Journal of Statistical
Theory & Methods
1995, Vol. 6, pp 41~50

Joint Estimation of the Outliers Effect and the Model Parameters in ARMA Process

Kwang Ho Lee[1] and Hye Jung Shin[1]

Abstract In this paper, an iterative procedure, which detects the location of the outliers and the joint estimates of the outliers effects and the model parameters in the autoregressive moving average model with two types of outliers, is proposed. The performance of the procedure is compared with the one in Chen and Liu(1993) through the Monte Carlo simulation. The proposed procedure is very robust in the sense that applies the procedures to the stationary time series model with any types of outliers.

1. Introduction

In time series analysis, observations are often subject to the influence of nonrepetitive exogenous intervention - for example, strikes, outbreak of wars, sudden change in the market structure of a commodity, unexpected changes if certain conditions on physical system, and so forth - and as a results some observations become outliers. Outlier detection and the development of method of parameter estimation insensitive to the presence of outliers are important in statistical practice.

Several authors including Fox(1972), Abraham and Box(1979), Tsay(1986), Chang et al.(1988), Chen and Liu(1993), Abraham and Chuang(1993) have considered the problems that detects the location of the outliers and estimates the model parameters and the outlier effects. A common approach to deal with outliers in time series analysis is to identify the location and the types of outliers and then apply the intervention models discussed in Box and Tiao(1975). It is known that the iterative procedure proposed by Chang et al.(1988) is quite effective in the problem that detects the locations and estimates the effects of the outliers.

Dempster et al.(1977) suggested the Expection-maximization(EM) algorithm, a

[1] Department of Statistics, Yeungnam University, Kyungbuk, Republic of Korea.

method of iterative searching for the maximum likelihood estimate(MLE) in the context of incomplete data. Abranham and Chung(1993) developed an estimation method which was devoted implementation of the EM algorithm in the context of time series models with outliers. Chen and Liu(1993) used also an iterative procedure to obtain the joint estimates of model parameters and outlier effects. They considered four types of outliers and issues of spurious and masking effects were disscussed.

In this paper, we proposes an iterative procedure to detect the location of the outliers and to obtain the joint estimates of the model parameters and the outliers effects in the autoregressive moving average model (ARMA) with two types of outliers. We will show that the procedure performs well in terms of detecting outliers and estimating the model parameters jointly, and compare the performance of the proposed prpcedure with the ones of the conditional least square method and with the ones of the procedure proposed by Chen and Liu(1993) under the considering model through the Monte Carlo simulation method.

In Section 2, two types of outliers are defined.Because of most of all outliers occuring in the time series analysis are represented by the linear combination of these two types of outliers, the two types of the outliers are considered in the paper. To detect the locations of the outliers and to estimate the model parameters and the outliers effects jointly, an iterative procedure is proposed in Section 3. In Section 4, The performance of the proposed procedure will be compared with the conditional least square procedure and with the one given by Chen and Liu (1993) through the Monte Carlo Simulation method. Finally some remarks and conclusions will be given in secion 5.

2. Types of Outliers

In this paper the proposed procedure may be applied to the stationary ARMA process. But to simplify the presentation, we consider only AR(p) process.

Let $\{Z_t\}$ be a stationary AR process with order p defined by

$$\phi(B)Z_t = a_t, \qquad t = 0, \pm 1, \pm 2, \cdots, \tag{2.1}$$

where $\phi(B)$ is $-\phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and B is backshift operator such that $Bz_i = z_{t-1}$. The roots of $\phi(B)$ are outside the unit circle, and a_t is a sequence of independent and identically distributed normal random variables with mean zero and variance σ^2.

Here we introduce two types of outliers which are often occurred in the time series process. The one is innovational outlier and the other is addtive outlier. These types of outliers are discussed by Fox(1972) and Abraham and Box(1979).

