TECHNOMETRICS, MAY 1988, VOL. 30, NO. 2

Estimation of Time Series Parameters in the Presence of Outliers

Ih Chang
Pharmaceutical Division
Pennwalt Corporation
Rochester, NY 14623

George C. Tiao
Graduate School of Business
University of Chicago
Chicago, IL 60637

Chung Chen
Department of Management and Systems
Washington State University
Pullman, WA 99164-4726

Outliers in time series can be regarded as being generated by dynamic intervention models at unknown time points. Two special cases, innovational outlier (IO) and additive outlier (AO), are studied in this article. The likelihood ratio criteria for testing the existence of outliers of both types, and the criteria for distinguishing between them are derived. An iterative procedure is proposed for detecting IO and AO in practice and for estimating the time series parameters in autoregressive-integrated-moving-average models in the presence of outliers. The powers of the procedure in detecting outliers are investigated by simulation experiments. The performance of the proposed procedure for estimating the autoregressive coefficient of a simple AR(1) model compares favorably with robust estimation procedures proposed in the literature. Two real examples are presented.

KEY WORDS: Additive outlier; Innovational outlier; ARIMA model; Intervention; Robust estimate.

1. INTRODUCTION

Unexpected extraordinary observations that look discordant from most observations in a data set are often encountered in various kinds of data analysis. Time series analysis is no exception. In addition to possible gross errors, time series observations are often subject to the influence of nonrepetitive exogenous interventions—for example, strikes, outbreak of wars, sudden changes in the market structure of a commodity, unexpected changes of certain conditions in a physical system, and so forth—and as a result some observations become outliers.

Let x_t be a stochastic process following an autoregressive-integrated-moving average (ARIMA) model or order p, d, and q; that is,

$$\phi(B)\alpha(B)x_t = \theta(B)a_t, \tag{1.1}$$

where B is the backshift operator such that $Bx_t = x_{t-1}$; $\phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$ and $\theta(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$ are two polynomials in B with zeros all outside the unit circle; $\alpha(B) = (1 - B)^{d_1}(1 - B^s)^{d_2}$; $d = d_1 + sd_2$; and $\{a_t\}$ are $\overset{\text{iid}}{\approx} N(0, \sigma_a^2)$. The effect of an exogenous intervention on a time series can often be represented by a dynamic model (see Box and Tiao 1975) as follows:

$$z_t = \frac{\omega(B)}{\beta(B)}\zeta_t^{(T)} + x_t, \tag{1.2}$$

where x_t follows the model (1.1),

$$\zeta_t^{(T)} = 1 \quad \text{for } t = T$$
$$= 0 \quad \text{otherwise}$$

signifies the time of occurrence of the intervention and $\omega(B) = (\omega_0 - \omega_1 B - \cdots - \omega_s B^s)$ and $\beta(B) = (\beta_0 - \beta_1 B - \cdots - \beta_r B^r)$ are two polynomials in B. The ratio $\omega(B)/\beta(B)$ describes the dynamic behavior of the intervention.

Guttman and Tiao (1978), Miller (1980), and Chang (1982) showed that the effect of such interventions may cause serious bias in estimating autocorrelations, partial autocorrelations, and autoregressive moving average (ARMA) parameters. It is, therefore, important to be able to identify these interventions and remove their effects from the observations to better understand the underlying structure of the series. When the time of occurrence T of an intervention is known, its effect can often be accounted for using the intervention analysis techniques proposed by Box and Tiao (1975). In practice, however, information about the timing is seldom available. In this article, we shall develop a procedure for estimating ARMA parameters under that circumstance.

Following Fox (1972), Denby and Martin (1979), and Abraham and Box (1979), we shall focus our attentions on two simple intervention models that are believed to represent a large portion of the outliers likely to be found in practice. They are referred to as *innovational outlier* (IO) and *additive outlier* (AO), respectively. The model for an IO is

$$z_t = x_t + \frac{\theta(B)}{\phi(B)\alpha(B)} \omega\zeta_t^{(T)}, \qquad (1.3)$$

and that of an AO is

$$z_t = x_t + \omega\zeta_t^{(T)}, \qquad (1.4)$$

where $\theta(B)$, $\phi(B)$, and $\alpha(B)$ are defined as in (1.1). These models can be rewritten in terms of the innovation sequence a_t's as follows:

$$\text{(IO)} \quad z_t = \frac{\theta(B)}{\phi(B)\alpha(B)} \{a_t + \omega\zeta_t^{(T)}\}, \qquad (1.5)$$

and

$$\text{(AO)} \quad z_t = \frac{\theta(B)}{\phi(B)\alpha(B)} a_t + \omega\zeta_t^{(T)}. \qquad (1.6)$$

Thus the AO case may be called a *gross error* model, since only the level of the Tth observation is affected. On the other hand, an IO represents an extraordinary shock at time point T influencing $z_T, z_{T+1}, \ldots$ through the dynamic system described by $\theta(B)/[\phi(B)\alpha(B)]$. We hope that our study of these simple models might shed some light on the problems of more complex structure.

Section 2 discusses the likelihood ratio criteria for testing these two types of outliers. Section 3 presents a procedure for estimating ARMA parameters in the presence of these outliers. The power and performance of this procedure are examined and compared with those of some other approaches through simulation methods in Section 4. The proposed procedure is then applied to two examples in Section 5. Finally, some concluding remarks are given in Section 6.

2. LIKELIHOOD RATIO CRITERIA FOR DETECTING AN OUTLIER AND DETERMINING ITS NATURE

2.1 When ARMA Parameters and σ_a^2 Are Known

We shall first consider estimating the impact ω of an IO (1.3) and that of an AO (1.4), respectively, in a hypothetical situation in which all of the time series parameters, as well as the innovation variance σ_a^2 of the underlying process x_t, are known. Moreover, we assume that observations are available from $t = -J$ to $t = n$, where J is an integer much larger than $p + d + q$ and that $1 \le T \le n$. Let $\pi(B) = \phi(B)\alpha(B)/\theta(B) = 1 - \pi_1 B - \pi_2 B^2 - \cdots$. Because the zeros of $\theta(B)$ are all outside the unit circle, the weights π_j's for j beyond J would in practice become essentially equal to zero with J of moderate size.

