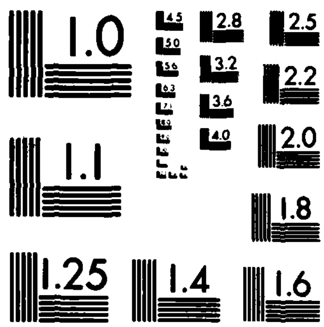
AD-A144 719 INFLUENTIAL OBSERVATIONS IN TIME SERIES(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER D PENA JUL 84 MRC-TSR-2718 DAAG29-80-C-0041 1/1

UNCLASSIFIED F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A144 719

MRC Technical Summary Report # 2718

INFLUENTIAL OBSERVATIONS IN TIME SERIES

Daniel Peña

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

July 1984

(Received June 7, 1984)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

INFLUENTIAL OBSERVATIONS IN TIME SERIES

Daniel Peña*

Technical Summary Report #2718

July 1984

ABSTRACT

This paper studies how to identify influential observations in univariate ARIMA time series models and how to measure their effects on the estimated parameters of the model. The sensitivity of the parameters to the presence of either additive or innovational outliers is analyzed, and influence statistics based on the Mahalanobis distance are presented. The statistic linked to additive outliers is shown to be very useful to indicate the robustness of the fitted model to the given data set. Its application is illustrated using simulation results and a relevant set of historical data.

AMS (MOS) Subject Classifications: 62M10, 62F35, 62-07

Key Words: influential observations, robustness of ARIMA models, diagnostic checks

Work Unit Number 4 - Statistics and Probability

---

* ETSII, Universidad Politecnica de Madrid, Spain. At the Mathematics Research Center and Statistics Department, University of Wisconsin-Madison at the time of this report.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and the United States-Spanish Joint Committee for Educational and Cultural Affairs.

SIGNIFICANCE AND EXPLANATION

Observed time series almost always have atypical points that are produced by nonsystematic changes in the variables that are driving the series. As the univariate forecasts from any univariate time series model are based on the extrapolation of the historical patterns, if the parameters of the series are very much dependent on a few atypical observations linked to non-repeatable events then, the quality of these forecasts can be very poor. Furthermore, the identification of these observations is very important in order to check the robustness of the fitted model.

This paper studies how to measure the influence of each observation on the estimated parameters of a time series ARIMA model. The effect of either additive or innovational outliers is analyzed, and simple expressions are obtained to measure their effects. A statistic is introduced that seems to be very useful to indicate influential observations and to judge the general robustness of the fitted model.

A-1

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

INFLUENTIAL OBSERVATIONS IN TIME SERIES

Daniel Peña*

1. INTRODUCTION AND SUMMARY

Observed time series almost always have atypical points. These anomalous values are produced by nonsystematic changes in the variables that are driving the series or affecting them. As the forecasts from any time series model are based on the extrapolation of the historical patterns, if the parameters of the series are very much dependent on a few atypical observations, resulting from isolated or non-repeatible events, then, the quality of the forecasts can be expected to be poor. Also, if the parameters of the model have physical or economical interpretation, the presence of undetected influential observations can mislead the scientist about the properties of the model. Finally, the study of these observations, that is the sensitivity of the model to the given data set, provides meaningful information about the robustness of the fitted model.

This problem is related to, although different from, the study of outliers because it is well known that the fact that an observation is an outlier does not imply this observation affects substantially the parameter estimates of the assumed model, although in general it will affect the variance of the estimates.

Cook and Weisberg (1982) and Belsley, Kuh and Welsch (1980) present an overview of influential observations in the regression model. This study has been extended to some other members of the generalized linear model family. (See Pregibon (1981).) Briefly, the main idea of this approach is to delete suspicious observations and build a measure of the change that this deletion produces in relevant features of the model, such as the estimated parameter values, or the forecasts.

* ETSII, Universidad Politecnica de Madrid, Spain. At the Mathematics Research Center and Statistics Department, University of Wisconsin-Madison at the time of this report.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and the United States-Spanish Joint Committee for Educational and Cultural Affairs.