Suppose that an outlier occurs at $t=T$ in the given time seires process $\{Z_t\}$. Define a new process $\{Y_t\}$ as follows :

$$Y_t = Z_t + \delta A(B)\xi_t \tag{2.2}$$

where $\{Z_t\}$ follows an AR(p) process described in (2.1), δ represents the magnitude of the outliers in the process $\{Z_t\}$. And ξ_t is a time indicator signifying the occurrence of the ourlier which is defined as

$$\xi_t = \begin{cases} 0, & t \neq T, \\ 1, & t = T, \end{cases}$$

A(B) in (2.2) denotes the dynamic pattern of signifying the occurence of the outlier. That is, if $A(B) = 1/\phi(B)$ then (2.2) can be representd as follow

$$\phi(B)Y_t = a_t + \delta\xi_t. \tag{2.3}$$

If $A(B)=1$ then (2.2) can be represented as

$$Y_t = Z_t + \delta\xi_t. \tag{2.4}$$

In the case of $A(B)=1/\phi(B)$, the time series process $\{Y_t\}$ is called an innovational outliers (IO) model and in the case of $A(B)=1$, we refer to $\{Y_t\}$ as an additive outliers (AO) model. In general, most of all outliers in the time series data are represented as the linear combimation of these two types of outliers. Thus we only consider these two types of ourliers in this paper.

3. Joint Estimation of the Patameters and the Outlier Effect

Suppose that $Y=(y_1,\cdots,y_n)'$ is a vector of observations and each y_t is associated with one of two unobserable states(the outlier state and the non-outlier state). Thus there is an unobserable indicator vector $\xi = (\xi_1,\xi_2,\cdots,\xi_n)'$ whrer ξ_t is as defined in section 2. In fact, $\{\xi_t\}$ is considered as a sequence of Bernoulli ramdom variables with

$$P(\xi_t = 1) = \alpha, \quad P(\xi_t = 0) = 1-\alpha, \quad t = p+1,\cdots,n.$$

where, α is small prior probability. Further, assume that $\xi_1 = \xi_2 = \cdots = \xi_p = 0$, that is, the first p observations are not ourliers. For the notational conveniences let $\phi = (\phi_1,\phi_2,\cdots,\phi_p)'$, $Y_0 = (y_1,y_2,\cdots,y_p)'$, $Y = (y_{p+1},y_{p+2},\cdots,y_n)'$, $r = \sum_{t=p+1}^{n}\xi_t$, $\xi = (\xi_{p+1},\xi_{p+2},\cdots,\xi_n)'$.

$$V = \begin{pmatrix} y_p & y_{p-1} & \cdots & y_1 \\ y_{p+1} & y_p & \cdots & y_2 \\ & & \vdots & \\ y_{n-1} & y_{n-2} & \cdots & y_{n-p} \end{pmatrix}$$

and

$$U = \begin{pmatrix} \xi_p & \xi_{p-1} & \cdots & \xi_1 \\ \xi_{p+1} & \xi_p & \cdots & \xi_2 \\ & & \vdots & \\ \xi_{n-1} & \xi_{n-2} & \cdots & \xi_{n-p} \end{pmatrix}$$

Let β represent a vector of parameters. Y is refered to as the incomplete data and $X=(Y_0,Y,\xi)$ as the complete data. We apply the basis idea behind the EM algorithm that is to maximize the incomplete data likelihood $L(Y,\beta)$ by maximizing the conditional expectation of the complete data likilihood $L(X,\beta)$ given the incomplete data Y in each iteration.

The complete data likelihood given the initial values $z_1,z_2,\cdots,z_p$ is equivalent to that given by $Y_0=(y_1,y_2,\cdots,y_p)'$, can be expressed by

$$\begin{aligned} L(X,\beta,\alpha|Y_0) &= f(Y,\xi|\beta,\alpha,Y_0) \\ &= f(Y|\xi,\beta,\alpha,Y_0)f(\xi|\alpha,Y_0) \end{aligned} \tag{3.1}$$

where $\beta'=(\alpha',\sigma^2,\delta)'$. Here we are considering only the likelihood distributional on the initial valued Y_0. Since distribution of ξ is not dependent of β, (3.1) can be expressed as

$$\begin{aligned} L(X,\beta,\alpha|Y_0) &= f(Y,\xi|\beta,\alpha,Y_0) \\ &\propto \sigma^{-(N-P)}\exp[-(2\sigma^2)^{-1}\{S(\phi)-2\delta D(\phi) \\ &\quad +\delta^2 C(\phi)\}]\alpha^r(1-\alpha)^{n-p-r}. \end{aligned} \tag{3.2}$$

where $S(\phi)=(Y-V\phi)'(Y-V\phi), C(\phi)=(\xi-U\phi)'(\xi-U\phi), D(\phi)=(Y-V\phi)'(\xi-U\phi)$.