Let $e_t = \pi(B)z_t$ for $t = 1, \ldots, n$. We can rewrite (1.5) and (1.6), respectively, as

$$\text{(IO)} \quad e_t = \omega\zeta_t^{(T)} + a_t,$$
$$\text{(AO)} \quad e_t = \omega\pi(B)\zeta_t^{(T)} + a_t. \qquad (2.1)$$

In other words, the information about an IO is contained in the "residual" e_T right at that particular point T, whereas that of an AO is scattered over a string of "residuals" $e_T, e_{T+1}, \ldots$. Write $X_t = \pi(B)\zeta_t^{(T)}$ so that $X_t = 0$ for $t < T$, $X_t = 1$ for $t = T$, and $X_t = -\pi_{t-T}$ for $t > T$. From least squares theory, the estimators of the impact, ω, in these two models are

$$\text{(IO)} \quad \tilde{\omega}_I = e_T,$$
$$\text{(AO)} \quad \tilde{\omega}_A = \rho^2\pi(F)e_T$$
$$= \rho^2(1 - \pi_1 F - \pi_2 F^2 - \cdots - \pi_{n-T}F^{n-T})e_T, \qquad (2.2a)$$

respectively, where $\rho^2 = (1 + \pi_1^2 + \pi_2^2 + \cdots + \pi_{n-T}^2)^{-1}$ and F is the forward-shift operator such that $Fe_t = e_{t+1}$. Thus, not surprisingly, the effect of an IO at time T can be best estimated by the residual e_T at that point, but the best estimate of the AO effect is a linear combination of $e_T, e_{T+1}, \ldots$, with weights depending on the model from which the underlying process x_t is generated. We note that, in terms of the observations z_t, the estimate $\tilde{\omega}_A$ in (2.2a) could be written as

$$\tilde{\omega}_A = \rho^2\pi(F)\pi(B)z_t. \qquad (2.2b)$$

For the case of (autoregressive) AR(p) model, this expression reduces to equation (2.5) of Fox (1972).

The variances of these estimators are

$$\text{var}(\tilde{\omega}_I) = \sigma_a^2, \qquad \text{var}(\tilde{\omega}_A) = \rho^2\sigma_a^2, \qquad (2.3)$$

respectively. Note that since $\rho^2 \le 1$, the variance of $\tilde{\omega}_A$ is at most as large as that of $\tilde{\omega}_I$, and in some cases it can be much smaller than σ_a^2. For example, when the underlying process x_t follows a first-order moving average (MA) model, $x_t = (1 - \theta B)a_t$, the variance for $\tilde{\omega}_A$ would be nearly $\sigma_a^2(1 - \theta^2)$ if the outlier does not occur near the end of the series.

Let H_0 denote the hypothesis that $\omega = 0$ in (1.5) and (1.6), H_1 denote the situation $\omega \ne 0$ in (1.5), and H_2 denote the situation $\omega \ne 0$ in (1.6). The likelihood ratio criteria can be derived for testing one hypothesis versus another among the following:

H_0 vs. H_1: $\lambda_{1,T} = \tilde{\omega}_I/\sigma_a$,

H_0 vs. H_2: $\lambda_{2,T} = \tilde{\omega}_A/\rho\sigma_a$,

Table 1. Estimated Level of Significance of Tests (2.7) and (2.8): 1,000 Replications

Time series model	C	T = 50 IO	T = 50 AO	T = 100 IO	T = 100 AO	T = 150 IO	T = 150 AO
AR(1), $\phi = .6$	3.0	23.3	18.3	33.9	31.3	38.7	35.0
	3.5	5.6	3.9	8.4	4.8	8.4	7.4
	4.0	1.3	1.1	1.2	.7	1.2	1.0
MA(1), $\theta = .6$	3.0	19.9	10.9	31.5	23.3	39.2	32.4
	3.5	4.3	2.9	7.1	4.2	9.4	5.6
	4.0	1.1	.6	2.1	.7	1.5	.8

NOTE: The estimated standard deviation of the estimated level of significance is $(\hat{p}(1 - \hat{p})/N)^{1/2}$, where N is the number of replications and $\hat{p}$ is the estimated significance level.

and

$$H_1 \text{ vs. } H_2: \quad \lambda_{3,T} = [\rho^{-2}\tilde{\omega}_A^2 - \tilde{\omega}_I^2]/[2\sigma_a^2(1 - \rho^2)^{1/2}]. \tag{2.4}$$

Under hypothesis H_0, the statistics $\lambda_{1,T}$ and $\lambda_{2,T}$ both have the standard normal distribution. To derive the distribution of $\lambda_{3,T}$, let

$$u = [\rho^{-1}\tilde{\omega}_A + \tilde{\omega}_I]/[\sigma_a[2(1 + \rho)]^{1/2}]$$

and

$$v = [\rho^{-1}\tilde{\omega}_A - \tilde{\omega}_I]/[\sigma_a[2(1 - \rho)]^{1/2}].$$

Then we have that $\lambda_{3,T} = uv$. The vector $(u, v)'$ has a bivariate normal distribution with mean $(m_1, m_2)'$ and covariance matrix I, where $m_1 = \omega[(1 + \rho)/(2\sigma_a^2)]^{1/2}$ and $m_2 = -\omega[(1 - \rho)/(2\sigma_a^2)]^{1/2}$ under the hypothesis H_1, and where $m_1 = \omega[(1 + \rho)/(2\rho\sigma_a^2)]^{1/2}$ and $m_2 = -\omega[(1 - \rho)/(2\rho\sigma_a^2)]^{1/2}$ under the hypothesis H_2. It is noted that the distribution of $\lambda_{3,T}$ depends only on the two parameters, ρ and the ratio ω/σ_a, in either case. Craig (1936) showed that the distribution of the product of two such independent and normally distributed random variables can be expressed as a convergent expression of Bessel functions. The percentage points of the distribution of $\lambda_{3,T}$ under hypothesis H_1 were tabulated via numerical integration for a range of values of ω/σ_a and ρ by Chang (1982).

The likelihood ratio method further leads to the criteria

$$\text{(IO)} \quad \max_{t=1,\ldots,n} |\lambda_{1,t}|, \qquad \text{(AO)} \quad \max_{t=1,\ldots,n} |\lambda_{2,t}|, \tag{2.5}$$

for testing the possibility of an IO or an AO, respectively, at an unknown position in the series $z_1, \ldots, z_n$.

2.2 When ARMA Parameters and σ_a^2 Are Unknown

In practice, the ARMA parameters and σ_a^2 are usually unknown. Estimates of these parameters, together with that of ω under either the IO or the AO case, can be obtained by maximizing the likelihood function of $(\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \omega, \sigma_a^2)$ in the same fashion as that described by Box and Jenkins (1976). Based on these estimates, the likelihood ratios can be computed accordingly for testing the hypotheses, one against another, in (2.4). The criteria for detecting an outlier at an unknown position then follow. Fox (1972) first discussed these ratio criteria in an AR context.

