DETECTION OF OUTLIERS IN TIME SERIES DATA

Mintu Kr. Das[1] and Bipin Gogoi[2]
[1]Research Scholar, [2]Professor, Department of Statistics, Dibrugarh University

Abstract

Time series data are often subjected to outlying observations. Depending on their type/nature, time series outliers may have a moderate to significant impact on the various aspects for time series analysis such as, it can mislead the model identification, it can yield biased parameter estimates, and may results into poor forecasting. This papers provides a comparative analysis of two outlier detection procedures. Simulation study is carried out to assess the relative performance and conclusions are drawn accordingly.

Keywords: Additive outlier, innovational outlier, ARMA model, stationarity.

1. Introduction:

A time series is a collection of observations made sequentially through time. In almost all fields of inquiry, it is customary to work with data recorded over time. Examples occur in a variety of fields/disciplines including physics, astronomy, oceanography, engineering, economics, demography and management science. Theforecasting of time series is an important area in different applied fields, in which past observations of the

same variable are collected and analyzed to develop a model describing the underlying relationship. The model is then used to extrapolate the time series into the future. The significance of this modeling approach lies in the fact that when little knowledge is available on the underlying data generating process or when there is no satisfactory explanatory model that which the prediction variable to other explanatory variables (Zhang, 2003).

An outlier is any value that is numerically distant from most of the other data points in a set of data. Grubbs (1969) defined outlier as an observation which appears to deviate markedly from other members of the sample in which it occurs. Owing to various reasons, time series data are often subject to uncontrolled or unanticipated interventions, from which various types of outlying observations are produced. Besides possible gross errors, time series data are often caused to experience the influence some non-repetitive events, such as, implementation of a new regulation, major changes in political or economic policy, or occurrence of a disaster, strikes, outbreaks of wars, sudden changes in the market structure of a commodity, unexpected changes of certain condition in a physical system etc. As a result, discordant observations and various types of structural changes occur frequently in time series data (Chen and Liu, 1993).

The forecasting may be unduly effected by the presence of outlying values at one or more time points. Generally the forecasting methods assumes that the underlying time series are stationary or they can be made stationary with appropriate transformations. Generally there are five methods of forecasting based on time series, viz., exponential smoothing method, single-equation regression method, simultaneous-equation regression method, autoregressive integrated moving average (ARIMA) method and vector autoregressive (VAR) method.In the outlier detection matter our work will mainly deal with univariate autoregressive integrated moving average models.

Here we are interested to evaluate the empirical powers of two detection procedures for AO and IO. The rest of the paper is organized as follows: In Section 2, the outlier models for time series are presented. In Sec 3, the detection procedures to be compared are briefly described.The simulation study are given in Section 4 and it results are followed in the next Sections. Finally concluding remarks are made in Section 6.

2. Notation and model:

Considering Y_t as the state or output of a stochastic system at time t, we can write a simple stationary univariate model observed over the sequence of time $t = 1, 2, \ldots, T$, as follows:

$$Y_t = \phi Y_{t-1} + u_t \tag{2.1}$$

with mean: $E(Y_t) = \mu$; variance: $var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$; autocovariance: $\gamma_l = E[(Y_t - \mu)(Y_{t+1} - \mu)]$ at lag l; where $|\phi| < 1$(to be stationary), u_t is a white noise error term with zero mean and constant variance σ^2. This model resembles with the Markov first-order autoregressive or simply AR(1) model (Gujarati et al. 2011). A moving average process is a linear combination of white noise disturbance terms. Any stationary AR (p) model can be expressed as an MA(∞) model. A time series Y_t having both the characteristics of autoregressive and moving average, then the series is said to follow an autoregressive moving average (ARMA) process. For example an ARMA process with a constant and with one term each is defined as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \theta_0 u_t + \theta_1 u_{t-1} \tag{2.2}$$

Similarly a process with p AR terms and q MA terms or the ARMA (p, q) process can be defined. In real life many economic time series is nonstationary. And to deal with these situations, these have to be integrated. If we have to difference a time series d times to make it stationary and then apply *ARMA* (p, q) model to it, then the original time series is called *ARMA* (p, d, q), i.e., autoregressive integrated moving average time series and is given as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_0 u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q} \tag{2.3}$$

The above model can be condensed with the use of backshift notation $\boldsymbol{B}$, such that $\boldsymbol{B}Y_t = Y_{t-1}$;

$$(1 - \phi_1 \boldsymbol{B} - \phi_2 \boldsymbol{B}^2 - \cdots - \phi_p \boldsymbol{B}^p)(1 - \boldsymbol{B})^d Y_t = \phi_0 + (1 - \phi_1 \boldsymbol{B} - \phi_2 \boldsymbol{B}^2 - \cdots - \phi_q \boldsymbol{B}^q) u_t \tag{2.4}$$