Detection Step

From the E step of the EM algorithm we can approximately estimate the parameters α and β, say, $\hat{\alpha}$ and $\hat{\beta}$ respectively. Since the ξ_ts are Bernoulli trials, the expectation of ξ_t, given $(Y_0,\hat{\beta},\hat{\alpha})$ and Y at the (k+1)th iteration is in general

$$\hat{\xi}_t^{(k+1)} = E(\xi_t|Y_0,Y,\hat{\beta},\hat{\alpha}),$$

where the expectation is taken over the joint distribution of given $G = (Y_0, \beta, \hat{\alpha})$ and Y. Hence

$$\hat{\xi}_t^{(k+1)} = \frac{\sum_{\xi} \xi_t f(Y,\xi|G)}{\sum_{\xi} f(Y,\xi|G)} = \frac{\sum_{\xi_{(t)}} f(Y,\xi,\xi_t,\xi_t = 1|G)}{\sum_{\xi} f(Y,\xi|G)}$$

where $\xi_{(t)}$ is the vector without the element ξ_t and $f(Y,\xi|G)$ is as given in (3.2) with (β,α) replaced by $(\beta^{(k)},\hat{\alpha}^{(k)})$. Thus we take

$$\hat{\xi}_t^{(k+1)} \cong \frac{\sum_{\xi_t} f(Y,\xi_*,\xi_p = \hat{\xi}_p^{(k)},\xi_t = 1|G)}{\sum_{\xi} f(Y,\xi_*,\xi_p = \hat{\xi}_p^{(k)},\xi_t = 1|G) + \sum_{\xi} f(Y,\xi_*,\xi_p = \hat{\xi}_p^{(k)},\xi = 0|G)}$$

where $\xi_p = (\xi_{t-p},\cdots,\xi_{t-1},\xi_{t+1},\cdots,\xi_{t+p})$ and $\xi_* = (\xi_{p+1},\cdots,\xi_{t-p+1},\xi_{t+p+1},\cdots,\xi_n)$. It can be simplified as

$$\hat{\xi}_t^{(k+1)} \cong (1+h_t)^{-1}, \qquad t = p+1,\cdots,n, \tag{3.3}$$

where

$$h_t = \frac{1-\hat{\alpha}^{(k)}}{\hat{\alpha}^{(k)}} \exp\left[\frac{\hat{\eta}^{(k)}\hat{\delta}^{(k)}}{\sigma^{2(k)}}\left\{\hat{\delta}^{(k)}\left(\frac{1}{2} - \bar{\xi}_t^{(k)}\right) - \left(y_t - \hat{y}_t^{(k)}\right)\right\}\right]$$

$$\hat{\eta}^{(k)} = 1 + \hat{\phi}_1^{(k)2} + \cdots + \hat{\phi}_p^{(k)2},$$

$$\bar{\xi}_t^{(k)} = \sum_{j=1}^{p} \hat{\phi}_j^{(k)}\left(\hat{\xi}_{t-j}^{(k)} + \hat{\xi}_{t+j}^{(k)}\right),$$

$$\hat{y}_t^{(k)} = \sum_{j=1}^{p} \hat{\phi}_j^{(k)}\left(y_{t-j} + y_{t+j}\right),$$

$$\hat{\phi}_j^{(k)} = \frac{\hat{\phi}_j^{(k)} - \sum_{i=1}^{p-j} \hat{\phi}_i^{(k)}\hat{\phi}_{i+j}^{(k)}}{\hat{\eta}^{(k)}},$$

$$\hat{\alpha}^{(k)} = \frac{\sum_{t+p+1}^{n} \hat{\xi}_t^{(k)}}{n-p}, \qquad j = 1,2,\cdots,p.$$

We obtain ξ_t that the value means outlier exist or not at time point t. So we can detect the outliers by the values of ξ_t's.