The study of influential observations has been limited so far to independent data. This paper attempts to extend these ideas to dependent observations in the context of time series analysis and is organized as follows. Section 2 summarizes the literature of outliers in time series models and discusses the two basic types of outliers that can occur in a dynamic situation. Sections 3 and 4 show how to build measures of influence for additive outliers and for innovational outliers. Section 5 presents some simulated examples of the behavior of these statistics that are then applied in section 6 to the study of the robustness of a time series model.

The main result from this paper is to present a statistic that can be easily computed and seems to be very useful in indicating the observations that have strong influence on the estimated parameter values. Thus, this statistic can be incorporated easily into standard time series analysis practice and provides a quick and simple way to judge the robustness of the fitted model. Second, a simple expression has been found that relates the parameter values estimated with and without outliers in ARIMA models. This expression is used to prove that additive outliers are expected to be much more influential than inovational outliers. Third, on the assumption of innovational outliers the relevant statistics for influential observations are identical to those desired in the regression situation but their usefulness seems to be small in the time series context.

2. OUTLIERS IN TIME SERIES

Fox (1972) defines two types of outlier which may occur in time series data. The first, called *type I outlier* by Fox, corresponds to a modification of the value of the observed series due to some external cause, as a gross recording error or a intervention at some point. Assuming that the observed series z_t follows and autoregression moving average model, the model for a type I outlier is:

$$\phi(B)y_t = \theta(B)a_t$$

$$y_t = \begin{cases} z_t & t \neq T \\ z_t - w & t = T \end{cases}$$

where B is the backshift operator, $B^k y_t = y_{t-k}$ and $\phi(B) = 1-\phi_1 B - \ldots - \phi_p B^p$ and $\theta(B) = 1-\theta_1 B - \ldots - \theta_q B^q$ are the autoregressive and moving average polynomials. This model could also be written as:

$$z_t = w\zeta_t^{(T)} + \frac{\theta(B)}{\phi(B)} a_t$$

$$\zeta_t^{(T)} = \begin{cases} 0 & t \neq T \\ 1 & t = T \end{cases} \tag{2.1}$$

which points out that this model is a special case of intervention analysis (Box and Tiao (1975)) with a instantaneous reponse function, w. Model (2.1) has been called the additive outlier model by Denby and Martin (1979), and Chang and Tiao (1983), and the aberrant observation model by Abraham and Box (1979).

The type I outlier can be interpreted as the effect on the series of some external even or exogenous change in the system. On the other hand, the second type of outlier can be considered as the effect of some internal change or endogenous effect. If we think of a univariate time series model as an aggregate representation of the pattern of behavior of a vector x_t of explicative time series that are causing the observed series z_t, the noise of the univariate model represents the aggregate of the nonsystematic variation of the components x_t. An exogeneous intervention outlier in any of the components will produce an anomalous value on the noise of the univariate process. The model will be

$$\phi(B)z_t = \theta(B)(a_t + w\zeta_t^{(T)}) \tag{2.2}$$

where the atypical behavior appears on the innovation. This model has been called the innovational outlier (Cheng and Tiao (1983)) or the aberrant innovation model (Abraham and Box (1979)).

Calling $\psi(B) = \phi(B)^{-1}\theta(B)$ both types of outliers can be written as

$$z_t = \nu(B)w\zeta_t^{(T)} + \psi(B)a_t \tag{2.3}$$

where $\nu(B) = 1+\nu_1 B + \ldots$. Expression (2.3) shows that, as shown by Cheng and Tiao (1983), both types of outliers are particular cases in the general intervention analysis model (2.3). The cases $\nu(B) = 1$, additive outlier, and $\nu(B) = \psi(B)$, innovational outlier, are extreme cases in this representation and it is sensible to think of a third category of time series outliers in which $\nu(B)$ is any dynamic transfer function response. The study of this general class of outliers is still to be made.