Or, $$\phi(\boldsymbol{B})Y_t = \theta(\boldsymbol{B})u_t \qquad (2.5)$$

where, $(1- \boldsymbol{B})^d Y_t = Y_t - Y_{t-d}$;

considering $d = 0$, we may express other two forms of the ARMA model by long division

$$\frac{\theta(\boldsymbol{B})}{\phi(\boldsymbol{B})} = 1 + \psi_1 \boldsymbol{B} + \psi_2 \boldsymbol{B}^2 + \cdots \equiv \psi(\boldsymbol{B}) \qquad (2.6a)$$

$$\frac{\phi(\boldsymbol{B})}{\theta(\boldsymbol{B})} = 1 - \pi_1 \boldsymbol{B} - \pi_2 \boldsymbol{B}^2 - \cdots \equiv \pi(\boldsymbol{B}) \qquad (2.6b)$$

The time series outliers are classified into four categories depending on the nature of their impacts on the time series, namely, additive, innovational, level shift and transitory change *outlier* (Tsay, 2005). A*n*additive outlier (AO) represents an isolated spike, a level shift outlier (LSO) a step function, a temporary or transitory change outlier (TCO) a spike that takes a few periods to disappear and an innovational outlier (IO) a shock in the innovations of the model. Depending upon the knowledge of outlier type we need to deal with separate outlier model.

Chen and Liu (1993) used the following model to describe a time series subject to the influence of a non-repetitive event:

$$X_t = Y_t + \omega \frac{A(B)}{G(B)H(B)} \xi_t^T \qquad (2.7)$$

where, X_t denotes the observed series, Y_t follows a general ARMA process described by

$$Y_t = \frac{\theta(\boldsymbol{B})}{\phi(\boldsymbol{B})\alpha(B)} u_t, \quad t = 1, 2, \cdots, n \qquad (2.8)$$

with $\theta(\boldsymbol{B})$, $\phi(\boldsymbol{B})$ and $\alpha(\boldsymbol{B})$ are polynomials of $\boldsymbol{B}$ as defined earlier. ξ_t^T is an indicator function for the occurrence of any outlier such that $\xi_t^T = 1$ if $t = T$ and $\xi_t^T = 0$, otherwise.

ω denotes the magnitude of the outlier and the ratio $\frac{A(B)}{G(B)H(B)}$ represents the dynamic pattern of the outlier effect.

In relation to the outlier model the four types of outlier can be defined as:

IO: $$\frac{A(B)}{G(B)H(B)} = \frac{\theta(B)}{\alpha(B)\phi(B)} \tag{2.9a}$$

AO: $$\frac{A(B)}{G(B)H(B)} = 1 \tag{2.9b}$$

TCO: $$\frac{A(B)}{G(B)H(B)} = \frac{1}{(1-\delta B)} \tag{2.9c}$$

LSO: $$\frac{A(B)}{G(B)H(B)} = \frac{1}{(1-B)} \tag{2.9d}$$

3. Procedures to be compared:

Here we are considering two test procedures for detection of AO and IO in ARMA models using the procedures described below:

3.1. Iterative procedure proposed by Chang et al. (1988):

Considering the time series model, ARIMA (p, d, q) as

$$\left(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p\right)(1-B)^{d_1}(1-B)^{d_1}\left(1-B^s\right)^{d_2} Y_t = \phi_0 + \left(1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q\right)u_t$$

with $d_1 + sd_1 = d$, or writing these in compact form we have

$$\phi(\boldsymbol{B})\alpha(\boldsymbol{B})Y_t = \theta(\boldsymbol{B})u_t, \quad t = 1, 2, \ldots, n \tag{3.1}$$

The two polynomials $\phi(\boldsymbol{B})$ and $\theta(\boldsymbol{B})$have zeros all outside the unit circle. Here the differencing filter $\alpha(\boldsymbol{B})$ renders the data stationary and is an AR polynomial that contains

all its roots on the unit circle. Following Box and Tiao (1975), the exogenous intervention effect on a time series is given by the dynamic model

$$X_t = \frac{A(B)}{G(B)H(B)}\xi_t^T + Y_t \qquad (3.2)$$

where, X_t is the observed series, ξ_t^T indicates the occurrence of the intervention, such that $\xi_t^T = 1$, for $t = T$ and zero otherwise. The ratio $\frac{A(B)}{G(B)H(B)}$ represents the dynamic pattern of the outlier effect or the intervention. The authors focused on two types of outliers additive and innovational (AO and IO). For AO and IO the respective outlier models are

$$X_t = Y_t + \omega\xi_t^T \qquad (3.3a)$$

and

$$X_t = Y_t \frac{\theta(B)}{\phi(B)\alpha(B)}\omega\xi_t^T \qquad (3.3b)$$

In terms of the innovation sequence u_t the models may be written as

$$X_t = \frac{\theta(\boldsymbol{B})}{\phi(\boldsymbol{B})\alpha(\boldsymbol{B})}u_t + \omega\xi_t^T \qquad (3.4a)$$

and

$$X_t = \frac{\theta(\boldsymbol{B})}{\phi(\boldsymbol{B})\alpha(\boldsymbol{B})}\left\{u_t + \omega\xi_t^T\right\} \qquad (3.4b)$$