Estimation Step

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

$$Y_t^* = \sum_{j=1}^{m} \delta_j L_j(B) \xi_{tj} + \frac{1}{\phi(B)} a_t, \tag{3.4}$$

where $L_j(B) = 1/\phi(B)$ or $L_j(B) = 1$. Without distinguishing notations of the estimated and the true paramaters, the residuals $\{\hat{e}_t\}$ by fitting an AR(p) model to Y_t^* may be expressed as

$$\begin{aligned} \hat{e}_t &= \sum_{j=1}^{m} \delta_j \phi(B) L_j(B) \xi_{tj} + t, \\ \hat{\delta}_t &= \hat{e}_t \end{aligned} \tag{3.5}$$

If the effects of outliers and their locations are available, then we can adjust the outlier effects based on equation (3.4) and subsequence estimate the model parameters. On the other hand, when the model patameters are known, we can identify outliers and estimate their effects using equation (3.5). It is difficult, if not impossible, to acheive our stated goals in a single step.

Now we are led to the following procedure to handle situation in which there may exist an unknown number of IO's or AO's. The procedure begins with modeling the original series $\{Y_t\}$ by supposing that there are no outliers. Then the outlier detection step and the parameter estimation step will be alternatively followed.

STEP 1. Compute the maximum likelihood estimates of model parameters based on the original and obtain ξ_t by (3.3). Then we identify the outliers location in the step.

STEP 2. Suppose that m time points $t_1, t_2, \cdots, t_m$ are identified as outliers locations. The outliers effects δ_j's can be estimates jointly using the multiple regression model described in (3.5), where $\{\hat{e}_t\}$ is regarded as the output variable and $L_j(B)t_i\}$ are the input variables.

STEP 3. Obtain the adjusted series by removing the ourliers effects, using most recent estimates of δ_j's at step 2. Compute the maximum likelihood estimates of the model parameters based on the adjusted series.

4. Monte Carlo Simulation

In this section, we investigate the performance of the conditional least square estimates (CLSE), Chen & Liu's estimates (C&LE), and the proposed estimates (PPE) through the Monte Carlo simulation.

Three time series data with outliers were generated using the following AR(1) model,

$$(a) y_t = z_t + \delta\xi_t \quad \text{and} \quad z_t = \phi z_{t-1} + a_t, \quad \text{(AO)}$$
$$(b) y_t = \phi y_{t-1}(a_t + \delta\xi_t), \quad \text{(IO)}$$
$$(c) y_t = z_t + A(B)\delta\xi_t, \quad \text{(Mixed)}$$

where

$$\xi_t = \begin{cases} 0, & t \neq t_1, t_2, \cdots, t_m, \\ 1, & t = t_1, t_2, \cdots, t_m, \end{cases}$$
$$A(B) = \begin{cases} 1/(1-\phi B), & \text{if } y_t \text{ is IO}, \\ 1, & \text{if } y_t \text{ is AO}, \end{cases}$$

and $\{a_t\}$ is a white noise sequence with mean zero and variance one.

In a case of AR(2) model, three time series with outliers each was generated using the following models,

$$(a) y_t = z_t + \delta\xi_t \quad \text{and}$$
$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} a_t, \quad \text{(AO)}$$
$$(b) y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2}(a_t + \delta\xi_t), \quad \text{(IO)}$$
$$(c) y_t = z_t + A(B)\delta\xi_t, \quad \text{(Mixed)}$$

where ξ_t and $\{a_t\}$ are same as AR(1) and

$$A(B) = \begin{cases} 1/(1-\phi_1 B - \phi_2 B^2), & \text{if } y_t \text{ is IO}, \\ 1, & \text{if } y_t \text{ is AO}, \end{cases}$$

The same sequence $\{a_t\}$ is used for each of the three models. It should be also noted that some initial observations were already discarded to aviod transient starting effects.