The maximum likelihood estimates for the parameters, as well as the likelihood ratios, do not, however, reduce to explicit expressions in general because

Table 2. Estimated Percentiles of the Test Statistics of Tests (2.7) and (2.8): 1,000 Replications

Time series model	%	T = 50 IO	T = 50 AO	T = 100 IO	T = 100 AO	T = 150 IO	T = 150 AO
AR(1), $\phi = .6$	10	3.29	3.21	3.45	3.36	3.45	3.43
	5	3.55	3.39	3.64	3.49	3.63	3.60
	1	4.11	4.05	4.03	3.88	4.13	4.00
MA(1), $\theta = .6$	10	3.25	3.02	3.40	3.24	3.48	3.36
	5	3.47	3.22	3.62	3.46	3.66	3.53
	1	4.02	3.80	4.27	3.88	4.04	3.95

NOTE: The standard deviation of the estimated pth percentile of the test statistics λ_p is approximately $(p(1 - p)/[Nf^2(\lambda_p)])^{1/2}$, where N is the number of replications and $f(\lambda)$ is the probability density function of the test statistics. Taking the normal approximation of $f(\lambda)$, we have found that standard deviations of the estimated 10% and 5% percentiles are less than .03 and those of the estimated 1% percentiles are less than .06.

of the nonlinear nature of the AO model (2.1) and that of the general ARMA models. Nonlinear estimation algorithms are, therefore, necessary to implement the likelihood ratio tests in this circumstance, and the computational burden will become very heavy when the timing of the outlier is unknown. Nevertheless, several simpler statistics arise naturally as possible approximations to these likelihood ratio criteria.

Let $\hat{\phi}_1, \ldots, \hat{\phi}_p; \hat{\theta}_1, \ldots, \hat{\theta}_q$; and $\hat{\sigma}_a^2$ be the maximum likelihood estimate of the parameters $\phi_1, \ldots, \phi_p; \theta_1, \ldots, \theta_q$; and σ_a^2, respectively, obtained by treating the time series z_t as containing no outliers. Moreover, let $\hat{e}_t$ be the residuals computed from such an estimated model and $\hat{\pi}(B) = \hat{\phi}(B)\alpha(B)/\hat{\theta}(B)$. We shall consider the following statistics:

$$\hat{\lambda}_{1,T} = \hat{\omega}_I/\hat{\sigma}_a, \qquad \hat{\lambda}_{2,T} = \hat{\omega}_A/(\hat{\rho}\hat{\sigma}_a), \tag{2.6}$$

and

$$\hat{\lambda}_{3,T} = \frac{\hat{\rho}^{-2}\hat{\omega}_A^2 - \hat{\omega}_I^2}{2\hat{\sigma}_a^2(1-\hat{\rho}^2)^{1/2}},$$

where

$$\hat{\omega}_I = \hat{e}_T,$$

$$\hat{\omega}_A = \hat{\rho}^2(1 - \hat{\pi}_1 F - \hat{\pi}_2 F^2 - \cdots - \hat{\pi}_{n-T} F^{n-T})\hat{e}_T,$$

and

$$\hat{\rho}^2 = (1 + \hat{\pi}_1^2 + \hat{\pi}_2^2 + \cdots + \hat{\pi}_{n-T}^2)^{-1}.$$

It can be shown that $\hat{\lambda}_{1,T}$, $\hat{\lambda}_{2,T}$, and $\hat{\lambda}_{3,T}$ are asymptotically equivalent to the likelihood ratio criteria $\lambda_{1,T}$, $\lambda_{2,T}$, and $\lambda_{3,T}$ in (2.4) for testing hypotheses H_0 versus H_1, H_0 versus H_2, and H_1 versus H_2, respectively, at a given point T (see Chang 1982).

To detect an IO or an AO at an unknown position, we can then scan through the sequence of $\hat{\lambda}_{1,t}$, $t = 1, \ldots, n$ or the sequence of $\hat{\lambda}_{2,t}$, $t = 1, \ldots, n$, respectively. In other words, the general possibility of an IO in the series or that of an AO in the series can be tested by

$$\hat{\eta}_{\mathrm{IO}} = \max_{t=1,\ldots,n} |\lambda_{1,t}| > C \tag{2.7}$$

or

$$\hat{\eta}_{\mathrm{AO}} = \max_{t=1,\ldots,n} |\lambda_{2,t}| > C, \tag{2.8}$$

respectively, where C is some suitably chosen positive constant. A simulation experiment was conducted on an AR model and an MA model to estimate the significance level of the tests (2.7) and (2.8) at critical values $C = 3$, 3.5, and 4 for sample sizes $n = 50$, 100, and 150. The results are given in Tables 1 and 2. Because of the extreme-value nature of the statistics (they are maximum of a set of random variables), the tail probability of the statistics is expected to increase as the sample size increases. This study shows that the estimated percentiles only increase moderately when n changes from 50 to 150; for instance, the 10% percentile of $\hat{\eta}_{\mathrm{AO}}$ changes from 3.21 to 3.43 in the AR(1) case and from 3.02 to 3.36 in the MA(1) case. We also observe that the estimated 1%, 5%, and 10% percentiles of $\hat{\eta}_{\mathrm{IO}}$ are slightly higher than those of $\hat{\eta}_{\mathrm{AO}}$. In practice, we recommend using $C = 3.0$ for high sensitivity, $C = 3.5$ for medium sensitivity, and $C = 4.0$ for low sensitivity in the outlier-detecting procedure when the length of the series is less than 200.

2.3 Distinguishing an AO From an IO

There is usually little information available in practice about what type the possible outlier might be; thus it is unclear which detection test, (2.7) or (2.8), is more appropriate for a given situation. When a test of an inappropriate type is used, the detecting power of the test could be substantially reduced. Furthermore, even if it is known that an outlier has occurred at a particular point, the possibly adverse effect of the outlier may not be easy to remove unless its nature is properly identified. The statistic $\hat{\lambda}_{3,T}$ in (2.6) is derived for making the distinction between an IO and an AO at a given point T. When the position of the possible outlier is unknown, however, we may need to perform such a test repeatedly at various time points, and it can become a cumbersome task.

To simplify the problem, we consider a simple rule mentioned by Fox (1972) as a possible way to distinguish between an IO and an AO. At any suspected point T, the possible outlier is classified as an IO if $|\hat{\lambda}_{1,T}| > |\hat{\lambda}_{2,T}|$, and it is classified as an AO if $|\hat{\lambda}_{1,T}| \le |\hat{\lambda}_{2,T}|$.

From a simulation experiment similar to that described in Section 2.2 on AR series of moderate length, $n = 50$, we have noted that on the average this simple rule can correctly identify the nature of an outlier, either IO or AO, at a particular point 78.6% of the time when the magnitude ω of the outlier is $3\sigma_a$, whereas the test based on $\hat{\lambda}_{3,T}$ at the nominal 10% level has an average success rate of 87.3%. When the magnitude ω is increased to $5\sigma_a$, however, this simple rule makes correct identification, based on the simulation study, 92% of the time, which is about the same as the 92.4% achieved by the test based on $\hat{\lambda}_{3,T}$ at the 10% level (Chang 1982).