Situation - I: When ARMA parameters and the σ_u^2 are known

Using the results of the long division of the polynomials let us consider for $e_t = \pi(\boldsymbol{B})X_t$ for $t = 1, 2, \ldots, n$, so that the above expressions can be rewritten as

(AO) $$e_t = \omega\pi(\boldsymbol{B})\ \xi_t^T + u_t \qquad (3.5a)$$

(IO) $$e_t = \omega\xi_t^T + u_t \qquad (3.5b)$$

By least squares estimation the estimated impact of ω of these models are given by

$$\text{(AO)} \quad \hat{\omega}_A = \rho^2\left(1-\pi_1 F-\pi_2 F^2-\cdots-\pi_{n-T}F^{n-T}\right)e_T = \rho^2\pi\left(F\right)=e_T \tag{3.6a}$$

$$\text{(IO)} \quad \hat{\omega}_I = e_T \tag{3.6b}$$

where $\rho^2=\left(1+\pi_1^2+\pi_2^2+\cdots+\pi_{n-T}^2\right)^{-1}$, and F is the forward shift operator such that $Fe_t = e_{t+1}$. The estimated magnitude ($\hat{\omega}_A$) of the additive outlier can be written in terms of the observations as $\hat{\omega}_A = \rho^2\pi(F)\pi(B)X_t$. The variances of these estimates are $var\left(\hat{\omega}_A\right)=\rho^2\sigma_u^2$ and $var\left(\hat{\omega}_I\right)=\sigma_u^2$. The variance of $\hat{\omega}_A$ is at most as large as that of $\hat{\omega}_I$ because $\rho^2 \le 1$, and in some cases it can be much smaller than σ_u^2. Here the null hypothesis stats that $\omega = 0$; against the two alternative hypothesis respectively for AO and IO, i.e.

$$H_0 \; vs \; H_1: \qquad \lambda_{I,T} = \hat{\omega}_I / \sigma_u$$

$$H_0 \; vs \; H_2: \qquad \lambda_{A,T} = \hat{\omega}_A / \rho\sigma_u$$

Under hypothesis H_0, both the statistics $\lambda_{A,T}$ and $\lambda_{I,T}$ have standard normal distribution. Finally for testing the possibility of an AO or an IO, at an unknown position in the series $X_1, X_2, ..., X_n$, the likelihood ratio method leads to the criteria:

$$\text{(IO)} \qquad \max_t = 1, 2, \ldots, n\left|\lambda_{I,t}\right|$$

$$\text{(AO)} \qquad \max_t = 1, 2, \ldots, n\left|\lambda_{A,t}\right|$$

Situation -II: When ARMA parameters and the σ_u^2 are unknown

In the unknown event, the estimates of the ARMA parameters and variance of the intervention under the IO or AO case, can be obtained by maximizing the likelihood function $(\phi_1,\ldots,\phi_p\theta_1,\ldots,\theta_q,\omega,\phi_1)$. If $\hat{\phi}_1,\ldots,\hat{\phi}_p;\hat{\theta}_1,\ldots,\hat{\theta}_q$ and $\hat{\sigma}_u^2$ are the maximum likelihood estimate of the parameters $\phi_1,\ldots,\phi_p\theta_1,\ldots,\theta_q$ and σ_u^2 respectively, then the above statistics becomes $\hat{\lambda}_{I,T} = \hat{\omega}_I/\hat{\sigma}_u; \hat{\lambda}_{A,T} = \hat{\omega}_A/\hat{\rho}\hat{\sigma}_u$.

In general the possibility of an IO or AO in the series is tested by

$$\hat{\eta}_{IO} = \max_{t=1,2,\ldots,n} \left|\lambda_{I,t}\right| > C \tag{3.7a}$$

$$\hat{\eta}_{AO} = \max_{t=1,2,\ldots,n} \left|\lambda_{A,t}\right| > C \tag{3.7b}$$

where C is a positive constant, suitably chosen. In practical applications the author recommended to use $C = 3.0$ for high sensitivity, $C = 3.5$ for medium sensitivity, and $C = 4.0$ for low sensitivity, when $n < 200$.

3.2. Iterative Procedure Proposed by Chen and Liu (1993):

Considering the above model $\phi(\boldsymbol{B})\alpha(\boldsymbol{B})Y_t = \theta(\boldsymbol{B})u_t$; $t = 1, 2, \ldots, n$, Chen and Liu (1993) tried to jointly estimate the model parameters and outlier effects in the series. Using the polynomial $\pi(\boldsymbol{B})$, the estimated residuals of the assumed model can be expressed as $\hat{e}_t = \pi(\boldsymbol{B})X_t$ for $t = 1, 2, \ldots,$. Suppose ω is the magnitude of the outlier and δ be the parameter designed to model the pace of the dynamic dampening effect, then the four types of outlier can be modelled as:

(IO) $$\hat{e}_t = \omega\xi_t^T + u_t \tag{3.8a}$$

(AO) $$\hat{e}_t = \omega\pi(\boldsymbol{B})\xi_t^T + u_t \tag{3.8b}$$

(TCO) $$\hat{e}_t = \omega\left\{\frac{\pi(\boldsymbol{B})}{(1-\delta\boldsymbol{B})}\right\}\xi_t^T + u_t \tag{3.8c}$$

(LSO) $$\hat{e}_t = \omega\left\{\frac{\pi(\boldsymbol{B})}{(1-\boldsymbol{B})}\right\}\xi_t^T + u_t \tag{3.8d}$$

These four equations can be expressed as

$$\hat{e}_t = \omega z_{it} + u_t \quad t = T, T+1, \ldots, n \tag{3.9}$$

Here $z_{it} = 0$ for all i and $t < T$

$z_{it} = 1$ for all i and $l \geq 1$,

For the effect of a single outlier at $t = T$, the least squares estimate are

$$\hat{\omega}_I(T) = \hat{e}_T; \hat{\omega}_A(T) = \frac{\sum_{t=T}^{n} \hat{e}_T z_{2t}}{\sum_{t=T}^{n} z_{2t}^2};\ \hat{\omega}_L(T) = \frac{\sum_{t=T}^{n} \hat{e}_T z_{3t}}{\sum_{t=T}^{n} z_{3t}^2} \text{ and } \hat{\omega}_T(T) = \frac{\sum_{t=T}^{n} \hat{e}_T z_{4t}}{\sum_{t=T}^{n} z_{4t}^2}$$

For the last observation, it becomes impossible to distinguish empirically the type of outlier at the very end of a series because for $T = n$; $\hat{\omega}_I(n) = \hat{\omega}_A(n) = \hat{\omega}_L(n) = \hat{\omega}_T(n)$. Accordingly the four standardized statistics are designed as:

$$\hat{\lambda}_I(T) = \frac{\hat{\omega}_I(T)}{\hat{\sigma}_u};\ \hat{\lambda}_A(T) = \left\{\frac{\hat{\omega}_A(T)}{\hat{\sigma}_u}\right\}\left(\sum_{t=T}^{n} z_{2t}^2\right)^{\frac{1}{2}};\ \hat{\lambda}_L(T) = \left\{\frac{\hat{\omega}_L(T)}{\hat{\sigma}_u}\right\}\left(\sum_{t=T}^{n} z_{3t}^2\right)^{\frac{1}{2}};\ \hat{\lambda}_T(T) = \left\{\frac{\hat{\omega}_T(T)}{\hat{\sigma}_u}\right\}\left(\sum_{t=T}^{n} z_{4t}^2\right)^{\frac{1}{2}}$$

All these four standardized statistics follow an approximately normal distribution.

4. Simulation Study:

Here we discuss two procedures of identifying and detecting outliers in a time series. The procedures of Chang et al. and Chen and Liu are named as I and II respectively. For this we are simulating three sets of data for each procedures, i.e, data from pure AR, pure MA and ARMA model. Then we examine the detection of deliberately planted outlier at various positions of the generated series. The magnitude of the planted outliers are incorporated as $\omega = a * \hat{\sigma}_u$, *a being a constant takes non-negative values*. Then the values of the test statistics $\lambda_{i,T}$; $i = AO, IO$, for those particular locations are compared with the threshold value 3.5 and 4.0. And these process is repeated for a good number of times. Different combinations of the AR and MA coefficients are used to evaluate the power. Also the determination of outliers can be sensitive to the estimate of residual standard deviation $\hat{\sigma}_u$. For a series contaminated by outlier the $\hat{\sigma}_u$ may be overestimated if the usual sample standard deviation is used. So here we have estimated $\hat{\sigma}_u$ by mean absolute deviation (MAD) method for the iterations.

In the Tables 1-2 the magnitude ω *are kept fixed for* all combinations, while in the subsequent tables $\omega = 1.0, 1.5, 2.5, 3.0, 4.5, 6.0, 7.0$ *were taken to* assess the performance for small or large shift. The following tables and figures shows the simulation result.

Table-1: Proportion of rejection in presence of two outliers when $k = 2$, $n = 100$

(p, d, q)	*AR coefficients* (ϕ)	*MA coefficients* (θ)	*Procedure I*		*Procedure II*	
			Additive	*Innovational*	*Additive*	*Innovational*
(1,0,0)	*0.10*	--	*0.267*	*0.897*	*0.870*	*0.847*
	0.50	--	*0.021*	*0.630*	*0.613*	*0.447*
	0.90	--	*0.049*	*0.503*	*0.723*	*0.377*
	-0.50	--	*0.998*	*0.994*	*0.991*	*0.990*
	-0.90	--	*1.000*	*0.986*	*1.000*	*1.000*
(0,0,1)	--	*0.10*	*0.076*	*0.837*	*0.862*	*0.841*
	--	*0.50*	*0.002*	*0.448*	*0.423*	*0.304*
	--	*0.88*	*0.000*	*0.500*	*0.072*	*0.039*
	--	*-0.50*	*0.988*	*0.985*	*0.992*	*0.992*
	--	*-0.88*	*0.994*	*0.969*	*0.935*	*0.935*
(1,0,1)	*0.10*	*0.10*	*0.008*	*0.787*	*0.788*	*0.734*
	0.50	*0.50*	*0.000*	*0.500*	*0.172*	*0.090*
	0.90	*0.88*	*0.000*	*0.129*	*0.059*	*0.288*
	-0.50	*-0..50*	*1.000*	*0.512*	*0.987*	*0.987*
	-0.90	*-0.88*	*1.000*	*0.838*	*1.000*	*1.000*