At frist we generate the random sample of size N=100 from the AR(1) process with the parameter ϕ=0.3(ϕ=0.6) and the outliers are occured at the time t=35, 59 and 87 with the effects δ=5. Finally we also consider the AR(2) process with the parameters ϕ_1=0.5 and ϕ_2=0.3(ϕ_1=-0.4 and ϕ_2=-0.6) with the same location of the outliers and the same outliers effects.

Simulations were performed to investigate the behaviours of CLSE, C&LE and PPE. For the each model (a) - (c) all given in case 1 and 2 are repeated times. The average and the mean square error from these repetitions are shown in table 1, table 2.

Through the Table 1 we can see the followings. For ϕ=0.3, 0.6, the proposed procedure estimator was better than CLSE and C&LE regardless the ourliers types. Also the outliers effects were estimated appropriately.

From the Table 2 we know the following facts. For ϕ_1=0.5, ϕ_2=0.3 and ϕ_1=-0.4, ϕ_2=-0.6, as AR(1) model, the proposed procedure estimator $\hat{\phi}_1$ and $\hat{\phi}_2$ were better than CLSE and C&LE, regardless the outliers types. Also the outliers effects were estimated appropriately.

5. Remarks and Conclusions

In this paper, we deal with the joint estimation problem in the autoregressive process with the two types of outliers.

We know that the proposed procedure is very effective in the sense of bias and mean square error. Regardless of the types of the outliers, the procedure can be applied to detect of the location of the outliers and to jointly estimates of the model parameters and the outliers effect. We recommand that you use this procedure to estimate the outliers effect and the parameters in the autoregressive time series model with any types of outliers.

In the autoregressive moving average process with the same condition given in this paper, the iterative procedure becomes more complicated but the outlier detection problems and the estimation problems of the outleirs effect and the model parameters can be dealt as the same manners.

Table 1

Comparisons of the CLSE, the C&LE and the PPE(N=100, δ=5, t=35,59,87)

1)ϕ=0.3

	CLSE		C&LE		PPE				
	$\hat{\phi}$	$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\sigma}^2$	$\hat{\delta}_{35}$	$\hat{\delta}_{59}$	$\hat{\delta}_{87}$
AO(3)	.181 (.008)	1.831 (.049)	.283 (.007)	1.061 (.040)	.293 (.008)	.998 (.018)	5.006 (.008)	4.999 (.010)	5.000 (.009)
IO(3)	.180 (.008)	1.845 (.065)	.281 (.007)	1.074 (.048)	.299 (.009)	1.056 (.026)	4.993 (.010)	4.987 (.009)	5.014 (.012)
AO(2) IO(1)	.175 (.008)	1.827 (.050)	.275 (.006)	1.066 (.044)	.294 (.008)	1.054 (.022)	5.000 (.010)	4.999 (.009)	4.984 (.059)
AO(1) IO(2)	.178 (.008)	1.847 (.048)	.282 (.006)	1.083 (.048)	.304 (.008)	1.064 (.026)	4.988 (.009)	5.003 (.008)	5.001 (.008)

2)ϕ=0.6

	CLSE		C&LE		PPE				
	$\hat{\phi}$	$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\sigma}^2$	$\hat{\delta}_{35}$	$\hat{\delta}_{59}$	$\hat{\delta}_{87}$
AO(3)	.445	1.979	.562	1.363	.587	1.002	4.994	5.007	4.993
	(.006)	(.061)	(.002)	(.141)	(.007)	(.025)	(.011)	(.009)	(.009)
IO(3)	.441	1.940	.562	1.311	.588	1.253	4.982	4.983	4.980
	(.006)	(.141)	(.002)	(.114)	(.006)	(.030)	(.009)	(.011)	(.009)
AO(2)	.437	1.957	.561	1.304	.581	1.260	4.984	4.964	4.983
IO(1)	(.005)	(.051)	(.001)	(.120)	(.006)	(.031)	(.012)	(.112)	(.009)
AO(1)	.436	1.971	.559	1.339	.597	1.266	4.976	4.991	4.993
IO(2)	(.006)	(.060)	(.002)	(.129)	(.005)	(.036)	(.009)	(.008)	(.008)