3. AN ITERATIVE PROCEDURE FOR OUTLIER DETECTION AND PARAMETER ESTIMATION

The considerations in Section 2 have led to the following iterative procedure to handle situations in which there may exist an unknown number of IO's or AO's. The procedure begins with modeling the original series z_t by supposing that there is no outlier. Then the outlier-detection steps and the parameter-estimation step will be alternatively followed. A detailed description of the procedure follows:

3.1 Outlier-Detection Stage

1. From the estimated model, compute the residuals $\hat{e}_t$, and let $\hat{\sigma}_a^2 = n^{-1} \sum_{t=1}^{n} \hat{e}_t^2$ be the estimate of σ_a^2. A possible robust alternative estimate of $\hat{\sigma}_a$ may be based on the median of the absolute values of residuals.

2. Compute $\hat{\lambda}_{1,T}$ and $\hat{\lambda}_{2,T}$ in (2.2a) and let $\eta_t = \max\{|\hat{\lambda}_{1,t}|, |\hat{\lambda}_{2,t}|\}$ for $t = 1, \ldots, n$. If $\max_t \eta_t = |\hat{\lambda}_{1,T}| > C$, where C is a predetermined positive constant, then there is the possibility of an IO at T. The impact ω of this possible IO is estimated by $\hat{\omega}_I$ in (2.2a). We then eliminate its effect by defining a new residual $\check{e}_T = \hat{e}_T - \hat{\omega}_I = 0$ at T. If $\max_t \eta_t = |\hat{\lambda}_{2,T}| > C$, then there is the possibility of an AO at T and its impact is estimated by $\hat{\omega}_A$ in (2.2a). The effect of this AO can be removed by defining new residuals $\check{e}_t = \hat{e}_t - \hat{\omega}_A \hat{\pi}(B)\zeta_t^{(T)}$ for $t \geq T$. In either of the preceding cases, a new estimate $\check{\sigma}_a^2$ is computed from the modified residuals.

3. If an IO or an AO is identified in step 2, recompute $\hat{\lambda}_{1,T}$ and $\hat{\lambda}_{2,T}$ based on the same initial estimates of the time series parameters, but using the modified residuals $\check{e}_t$'s and the estimate $\check{\sigma}_a^2$, and repeat step 2.

4. Continue to repeat steps 2 and 3 until no further outlier candidates can be identified.

3.2 Parameter Estimation Stage

5. Suppose at k time points $T_1, T_2, \ldots, T_k$ are pinpointed for possible IO's or AO's. Treat these times as known, and estimate the outlier parameters $\omega_1, \omega_2, \ldots, \omega_k$ and the time series parameters simultaneously, as described by Box and Tiao (1975), using models of the form

$$z_t = \sum_{j=1}^{k} \omega_j L_j(B)\zeta_t^{(T_j)} + \frac{\theta(B)}{\phi(B)\alpha(B)} a_t, \qquad (3.1)$$

where $L_j(B) = 1$ for an AO and $L_j(B) = \theta(B)/[\phi(B)\alpha(B)]$ for an IO at $t = T_j$.

Treating model (3.1) as the entertained model, we start the outlier detection stage again. The notations $\hat{\pi}_j$'s, $\hat{\omega}_j$'s and $\hat{e}_t$ represent the estimated values obtained from the joint estimation of all parameters of model (3.1). If no other outliers are found, we stop. Otherwise, the estimation stage is repeated, with the newly identified outliers incorporated into model (3.1), until no more outliers can be found and all of the outlier effects have been simultaneously estimated with the time series parameters.

4. PERFORMANCE OF THE ITERATIVE PROCEDURE

We have conducted simulation studies to obtain some information about the performance of the preceding procedure. Results on (a) the powers of detecting outliers and (b) the estimation of time series parameters will be reported in this section.

4.1 Power of the Iterative Outlier Detection Procedure

The proposed procedure is designed to detect the existence and to identify the types of outliers simultaneously. To assess the power of the procedure, we should be concerned with the probability of correctly detecting the location of outliers and the probability of correctly identifying the type of outliers. In this study, we focus on the issues of (a) single or multiple outliers, (b) time series structure of x_t, (c) types of outliers, (d) size of outliers, and (e) sample size of the series. For factor (a), we consider the cases of one outlier and two outliers.

In the one-outlier case, we will examine 24 situations that result from combinations of two series models for x_t, AR(1) with $\phi = .6$, $\sigma_a^2 = 1$ and MA(1) with $\theta = .6$, $\sigma_a^2 = 1$; two types of outliers, IO and AO; two different sizes of outliers, $\omega = 3\sigma_a$ and $\omega = 5\sigma_a$; and three sample sizes, $n = 50$, 100, and 150. The location of the outlier is set in the middle of the observational period, specifically $T = 26$ when $n = 50$, $T = 51$ for $n = 100$, and $T = 76$ for $n = 150$.

For the two outliers case, we consider the AR(1) model with $\phi = .6$, and $\sigma_a^2 = 1$ for x_t; three types of outliers, two AO's, two IO's, and one IO, one AO; two sizes, both outliers having $\omega = 3\sigma_a$ and both having $\omega = 5\sigma_a$; and three sample sizes, $n = 50$, 100, and 150. We assume that outliers occurred at the one-third point from the beginning period and the one-third point from the ending period, or $T_1 = 17$, $T_2 = 34$ for $n = 50$; $T_1 = 34$, $T_2 = 66$ for $n = 100$; and $T_1 = 51$, $T_2 = 101$ for $n = 150$.

In each case of the preceding design, time series data are generated according to the particular specification. Assuming that the model is known and that there is no outlier in the sample, the exact likelihood estimation method, considered by Hillmer and Tiao (1979), will be used to estimate time series parameters and to produce the residuals. Based on the estimated parameters and residuals, one iteration of the outlier-detection steps will be conducted with three critical values, $C = 3$, 3.5, and 4, individually. The outcomes of this detection procedure are then compared with the specification of the alternatives. Repeating such

Table 3. Frequency of Correct Detection of Outliers–Location (percentage of correct identification of types): Critical Value C = 4.0; 1,000 Replications