Table-2: Proportion of rejection in presence of three outliers when $k = 2$, $n = 100$

(p, d, q)	*AR coefficients*	*MA coefficients*	*Procedure I*		*Procedure II*	
	(ϕ)	(θ)	*Additive*	*Innovational*	*Additive*	*Innovational*
	0.10	--	*0.004*	*0.655*	*0.432*	*0.656*
(1, 0, 0)	*0.50*	--	*0.007*	*0.556*	*0.389*	*0.468*
	0.90	--	*0.039*	*0.499*	*0.428*	*0.429*
	-0.50	--	*0.049*	*0.897*	*0.914*	*0.928*
	-0.90	--	*0.999*	*0.874*	*0.740*	*0.725*
	--	*0.10*	*0.023*	*0.604*	*0.423*	*0.657*
(0, 0, 1)	--	*0.50*	*0.190*	*0.126*	*0.423*	*0.506*
	--	*0.88*	*0.497*	*0.124*	*0.450*	*0.463*
	--	*-0.50*	*0.165*	*0.899*	*0.940*	*0.964*
	--	*-0.88*	*0.532*	*0.832*	*0.928*	*0.929*
	0.10	*0.10*	*0.002*	*0.119*	*0.405*	*0.609*
(1, 0, 1)	*0.50*	*0.50*	*0.450*	*0.496*	*0.387*	*0.379*
	0.90	*0.88*	*0.509*	*0.295*	*0.460*	*0.432*
	-0.50	*-0..50*	*0.554*	*0.327*	*0.731*	*0.825*
	-0.90	*-0.88*	*0.624*	*0.748*	*0.834*	*0.835*

Table-3: Proportion of rejection in presence of two additive outliers.

(ϕ_1, θ_1)	(n, ω)	*Procedure I*		*Procedure II*	
		***cval* = 3.5**	***cval* = 4.0**	***cval* = 3.5**	***cval* = 4.0**
(-0.90, -0.88)	*(100, 1.0)*	*0.000*	*0.000*	*0.125*	*0.040*
(-0.90, -0.88)	*(100, 1.5)*	*0.000*	*0.000*	*0.266*	*0.105*
(-0.90, -0.88)	*(100, 2.5)*	*0.374*	*0.121*	*0.520*	*0.275*
(-0.90, -0.88)	*(100, 3.0)*	*0.493*	*0.184*	*0.855*	*0.625*
(-0.90, -0.88)	*(100, 4.5)*	*0.729*	*0.406*	*0.910*	*0.761*
(-0.90, -0.88)	*(100, 6.0)*	*0.847*	*0.566*	*0.925*	*0.835*
(-0.90, -0.88)	*(100, 7.0)*	*0.899*	*0.681*	*0.960*	*0.880*
(-0.90, -0.88)	*(200, 3.0)*	*0.960*	*0.890*	*0.965*	*0.870*
(-0.90, -0.88)	*(200, 4.5)*	*0.995*	*0.983*	*1.000*	*1.000*
(-0.90, -0.88)	*(200, 6.0)*	*0.998*	*0.996*	*1.000*	*1.000*
(-0.90, -0.88)	*(200, 7.0)*	*1.000*	*0.999*	*1.000*	*1.000*
(-0.90, 0.50)	*(200, 3.0)*	*0.361*	*0.224*	*0.465*	*0.265*
(-0.90, 0.50)	*(200, 4.5)*	*0.784*	*0.677*	*0.882*	*0.744*
(-0.90, 0.50)	*(200, 6.0)*	*0.714*	*0.606*	*0.985*	*0.936*
(-0.90, 0.50)	*(200, 7.0)*	*0.995*	*0.984*	*0.991*	*0.950*
(0.50, -0.88)	*(200, 3.0)*	*0.023*	*0.012*	*0.325*	*0.197*
(0.50, -0.88)	*(200, 4.5)*	*0.267*	*0.187*	*0.620*	*0.435*
(0.50, -0.88)	*(200, 6.0)*	*0.485*	*0.384*	*0.910*	*0.801*
(0.50, -0.88)	*(200, 7.0)*	*0.634*	*0.529*	*0.950*	*0.895*

Table-4: Proportion of rejection in presence of two innovative outliers.