Note. (•) represents the mean square error (MSE)

Table 2

Comparisons of the CLSE, the C&LE and the PPE(N=100, δ=5, t=35,59,87)

1)ϕ_1=0.5, ϕ_2=0.3

	CLSE			C&LE			PPE					
	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\delta}_{35}$	$\hat{\delta}_{59}$	$\hat{\delta}_{87}$
AO(3)	.392	.260	1.792	.471	.273	1.179	.497	.279	.974	4.991	5.013	4.996
	(.012)	(.010)	(.067)	(.011)	(.010)	(.142)	(.010)	(.010)	(.022)	(.023)	(.022)	(.020)
IO(3)	.385	.258	1.867	.466	.271	1.229	.502	.273	.976	4.998	4.991	5.009
	(.009)	(.009)	(.079)	(.013)	(.009)	(.187)	(.011)	(.010)	(.023)	(.020)	(.022)	(.023)
AO(2)	.392	.259	1.827	.481	.266	1.163	.495	.277	.984	4.999	5.000	4.996
IO(1)	(.009)	(.010)	(.075)	(.012)	(.010)	(.179)	(.010)	(.009)	(.022)	(.021)	(.022)	(.023)
AO(1)	.393	.259	1.824	.482	.269	1.135	.498	.284	.992	4.994	5.013	5.017
IO(2)	(.010)	(.010)	(.079)	(.012)	(.010)	(.145)	(.008)	(.010)	(.022)	(.022)	(.025)	(.022)

2)ϕ_1=-0.4 ϕ_2=-0.6

	CLSE			C&LE			PPE					
	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2$	$\hat{\delta}_{35}$	$\hat{\delta}_{59}$	$\hat{\delta}_{87}$
AO(3)	-.169	-.321	1.984	-.355	-.534	1.291	-.385	-.578	.980	4.988	4.992	4.993
	(.012)	(.013)	(.091)	(.010)	(.009)	(.412)	(.007)	(.007)	(.018)	(.122)	(.070)	(.072)
IO(3)	-.216	-.382	1.725	-.352	-.521	1.303	-.380	-.583	.978	4.971	4.987	4.958
	(.010)	(.010)	(.076)	(.009)	(.010)	(.310)	(.008)	(.007)	(.022)	(.120)	(.021)	(.222)
AO(2)	-.204	-.370	1.765	-.349	-.531	1.239	-.390	-.584	.976	4.986	4.997	4.984
IO(1)	(.011)	(.011)	(.087)	(.010)	(.009)	(.290)	(.007)	(.008)	(.212)	(.071)	(.025)	(.018)
AO(1)	-.197	-.343	1.901	-.366	-.541	1.238	-.389	-.592	.979	4.989	4.974	4.984
IO(2)	(.010)	(.012)	(.088)	(.009)	(.009)	(.250)	(.007)	(.007)	(.212)	(.076)	(.074)	(.070)

Note. (•) represents the mean square error (MSE)

References

Abraham, B. and Box, G.E.P. (1979). Bayesian analysis of some outlier problems in time series. *Biometrika*. 66, 229-236.

Abraham, B. and Chuang, A. (1993). Expectation - Maximization Algorithms and the estimation of the time series models in the presence of outliers. *Journal of Time series amalysis.* 14, 221-234.

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

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 Liu, L.M (1993). Joint Estimation of Model parameters and outliers effect in time series. *Journal of American Statistical Association*. 88, 284-297.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likilihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. B* 39(1), 1-39.

Fox, A.J. (1972). Outliers in time series. *Journal of the Royal Statistical Society. B* 34(3), 350-363.

Tsay, R.S. (1986). Time series model specfication in the presence utliers. *Journal of American Statistical Association*. 81, 132-141.