	$\omega = 3\sigma_a$			$\omega = 5\sigma_a$		
n	*50*	*100*	*150*	*50*	*100*	*150*
AR, 1 AO	236	266	278	857	945	948
	(.72)	(.80)	(.81)	(.91)	(.90)	(.92)
AR, 1 IO	188	185	174	777	837	820
	(.88)	(.87)	(.87)	(.95)	(.93)	(.91)
MA, 1 AO	172	248	283	814	914	952
	(.70)	(.82)	(.84)	(.83)	(.90)	(.93)
MA, 1 IO	157	194	185	769	832	837
	(.97)	(.92)	(.88)	(.99)	(.96)	(.97)
AR, 2 AO'S	39	70	75	681	854	847
	(.38)	(.54)	(.52)	(.53)	(.76)	(.80)
1st Outlier	135	227	247	774	921	914
	(.65)	(.74)	(.81)	(.76)	(.86)	(.92)
2nd Outlier	141	204	252	759	917	920
	(.64)	(.73)	(.78)	(.74)	(.88)	(.88)
AR, 2 IO'S	24	24	33	557	727	682
	(.58)	(.67)	(.55)	(.82)	(.86)	(.84)
1st Outlier	120	137	162	689	835	800
	(.84)	(.80)	(.85)	(.91)	(.94)	(.91)
2nd Outlier	102	136	159	674	831	833
	(.77)	(.88)	(.83)	(.92)	(.91)	(.92)
AR, IO AO	52	41	48	689	781	788
	(.65)	(.63)	(.69)	(.86)	(.87)	(.88)
1st Outlier	136	142	133	718	813	828
	(.90)	(.86)	(.84)	(.98)	(.95)	(.93)
2nd Outlier	184	216	268	842	929	941
	(.72)	(.81)	(.81)	(.88)	(.92)	(.93)

operations, we may estimate the probability of correctly detecting an outlier's location and the probability of correctly identifying the outlier's type given that the location has been correctly detected. Tables 3, 4, and 5 give the results of 1,000 replications of the preceding operations for each of the alternative specifications using critical values $C = 3$, $C = 3.5$, and $C = 4$, respectively. The tables give the frequency with which the location of an outlier was correctly detected. Numbers in parentheses are percentages of correct identification of an outlier's type given that the location has been detected correctly. In the two-outlier cases, rows labeled by 1st and 2nd show the results of detecting and identifying the first outlier and the second outlier, respectively.

In general, the power (probability of correct detection or identification of outliers) of the procedure increases when sample size increases and decreases as the critical value C increases. For large-size outliers, $\omega = 5\sigma_a$, the procedure seems to be fairly powerful. Probabilities of correct detection, using $C = 3.5$, range from 89.6% to 98.8% for one-outlier cases and from 79.2% to 95.2% for two-outlier cases. The percentages of correct identification of outlier types ranges from 76% to 98% except for the case of two AO's with $n = 50$. For medium-size outliers, $\omega = 3\sigma_a$, the performance of the procedure is not as good. Especially when there is more than one outlier, the procedure may very likely miss some of them. Note that the results here are from the first round of outlier-detection procedure. In practice, the parameter-estimation stage will be followed according to the outlier-detection results, and the second-round outlier detection may be conducted if necessary. Hence the actual power of the procedure may be higher than what we report here, especially for the two-outliers case. Finally, note that the estimated standard deviation of the estimated power is $[\hat{p}(1 - \hat{p})/N]^{1/2}$, where N is the number of replications and $\hat{p}$ is the estimated power.

4.2 Estimation in the Presence of Outliers

In this study the estimates of the time series parameters computed from the proposed procedure were compared with the M estimates and the generalized M estimates (GM) proposed by Denby and Martin (1979) and Martin (1980).

Time series were generated from an AR(1) model under several different situations. In addition to the null case, the IO case (1.3), and the AO case (1.4), we

Table 4. Frequency of Correct Detection of Outliers–Location (percentage of correct identification of types): Critical Value C = 3.5; 1,000 Replications

	$\omega = 3\sigma_a$			$\omega = 5\sigma_a$		
n	*50*	*100*	*150*	*50*	*100*	*150*
AR, 1 AO	443 (.75)	482 (.80)	481 (.82)	947 (.86)	982 (.90)	975 (.92)
AR, 1 IO	319 (.85)	335 (.88)	336 (.81)	901 (.94)	936 (.93)	917 (.90)
MA, 1 AO	326 (.75)	477 (.83)	492 (.84)	927 (.84)	975 (.91)	989 (.93)
MA, 1 IO	290 (.94)	346 (.90)	347 (.89)	896 (.98)	935 (.96)	924 (.96)
AR, 2 AO'S	159 (.44)	220 (.54)	229 (.62)	862 (.58)	943 (.76)	952 (.80)
1st Outlier	337 (.70)	428 (.75)	466 (.81)	907 (.78)	968 (.86)	974 (.92)
2nd Outlier	333 (.70)	400 (.76)	452 (.80)	911 (.77)	973 (.88)	977 (.88)
AR, 2 IO'S	89 (.66)	100 (.65)	127 (.61)	792 (.82)	881 (.85)	854 (.82)
1st Outlier	246 (.83)	298 (.81)	314 (.84)	877 (.91)	939 (.94)	911 (.91)
2nd Outlier	241 (.82)	281 (.85)	338 (.84)	870 (.91)	929 (.90)	930 (.91)
AR, IO AO	157 (.64)	124 (.69)	180 (.65)	861 (.86)	907 (.87)	901 (.88)
1st Outlier	293 (.88)	275 (.84)	316 (.84)	882 (.97)	920 (.95)	925 (.93)
2nd Outlier	379 (.75)	410 (.82)	506 (.79)	954 (.88)	979 (.92)	975 (.93)

also considered the following $100(\gamma_a + \gamma_v - \gamma_a\gamma_v)\%$ contamination model in which the number of outliers is determined by a random mechanism:

$$y_t = \phi y_{t-1} + a_t^*, \qquad z_t = y_t + v_t, \tag{4.1}$$

where z_t's are the observations, a_t^*'s are iid with a contaminated normal density $(1 - \gamma_a)N(0, \sigma_a^2) + \gamma_a N(0, \kappa_a^2\sigma_a^2)$, v_t's are iid with density $(1 - \gamma_v)\delta(0) + \gamma_v N(0, \kappa_v^2\sigma_a^2)$, $\delta(0)$ represents a degenerate density at 0, and v_t's and a_t^*'s are all independent. This model, considered also by Denby and Martin (1979), may generate IO's or AO's or both. We shall label it as a *mixture model* in the following text.

For each simulated series, the following estimates of the AR(1) parameter ϕ were computed:

LS—The least squares estimate.

M-H—*M* estimate with Huber-type "down-weighting" function for residuals. The tuning constant C_{a_H} is set at 1.5.

M-B—*M* estimate with a bisquare type "down-weighting" function for residuals. The tuning constant C_{a_B} is set at 6.0.

GM-H—Generalized *M* estimate with two Huber-type "down-weighting" functions for residuals and observations, respectively. The tuning constants are set as $C_{a_H} = 1.5$ and $C_{Z_H} = 1.0$.

GM-B—Generalized *M* estimate with two bisquare type "down-weighting" functions for residuals and observations, respectively. The tuning constants are set as $C_{a_B} = 6.0$ and $C_{Z_B} = 3.9$.

