L. (1955), "Permutation Theory in the Derivation of Robust Criteria and the Study of Departure from Assumptions," *Journal of the Royal Statistical Society,* Ser. B, 17, 1–34.

Box, G. E. P., and Jenkins, G. M. (1970), *Time Series Analysis: Forecasting and Control,* San Francisco: Holden-Day.

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

Burman, J. P. (1989), "Outlier Detection in Time Series," American Statistical Association Proceedings of Business and Economic Statistics Section, 402–407.

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

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

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.

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

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.

Hillmer, S. C. (1984), "Monitoring and Adjusting Forecasts in the Presence of Additive Outliers," *Journal of Forecasting,* 3, 205–215.

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,* Washington, DC: U.S. Bureau of the Census, pp. 74–100.

Ledolter, J. (1988), "The Effect of Additive Outliers on the Forecasts From ARIMA Models," *The International Journal of Forecasting,* 5, 231–240.

Liu, L-M., and Chen, C. (1991), "Recent Developments of Time Series Analysis in Environmental Impact Studies," *Journal of Environment Science and Health,* 26, 1217–1252.

Liu, L. M., Hudak, G., Box, G. E. P., Muller, M. E., and Tiao, G. C. (1986), *The SCA Statistical System: Reference Manual for Forecasting and Time Series Analysis,* DeKalb, IL: Scientific Computing Associates.

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

——— (1988), "Outliers, Level Shifts, and Variance Changes in Time Series," *Journal of Forecasting,* 7, 1–20.