Fox (1972) derived the maximum likelihood ratio test for both types of outliers for autoregressive processes. Abraham and Box (1979) used the normal contaminated model as the basic set-up to make inference in both types of model. Denby and Martin (1979) developed generalized M-estimators for the first order, autoregressive model and showed that, on the one hand, no great loss of efficiency is expected in estimating the parameter for least squares where there are innovational (type II) outliers but, on the other hand, if additive outliers are present, the loss of efficiency suffered can be large. Alba and Zartman (1980) have shown examples of robust estimation for ARIMA models and Chernick, Downing and Pike (1982) have studied the influence function for the autocorrelations of a stationary time series.

Finally, Chang and Tiao (1983) extended Fox's results to general ARIMA models and suggest a useful iterative procedure for outlier detection and parameter estimation. They recommend computing the likelihood ratio statistics $\lambda_{1,T}$ and $\lambda_{2,T}$ to check if the observation T is either an innovational outlier $(\lambda_{1,T})$ or an additive outlier $(\lambda_{2,T})$. These statistics are given by

$$\lambda_{1,T} = \frac{\hat{w}_I}{\hat{\sigma}_a} \qquad \lambda_{2,T} = \frac{\hat{w}_A}{\hat{\sigma}_a(1+\pi_1^2+\ldots \pi_h^2)^{-1/2}}$$

where $\hat{w}_I$ and $\hat{w}_A$ are the estimated values of the outlier w assuming that it belongs to the innovational type or additive type, and π_i are the parameters of the autoregressive representation of the process.

3. A MEASURE OF INFLUENCE FOR ADDITIVE OUTLIERS

3.1. The change in the parameter estimates

Suppose we have a stochastic process y_t that follows a univariate ARIMA (p, d, q). It is assumed in what follows that y_t represents deviations from some origin μ that will be the mean if the series is stationary, and that the moving average part has a characteristic equation with roots outside the unit circle so that the process is invertible. Then, the process can be represented as

$$y_t = \sum_{\ell=1}^{h} \pi_\ell y_{t-\ell} + a_t$$

for some lag h. If the process is purely autoregressive $h = p+d$, otherwise the π coefficients are obtained from $\pi(B) = \phi(B)(1-B)^d\theta(B)^{-1}$ and because of the invertibility of $\theta(B)$ these coefficients will decrease and eventually will become zero for some lag h.

Let us now assume that an additive outlier happens at time T and instead of observing y_t we observe z_t where $z_t = y_t (t \neq T)$ and $z_t = y_t + w (t=T)$. Then, as the Jacobian of the transformation from y_t to z_t is one, the likelihood function for the observed series z_t conditional to the first h values is

$$\ell(\underline{\pi}, \sigma^2, w) = -\left(\frac{n-h}{2}\right)\ln \sigma_a^2 - \frac{1}{2\sigma_a^2}\sum_{S_1}(z_t - \pi' x_t)^2 - \frac{1}{2\sigma_a^2}\sum_{\ell=0}^{h}(z_{T+\ell} + \pi_\ell w - \underline{\pi}' x_{T+\ell})^2$$

where σ_a^2 is the variance of the noise process a_t, $x_t' = (z_{t-1}, \ldots, z_{t-h})$, $\pi' = (\pi_1, \ldots, \pi_h)$, and $\pi_0 = -1$. The set of indicies S_1 is $(h+1, \ldots, T-1, T+h+1, \ldots n)$. The conditional maximum likelihood estimates of π and w are