P—The estimate computed from the procedure proposed in Section 3 with the critical value $C = 3.0$.

Following Denby and Martin (1979), we computed four iterations of the iterated weighted least squares (IWLS) algorithm (Beaton and Tukey 1974) for the M-H and GM-H estimates using the LS estimate as the starting point. The M-H and GM-H estimates were then used as starting values for another four iterations of IWLS to obtain the M-B and GM-B estimates, respectively. In these IWLS iterations, the scale parameter σ_a was estimated by the median of the absolute values of the residuals divided by .6745, and the scale parameter σ_z was estimated by the median of the absolute deviations of the observations from their sample median divided by .6745.

Since all estimates here, except the LS estimate, were computed from an iterative algorithm, a convergence criterion needs to be set. We considered an

Table 5. Frequency of Correct Detection of Outliers–Location (percentage of correct identification of types): Critical Value C = 3.0; 1,000 Replications

	$\omega = 3\sigma_a$			$\omega = 5\sigma_a$		
n	*50*	*100*	*150*	*50*	*100*	*150*
AR, 1 AO	644	703	692	984	993	990
	(.75)	(.81)	(.81)	(.86)	(.90)	(.92)
AR, 1 IO	504	530	564	959	985	975
	(.83)	(.82)	(.77)	(.93)	(.93)	(.90)
MA, 1 AO	547	672	703	976	993	996
	(.78)	(.83)	(.85)	(.84)	(.91)	(.93)
MA, 1 IO	486	536	552	958	981	973
	(.91)	(.87)	(.88)	(.98)	(.96)	(.97)
AR, 2 AO'S	321	460	453	953	985	986
	(.50)	(.59)	(.63)	(.60)	(.76)	(.80)
1st Outlier	531	652	676	975	991	993
	(.74)	(.77)	(.81)	(.79)	(.86)	(.91)
2nd Outlier	561	636	655	973	994	993
	(.72)	(.77)	(.80)	(.77)	(.88)	(.88)
AR, 2 IO'S	217	283	300	908	962	951
	(.61)	(.63)	(.65)	(.81)	(.85)	(.82)
1st Outlier	444	522	518	956	983	972
	(.82)	(.79)	(.85)	(.91)	(.93)	(.90)
2nd Outlier	449	498	552	946	975	978
	(.78)	(.82)	(.80)	(.90)	(.90)	(.91)
AR, IO AO	347	364	404	934	976	964
	(.65)	(.68)	(.65)	(.86)	(.86)	(.87)
1st Outlier	480	516	537	947	984	973
	(.87)	(.82)	(.84)	(.97)	(.94)	(.93)
2nd Outlier	645	658	712	986	1000	991
	(.76)	(.83)	(.79)	(.89)	(.91)	(.93)

algorithm as having converged here if the estimates given by the most recent two consecutive iterations differed from each other by less than .0001.

For each outlier model, 500 replications of the same length were generated with $\sigma_a^2 = 1$ and, for the mixture model (4.1), $\sigma_v^2 = 1$. The lengths were moderate, $n = 50$ or 75. The means and the standard deviations, SD's as well as the mean squared errors (MSE's) of the estimates of the parameter ϕ, were computed over the 500 replications. Tables 6 and 7 present the results in two different cases, $\phi = .6$ and $\phi = .9$, respectively. The values of the commonly adopted approximation, $((1 - \phi^2)/n)^{1/2}$, to the standard deviation of the LS estimate of ϕ in the null case are also shown at the bottom of both tables for reference.

Note from these results that the bias of the *M* estimates in the AO case is almost as severe as that of the LS estimate. This point was also made earlier by Denby and Martin (1979). On the other hand, the generalized *M* estimates are less efficient in the null and the IO cases. The estimate computed from our proposed procedure, however, gives the smallest MSE's in most of the cases studied here. It compares favorably with the robust procedures M-H, M-B, and GM-H in every case, and only in the AO situation with $\phi = .6$ the proposed procedure is sometimes slightly worse off relative to GM-B.

The MSE's of the estimates from the proposed procedure in most cases are fairly close to those of the LS estimates in the null case, possibly because of the reasonably high success rate of the procedure in detecting outliers (see Table 3). Moreover, note that almost all of the GM-B estimates were obtained here after eight iterations of IWLS and that, throughout this study, our proposed procedure stopped at or before the fourth estimation cycle. As to the precision of this Monte Carlo study, assuming the sampling distribution of $\hat{\phi}$ is normal, the estimated standard deviation of the estimated MSE's is approximately $(2/N)^{1/2}\hat{\sigma}_{\hat{\phi}}^2$, where N is the number of replications and $\hat{\sigma}_{\hat{\phi}}^2$ is the estimated variance of $\hat{\phi}$. The choice that $N = 500$ assures that the estimated SD(MSE) in all of the cases considered is less than .002.

5. ILLUSTRATIVE EXAMPLES

Two examples will be analyzed here to illustrate the application of the proposed procedure described in Section 3.

Table 6. Means, Standard Deviations, and Mean Squared Errors of the Estimates of the AR(1) Parameter (true $\phi = .6$)

Outlier model	LS	M-H	M-B	GM-H	GM-B	p
Null	.5885	.5883	.5885	.5947	.5940	.5933
$n = 50$	(.1143)	(.1157)	(.1147)	(.1223)	(.1225)	(.1165)
	.0132	.0135	.0133	.0150	.0150	.0136
1 IO (fixed)	.5804	.5808	.5814	.5813	.5756	.5932
$T = 24, n = 50,$	(.1110)	(.0989)	(.0978)	(.1125)	(.1251)	(.0968)
$\omega = 5$	.0127	.0101	.0099	.0130	.0162	.0094
1 AO (fixed)	.4456	.4746	.4721	.5441	.5724	.5802
$T = 24, n = 50,$	(.1270)	(.1294)	(.1332)	(.1298)	(.1283)	(.1304)
$\omega = 5$	.0399	.0324	.0341	.0199	.0172	.0174
2 AO'S (fixed)	.4013	.4239	.4202	.5123	.5513	.5496
$T = 17, 34;$	(.1369)	(.1386)	(.1418)	(.1356)	(.1343)	(.1428)
$\omega = -3, 5; n = 50$	.0582	.0502	.0524	.0260	.0204	.0229
2 IO'S (fixed)	.5809	.5847	.5849	.5886	.5854	.5998
$T = 17, 34;$	(.1123)	(.1008)	(.1014)	(.1110)	(.1192)	(.1011)
$\omega = -3, 5; n = 50$	.0130	.0104	.0105	.0124	.0144	.0102
3 AO'S (fixed)	.3265	.3350	.3283	.4505	.5151	.4971
$T = 14, 26, 38;$	(.1354)	(.1300)	(.1319)	(.1393)	(.1524)	(.1686)
$\omega = -4, 5, 4; n = 50$	.0931	.0871	.0912	.0417	.0304	.0390
3 AO'S (fixed)	.3863	.4060	.3990	.5075	.5541	.5649
$T = 20, 42, 61;$	(.1112)	(.1117)	(.1109)	(.1064)	(.1054)	(.1136)
$\omega = -4, 5, 4; n = 75$	.0580	.0501	.0527	.0199	.0132	.0141
Mixture	.4929	.5056	.5094	.5470	.5604	.5606
$\gamma_a = .02, \gamma_v = .02,$	(.1559)	(.1460)	(.1455)	(.1241)	(.1246)	(.1242)
$\kappa_a^2 = 25, \kappa_v^2 = 25, n = 50$	.0357	.0302	.0293	.0182	.0171	.0170
Mixture	.4540	.4720	.4727	.5183	.5360	.5448
$\gamma_a = .05, \gamma_v = .05,$	(.1617)	(.1478)	(.1491)	(.1308)	(.1354)	(.1289)
$\kappa_a^2 = 16, \kappa_v^2 = 16, n = 50$	.0474	.0382	.0384	.0237	.0224	.0196