(ϕ_1, θ_1)	(n,ω)	*Procedure I*		*Procedure II*	
		***cval* = 3.5**	***cval* = 4.0**	***cval* = 3.5**	***cval* = 4.0**
(-0.90, -0.88)	*(100, 1.0)*	*0.000*	*0.000*	*0.180*	*0.045*
(-0.90, -0.88)	*(100, 1.5)*	*0.000*	*0.000*	*0.335*	*0.170*
(-0.90, -0.88)	*(100, 2.5)*	*0.002*	*0.000*	*0.696*	*0.410*
(-0.90, -0.88)	*(100, 3.0)*	*0.025*	*0.012*	*0.790*	*0.460*
(-0.90, -0.88)	*(100, 4.5)*	*0.359*	*0.226*	*0.945*	*0.555*
(-0.90, -0.88)	*(100, 6.0)*	*0.481*	*0.331*	*0.960*	*0.650*
(-0.90, -0.88)	*(100, 7.0)*	*0.543*	*0.384*	*0.985*	*0.691*
(-0.90, -0.88)	*(200, 3.0)*	*0.385*	*0.317*	*0.950*	*0.510*
(-0.90, -0.88)	*(200, 4.5)*	*0.546*	*0.491*	*0.995*	*0.630*
(-0.90, -0.88)	*(200, 6.0)*	*0.661*	*0.587*	*1.000*	*0.715*
(-0.90, -0.88)	*(200, 7.0)*	*0.749*	*0.664*	*1.000*	*0.760*
(-0.90, 0.50)	*(200, 3.0)*	*0.139*	*0.091*	*0.695*	*0.525*
(-0.90, 0.50)	*(200, 4.5)*	*0.725*	*0.583*	*0.946*	*0.759*
(-0.90, 0.50)	*(200, 6.0)*	*0.943*	*0.879*	*1.000*	*0.944*
(-0.90, 0.50)	*(200, 7.0)*	*0.736*	*0.719*	*1.000*	*0.961*
(0.50, -0.88)	*(200, 3.0)*	*0.006*	*0.003*	*0.680*	*0.551*
(0.50, -0.88)	*(200, 4.5)*	*0.518*	*0.360*	*0.833*	*0.715*
(0.50, -0.88)	*(200, 6.0)*	*0.760*	*0.652*	*0.955*	*0.913*
(0.50, -0.88)	*(200, 7.0)*	*0.872*	*0.780*	*0.990*	*0.985*

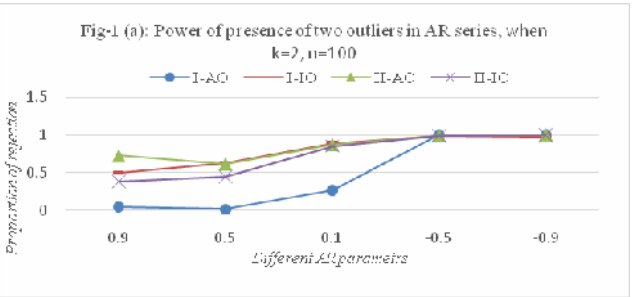


Fig-1 (a): Power of presence of two outliers in AR series, when k=2, n=100
I-AO
I-IO
II-AO
II-IO
1.5
1
0.5
0
0.9
0.5
0.1
-0.5
-0.9
Different AR parameters

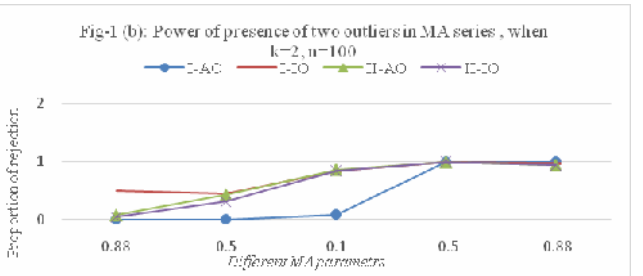


Fig-1 (b): Power of presence of two outliers in MA series , when k=2, n=100
I-AO
I-IO
II-AO
II-IO
2
1
0
0.88
0.5
0.1
0.5
0.88
Different MA parameters

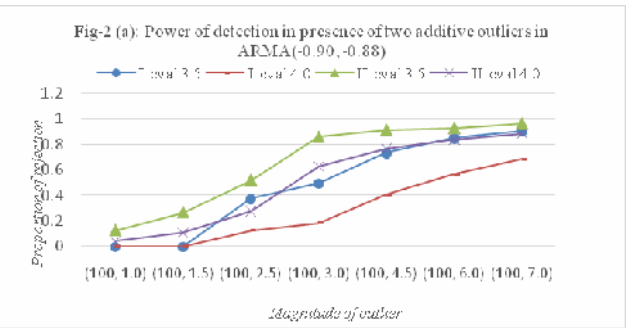


Fig-2 (a): Power of detection in presence of two additive outliers in ARMA(-0.90, -0.88)
1.2
1
0.8
0.6
0.4
0.2
0
(100, 1.0)
(100, 1.5)
(100, 2.5)
(100, 3.0)
(100, 4.5)
(100, 6.0)
(100, 7.0)
Magnitude of outlier