$$\hat{\pi} = (\hat{X}_y' \hat{X}_y)^{-1} \hat{X}_y \hat{Y} \qquad (3.1)$$

$$\hat{w} = z_T - \hat{z}_{T/n}$$

$$\hat{z}_{T/n} = \sum_{i=1}^{h} \delta_i (z_{T+i} + z_{T-i}) \qquad (3.2)$$

where

$$\hat{X}_y = \begin{bmatrix} \hat{y}_h & \hat{y}_{h-1} & \cdots & \hat{y}_1 \\ \vdots & \vdots & \vdots & \vdots \\ \hat{y}_{n-1} & \hat{y}_{n-2} & \cdot & \hat{y}_{n-h} \end{bmatrix} \qquad \hat{Y} = \begin{bmatrix} \hat{y}_{h+1} \\ \vdots \\ \hat{y}_n \end{bmatrix}$$

$$\hat{\delta}_i = (\hat{\pi}_i - \sum_{\ell=1}^{h} \hat{\pi}_\ell \hat{\pi}_{\ell+i}) / (\sum_{\ell=0}^{h} \hat{\pi}_\ell^2) \quad \text{with } \hat{\pi}_0 = -1$$

and $\hat{y}_j = z_j$ if $j \neq T$, $\hat{y}_T = z_T - \hat{w} = \hat{z}_{T/n}$.

Consequently, the estimated residual in time T is $\hat{z}_{T/n} - \sum_{i=1}^{h} \hat{\pi}_i Z_{t-i}$, and $\hat{z}_{T/n}$ can be interpreted as the best estimation of Z_T using all the sample information. Neither this residual nor $Z_{T/n}$ depends on the value Z_T.

The system of equations given by (3.1) and (3.2) has to be resolved iteratively. Starting with an initial value $\hat{\pi}_{(0)}$ for $\hat{\pi}$, that is normally the maximum likelihood estimator of π assuming no outliers, the weights $\hat{\delta}_i$ can be calculated and a value $\hat{w}(0)$ computed. This value is then used to compute $\hat{y}_{T(1)} = z_t - \hat{w}(0)$ which provides a new estimate $\hat{\pi}_{(1)}$. The process is repeated until convergence.

To study the effect of the outlying value on the estimated parameters, let us call $\hat{\pi}_0$ the conditional maximum likelihood estimator of π assuming no outliers. Then,

$$\hat{\pi}_0 = (X_z^1 X_z)^{-1} X_z' Z$$

and X_z and $\hat{X}_y$ are related by

$$X_z = \hat{X}_y + \hat{w}M \tag{3.3}$$

where

$$M' = \left[O_{h\times(T-h)} \quad I_{h\times h} \quad O_{h\times(n-h-T)} \right]$$

and $O_{a\times b}$ represents a rectangular matrix of dimension $a \times b$ with all its elements equal to zero, and $I_{h\times h}$ is the square identity matrix. Partition the matrix X_z as

$$X_z' = [Z_1' \quad Z_2' \quad Z_3']$$

where

$$Z_1 = \begin{bmatrix} z_h & z_{h-1} & \cdots & z_1 \\ \vdots & \vdots & & \vdots \\ z_{T-1} & z_{T-2} & \cdots & z_{T-h} \end{bmatrix}, \quad Z_2 = \begin{bmatrix} z_T & z_{T-1} & \cdots & z_{T-h+1} \\ \vdots & \vdots & & \vdots \\ z_{T+h-1} & z_{T+h-2} & \cdots & z_T \end{bmatrix}, \quad Z_3 = \begin{bmatrix} z_{T+h} & \cdots & z_{T-1} \\ z_{N-1} & \cdots & z_{N-h} \end{bmatrix}$$

Then

$$(\hat{X}_y'\hat{X}_y) = X_z'X_z + \hat{w}^2 I - \hat{w}(Z_2 + Z_2') = X_z'X_z - \hat{w} A_T \tag{3.4}$$

where $A_T = Z_2 + Z_2' - \hat{w}I$. Also, let us partition the vector $\hat{Y}$ accordingly as

$$\hat{Y} = Z - \hat{w}V$$

where

$$v' = [0,0,\ldots,0,1 \quad 0,\ldots,0, \quad 0,\ldots,0]$$

and

$$Z' = (z_{h+1},\ldots,z_n)$$

then using (3.3) and (3.4)

$$\hat{X}_y'\hat{Y} = X_z'Z - \hat{w}\ S_T$$

where $S_T' = e(z_{T+1} + z_{T-1},\ z_{T+2} + z_{T-2},\ldots,z_{T+h} + z_{T-h})$.