NOTE: $((1 - \phi^2)/n)^{1/2} \approx .1131$ when $n = 50$. The first entry in each case is the mean; the second entry (in parentheses) is the standard deviation; the third one is the mean squared error.

5.1 Box–Jenkins Series A

The first example is series A from Box and Jenkins (1976). The data are "uncontrolled" concentration readings of a chemical process recorded at every two-hour interval. Two models were suggested by Box and Jenkins (1976, p. 239):

$$x_t = c_0 + \frac{(1 - \theta B)}{(1 - \phi B)} a_t \tag{5.1}$$

and

$$(1 - B)x_t = (1 - \theta B)a_t. \tag{5.2}$$

In our iteration cycle 1, preliminary estimates of model (5.1) and model (5.2) are obtained assuming that there is no outlier, and the results are given in Table 8. Conducting steps 1 and 2 of the procedure in Section 3 using $C = 3.5$ based on the preliminary estimates leads to identification of an IO at $T = 64$ for both models and then the effect of this IO is removed by modifying the residuals accordingly. Following step 3 by using the modified residuals $\check{e}_t$'s and the estimated $\check{\sigma}_a^2$, an AO is found at $T = 43$ in both models. The next iteration of steps 2 and 3 shows that no outlier is identified. Then the time series parameters and the outlier parameters ω_1 and ω_2 are estimated simultaneously as described in the estimation stage. Time series parameters estimates are given under estimation cycle 2 in Table 8. Repeating the outlier-detection stage based on the newly estimated parameters and residuals, no other outlier is found.

For both models, adjusting the effect of two outliers results in changes of parameter estimates and reduction in the estimated σ_a^2. As might be expected, the procedure identifies the same types and locations of outliers for two somewhat different models of the same data.

5.2 The Los Angeles Oxidant Data

The monthly average of hourly readings of ozone in downtown Los Angeles from January 1955 to December 1972 is examined. We adopt the following

Table 7. Means, Standard Deviations, and Mean Squared Errors of the Estimates of the AR(1) Parameter (true $\phi = .9$)

Outlier model	LS	M-H	M-B	GM-H	GM-B	p
Null	.8638	.8639	.8639	.8682	.8657	.8651
$n = 50$	(.0798)	(.0811)	(.0803)	(.0838)	(.0876)	(.0793)
	.0077	.0079	.0077	.0080	.0088	.0075
1 IO (fixed)	.8688	.8752	.8771	.8786	.8783	.8813
$T = 24, n = 50,$	(.0787)	(.0708)	(.0703)	(.0711)	(.0717)	(.0673)
$\omega = 5$	.0071	.0056	.0055	.0055	.0056	.0049
1 AO (fixed)	.7686	.8163	.8264	.8440	.8555	.8658
$T = 24, n = 50,$	(.1155)	(.1073)	(.1137)	(.0938)	(.0894)	(.0819)
$\omega = 5$	.0306	.0185	.0183	.0119	.0100	.0079
2 AO'S (fixed)	.7483	.7927	.7962	.8267	.8353	.8601
$T = 17, 34;$	(.1189)	(.1106)	(.1173)	(.0952)	(.0924)	(.0855)
$\omega = -3, 5; n = 50$	.0371	.0237	.0245	.0144	.0127	.0089
2 IO'S (fixed)	.8678	.8745	.8768	.8768	.8765	.8802
$T = 17, 34;$	(.0693)	(.0631)	(.0629)	(.0645)	(.0698)	(.0635)
$\omega = -3, 5; n = 50$	.0058	.0046	.0045	.0047	.0054	.0044
3 AO'S (fixed)	.6845	.7360	.7340	.7932	.8078	.8554
$T = 14, 26, 38;$	(.1420)	(.1421)	(.1507)	(.1131)	(.1084)	(.1000)
$\omega = -4, 5, 4; n = 50$	.0665	.0470	.0502	.0242	.0202	.0120
3 AO'S (fixed)	.7480	.7956	.7961	.8301	.8393	.8743
$T = 20, 42, 61;$	(.1016)	(.0970)	(.1046)	(.0814)	(.0787)	(.0675)
$\omega = -4, 5, 4; n = 75$	.0334	.0203	.0217	.0115	.0099	.0052
Mixture	.8110	.8420	.8541	.8592	.8621	.8688
$\gamma_a = .02, \gamma_v = .02,$	(.1268)	(.0914)	(.0837)	(.0756)	(.0787)	(.0718)
$\kappa_a^2 = 25, \kappa_v^2 = 25, n = 50$	.0240	.0117	.0091	.0074	.0076	.0061
Mixture	.7760V.8115	.8196	.8344	.8403	.8551	
$\gamma_a = .05, \gamma_v = .05,$	(.1407)	(.1180)	(.1196)	(.0955)	(.0965)	(.0879)
$\kappa_a^2 = 16, \kappa_v^2 = 16, n = 50$	.0351	.0217	.0207	.0134	.0129	.0097

NOTE: $((1 - \phi^2)/n)^{1/2} \approx .0616$ when $n = 50$.

intervention model considered by Box and Tiao (1975):

$$y_t = \omega_1 \eta_{t1} + \omega_2 \frac{\eta_{t2}}{(1 - B^{12})} + \omega_3 \frac{\eta_{t3}}{(1 - B^{12})} + \frac{(1 - \theta_1 B)(1 - \theta_{12} B^{12})}{(1 - B^{12})} a_t, \quad (5.3)$$

where

$$\eta_{t1} = 0, \quad t < 60$$
$$= 1, \quad t \geq 61,$$
$$\eta_{t2} = 1, \quad t = 137 + 12l + k, \; k = 1, \ldots, 6, \; l = 0, 1, \ldots$$
$$= 0, \quad \text{otherwise},$$

Table 8. Box–Jenkins Series A: Outlier Detection and Parameter Estimation

	Parameters				Outliers			
Estimation cycle	c_0	$\hat{\phi}$	$\hat{\theta}$	$\hat{\sigma}_a^2 \times 10^3$	Time point	Type	$\hat{\omega}$	$\hat{\lambda}$
Model (5.1)								
1	17.1	.91	.59	97.53	64	IO	1.12	3.72
	(.11)	(.04)	(.08)		43	AO	−1.00	−3.68
2	17.1	.89	.47	83.52				
	(.09)	(.04)	(.09)					
Model (5.2)								
1			.70	100.74	64	IO	1.13	3.67
			(.05)		43	AO	−.98	−3.53
2			.63	88.04				
			(.05)					

NOTE: Numbers in parentheses are standard errors of the estimated parameters.