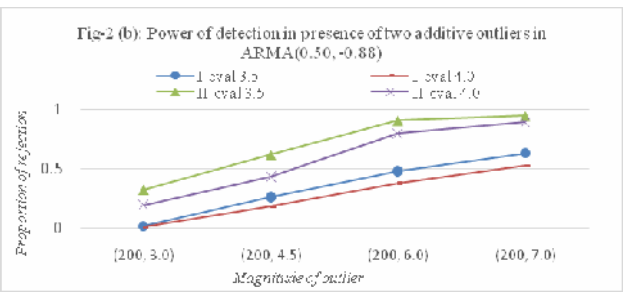
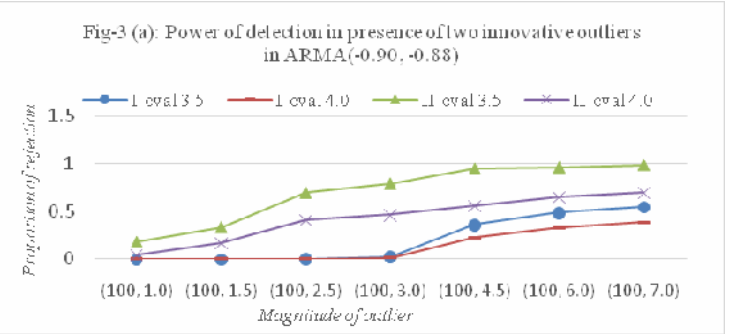
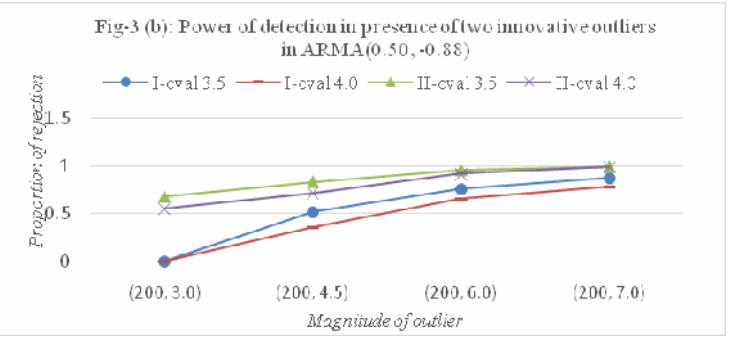


Fig-2 (b): Power of detection in presence of two additive outliers in ARMA(0.50, -0.88)
1
0.5
0
(200, 3.0)
(200, 4.5)
(200, 6.0)
(200, 7.0)
Magnitude of outlier





5. Results and discussion:

From the Table and Figures, we observe that for ARMA (1,0,0) with $\phi_1 = 0.10$, the procedure II detects more additive outlier (AO), while procedure I detects more innovational outlier (IO). Although for smaller value of ϕ_1 both procedures perform satisfactorily, the procedure-I possess higher power than the procedure-II for both type of outliers.

For ARMA (0,0,1) with small positive value of θ_1, the procedure II detects more both type of outliers, whereas with large positive value of θ_1 procedure II detects more IO. For large negative value of θ_1, both procedures perform satisfactorily. However, the procedure-I possess higher power than the procedure-II for large negative value of moving average coefficient.

For ARMA (1,0,1) with small positive AR and MA coefficients ($\phi_1 = 0.10$, $\theta_1 = 0.10$), the procedure II detects more AO but less IO than procedure. While for large negative value of AR and MA coefficients ($\phi_1 = -0.90$, $\theta_1 = -0.88$) both the procedure detects fairly high proportion of outliers.

From the two Tables 1-2 we see that both the procedures suffer from masking effect. Considering additive outlier, for ARMA (-0.90, -0.88) with $n = 100$, ω=7.0, the power of procedure I changes from 1.000 to 0.624 and that of procedure II changes from 1.000 to 0.834. Similarly, considering additive outlier, forARMA (-0.90, -0.88) with $n = 100$, $\omega = 7.0$, the power of procedure I changes from 0.838 to 0.748 and that of procedure II changes from 1.000 to 0.835.

Since both procedures can detect a fairly good proportion of outlier, so we next proceed to examine which one detects better for small or large shifts in the series. From Tables 3-4 and Figures 5.4(a), 5.4 (b); it is observed that for ARMA (-0.90, -0.88) with $n = 100$, $\omega = 1.0, 1.5, 2.5, 3.0, 4.5, 6.0, 7.0$, the procedure-II can detect more additive or innovative outlier than procedure-I. Also from even for small value of ω the procedure-II is better.

For ARMA (0.50, -0.88) with $n = 200$, $\omega = 3.0, 4.5, 6.0, 7.0$, the procedure-II can detect more additive or innovative outlier than procedure-I. Moreover for small value of ω there is a large difference in empirical power between procedure-I and procedure-II.

As expected the detection proportion reduces when the threshold value is increased for the test statistics. The type I error was not also controlled in both the cases. So the use of the critical value needs proper attention in locating the outliers.