Table 9. Los Angeles Oxidant Data: Outlier Detection and Parameter Estimation

Estimation cycle	Parameters						Outliers			
	$\hat{\theta}_1$	$\hat{\theta}_{12}$	$\hat{\sigma}_a^2 \times 10^3$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\omega}_3$	Time point	Type	ω	λ
1	−.27 (.07)	.78 (.04)	616.1	−1.34 (.19)	−.24 (.06)	−.10 (.05)	21 54	AO IO	2.40 2.21	3.61 3.14
2	−.30 (.06)	.80 (.04)	552.2	−1.29 (.18)	−.23 (.05)	−.10 (.05)				

and

$$\eta_{t3} = 1, \quad t = 142 + 12l + k, k = 1, \ldots, 6, l = 0, 1, \ldots$$
$$= 0, \quad \text{otherwise.}$$

The preliminary estimates of the intervention parameters (ω_1, ω_2, ω_3) and the time series parameters are given in Table 9 under estimation cycle 1 assuming no outlier. Let

$$\check{y}_t = y_t + \hat{\omega}_1 \eta_{t1} - \hat{\omega}_2 \eta_{t2}/(1 - B^{12}) - \hat{\omega}_3 \eta_{t3}/(1 - B^{12})$$

be the modified observations. The statistics $\hat{\lambda}_{1t}$ and $\hat{\lambda}_{2t}$ are calculated as described in step 2 of our procedure. From these statistics an AO is identified at $t = 21$. The effect of this AO can be removed by calculating the modified residuals accordingly. Following step 3 of the procedure, a new set of $\hat{\lambda}_{1t}$ and $\hat{\lambda}_{2t}$ are calculated. These indicate a potential IO at time $t = 54$. One more iteration of modifying residuals and recalculating $\hat{\lambda}_{1t}$ and $\hat{\lambda}_{2t}$ does not show evidence of further outliers. Then the time series parameters, intervention parameters as well as outlier parameters, are estimated simultaneously as described in step 4 of the procedure, and the time series and intervention parameter estimates are reported in Table 9 under estimation cycle (2). The next iteration of outlier-detection stage does not indicate the existence of additional outliers. In this example, adjusting the effect of two outliers results in slight changes of the estimated time series and intervention parameters and about a 10% reduction in $\hat{\sigma}_a^2$.

6. CONCLUDING REMARKS

A summary of the iterative outlier detection procedure proposed in Section 3 can be found in a review paper on modeling issues in the seasonal adjustment of economic time series by Hillmer, Bell, and Tiao (1983). They applied the procedure to the logarithms of monthly retail sales of variety stores from January 1967 to September 1979 obtained from the U.S. Bureau of the Census. The model for the series involves the differencing operation $(1 - B)(1 - B^{12})$. It is interesting to note that of the eight AO's and IO's detected, three consecutive IO's occurred beginning at the time point $t_0 = 112$ (April 1976). These three IO's corresponded to a level drop at $t_0 = 112$ and turned out to be associated with the event that a major variety-store chain went out of business at that time. Although the proposed procedure is designed to detect and to distinguish two simple types of outliers, the IO's and the AO's, in this example it actually revealed the effect of a more complex intervention. A more efficient way to detect level shifts at unknown time points was considered by Bell (1983) and Chen (1984).

In this article, the proposed procedure is demonstrated to be useful for estimating time series parameters when there is the possibility of outliers. It can be applied to all invertible ARIMA models. Moreover, it is flexible and easy to interpret, and it can be implemented with very few modifications to existing software packages capable of dealing with ARIMA and transfer function models. In practice, we suggest that this procedure be used in conjunction with other diagnostic tools for time series analysis to produce even better results.

Much further work is needed to investigate the variances and other sampling properties of the resulting estimates of time series parameters from the procedure proposed in Section 3. The simulation results in Section 4 on the AR(1) model seem to suggest that for moderate sample size the variance of such estimates might not be too far from those of the LS estimates obtained in the ideal null case.

Finally, the proposed procedure assumes that the form of the ARIMA model is correctly specified. Depending on the model specification (or identification) tool employed, the existence of AO's and IO's could conceivably lead to misspecification of the model form. As a possible solution to this problem, when an outlier is detected at $t = T$, one may simply adjust the observation z_T from (1.3) or (1.4) by employing the estimated effect of the outlier, and respecify the model accordingly.

ACKNOWLEDGMENTS

We express our appreciation to the editor, an associate editor, and the referees for their helpful comments. Our research was supported partially by grants from the Wisconsin Alumni Research Foundation and the Alcoa Foundation.

[Received July 1983. Revised August 1987.]

REFERENCES

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

Beaton, A. E., and Tukey, J. W. (1974), "The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data," *Technometrics*, 16, 147–185.

Bell, W. R. (1983), "A Computer Program for Detecting Outliers in Time Series," in *Proceedings of the Business and Economic Statistics Section, American Statistical Association*, pp. 634–639.

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 Economic and Environmental Problems," *Journal of the American Statistical Association*, 70, 70–79.

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

Chen, C. (1984), "Random Level Shift Time Series Model and a Diagnostic Method," unpublished Ph.D. thesis, University of Wisconsin–Madison, Dept. of Statistics.

Craig, C. C. (1936), "On the Frequency Function of *xy*," *Annals of Mathematical Statistics*, 7, 1–15.

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

Fox, A. J. (1972), "Outliers in Time Series," *Journal of the Royal Statistical Society*, Ser. B, 43, 350–363.

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.

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.

Hillmer, S. C., and Tiao, G. C. (1979), "Likelihood Function of Stationary Multiple Autoregressive Moving Average Models," *Journal of the American Statistical Association*, 74, 652–660.

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

Miller, R. B. (1980), Comment on "Robust Estimation of Autoregressive Models," in *Direction in Time Series*, eds. D. R. Brillinger and G. C. Tiao, Hayward, CA: Institute of Mathematical Statistics.