Example: To apply the above conspired methods, we generate an ARMA (2, 0, 2) series with $n = 200$ observations. Then three additive outliers was deliberately incorporated at the locations $T = 10, 11, 12$ and three innovative outliers at the locations $T = 25, 26, 27$. Then, taking the threshold value of test statistics as 3.5, and applying the test of Chang et al. (1988), we observed the following results:

Time point	Value of Test Statistic	Type
10	4.220	AO
12	3.689	AO
25	4.972	IO
26	7.487	IO
27	-10.724	IO

After omitting these five points the fitted model yields the following form:

Coefficients				$\hat{\sigma}_u$
ϕ_1	ϕ_2	θ_1	θ_2	
1.2301	-0.5294	-0.6250	0.0406	1.1264
(0.2336)	(0.1028)	(0.2501)	(0.1536)	

Whereas applying the test of Chen and Liu (1993), we observed the following results:

Time point	Value of Test Statistic	Type
10	5.074	AO
11	3.631	AO
12	4.644	AO
25	5.969	IO
26	9.124	IO
27	12.560	IO

After omitting these six points the fitted model yields the following form:

Coefficients				$\hat{\sigma}_u$
ϕ_1	ϕ_2	θ_1	θ_2	
1.3255	-0.5851	-0.6181	-0.0347	0.9816
(0.1445)	(0.0848)	(0.1629)	(0.1303)	

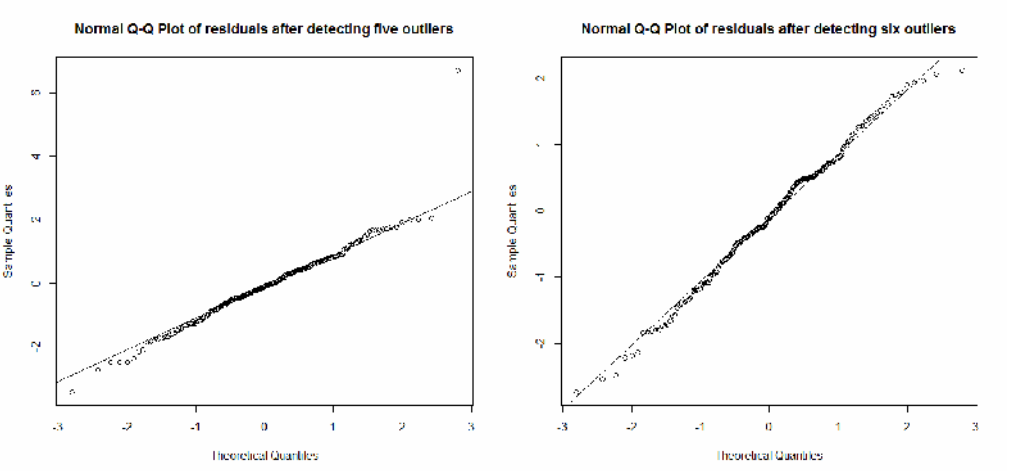


Fig-4. Normal Q-Q plot of the fitted residuals

The procedure proposed by Chang et al. (1988) was unable to label outlier in the 11th time point. But in the presence of this single outlier inflates the estimate of residual standard deviation from 0.9816 to 1.1264. Also there is a reduction of standard error of the coefficients. To inspect whether the residuals follow white noise process or not, we used the normal quantile-quantile plots as shown in Fig-4. These plot clearly reveals the effect of the outlier which was not detected by the first procedure on the series. Thus it is clear that the procedure proposed by Chen and Liu (1993) can detect more outlier.

6. Conclusion:

The simulation results shows that the detection proportion of procedure-II, i.e., the procedure proposed by Chen and Liu (1993) is always higher than procedure-I i.e., the procedure proposed by Chang et al.,(1988) for both types of outliers. Though both the procedures are effected by masking, the second procedure can detect small shifts in the time series data. The example of simulated data set also reveals the same results.

Acknowledgement: Authors are to referee for his valuable comment on the paper.

References:

1. Zhang, G. P., (2003): 'Time series forecasting using a hybrid ARIMA and neural network model', Neurocomputing, 50, pp.159 - 175.
2. 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, pp. 70-79.
3. Grubbs, F. E., (1969): 'Procedures for detecting outlying observations in samples', Technometrics, 11, pp. 1-21.
4. Tsay, R. S., (1988): 'Outliers, level shifts and variance changes in time series', J. Forecast., 7, pp. 1-20.
5. Tsay, R. S., (1986): 'Time series model specification in the presence of outliers', J. Am. Statist. Ass., 81, pp.132-141.
6. Tsay, R. S., (2005): 'Analysis of financial time series', Wiley-Interscience.
7. Chang, I., Tiao, G. C. and Chen, C. (1988): 'Estimation of time series parameters in the presence of outliers', Technometrics, 30, pp.193-204.
8. Chen, C. and Liu, L.M., (1993): 'Joint Estimation of Model Parameters and Outlier Effects in Time Series', Journal of the American Statistical Association, Vol. 88, No. 421 , pp. 284-297.
9. Gujarati, D., Porter, D. and Gunasekar, S., (2011): 'Basic Econometrics', McGraw Hill Education (India) Private Limited.