Expressing the estimated parameters $\hat{\pi}$ as a function of the original data

$$(X_z'X_z - \hat{w}\ A_T)\hat{\pi} = X_z'Z - \hat{w}\ S_T$$

which leads to

$$\hat{\pi} = \hat{\pi}_0 - \hat{w}(X_z'X_z)^{-1}(S_T - A_T\hat{\pi}) \tag{3.5}$$

and calling $\hat{E}_T$ the vector of "pseudo-residuals" given by

$$\hat{E}_T = S_T - A_T\hat{\pi}$$

we obtain

$$\hat{\pi}_0 = \hat{\pi} + \hat{w}(X_z'X_z)^{-1}\ \hat{E}_T \quad . \tag{3.6}$$

To study how additive outliers influence the parameter estimates, we will express $X_z'X_z$ and E_T in (3.6) in terms of the uncontaminated process y_t. As

$$A_T = Z_2 + Z_2' - \hat{w}I$$

using that

$$Z_2 = Y_2 + \hat{w}I$$

then

$$A_T = Y_2 + Y_2' + \hat{w}I \quad ,$$

and so

$$\hat{E}_T = S_T - (Y_2 + Y_2')\hat{\pi} - \hat{w}\ \hat{\pi} \quad .$$

Inserting this expression for $\hat{E}_T$ in (3.6)

$$\hat{\pi}_0 = (I - \hat{w}^2(X_z'\ X_z)^{-1})\hat{\pi} + \hat{w}(X_z'X_z)^{-1}(S_T - (Y_2 + Y_2')\hat{\pi}) \quad .$$

If now $w \to \infty$, assuming a fix sample size n, as

$$\lim \hat{w}^{-2}(X'X)_z = \lim \hat{w}^{-2}(\hat{X}_y\hat{X}_y) + \lim \hat{w}^{-1}(Y_2 + Y_2') + I$$

and, as $\hat{w}$ is a consistent estimator of w, when $w \to \infty$ the limit of $\hat{w}^2(X'X)_z^{-1}$ is also I. Also, $\hat{w}^{-1}(X'X)_z \to \infty$ and $\hat{w}(X'X)_z \to 0$. In practice this result means that when w is large, all the estimated coefficients $\hat{\pi}_0$ are pulled down towards zero, and the series will appear to be white noise. For instance, in the AR(1) case

$$\hat{\phi}_0 = \left(1 - \frac{\hat{w}^2}{\Sigma y_t^2 + \hat{w}^2 + 2\hat{w}y_T}\right)\hat{\phi} + \frac{\hat{w}}{\Sigma y_t^2 + \hat{w}^2 + 2\hat{w}y_T}(y_{T+1} + y_{T-1} - 2y_T\hat{\phi})$$

and if w^2 is large compared to Σy_t^2 the value of $\hat{\phi}_0$ will be much smaller than $\hat{\phi}$.

The result that gross errors pull all the autocorrelation coefficients, and so the estimated parameters, towards zero was noted by Treadway (1978) and Guttman and Tiao (1978). An example of this problem with economic time series can be found in Peña and Sanchez-Albornoz (1983).

This result is in agreement with the properties of the estimated parameter for a first-order autoregressive process with additive outlier given by Martin and Jong (1977) and Denby and Martin (1979).

3.2. A statistic to measure influential outliers

A natural way to measure the influence of observation z_T is to relate it to the change on the parameter estimates when this observation is assumed to be an outlier. As $\hat{\pi}_0$ and $\hat{\pi}$ are vectors, a useful way suggested by Cook (1977) is to measure the distance between both vectors relative to a relevant positive semidefinite matrix M. A natural selection is to choose M as the variance covariance matrix of either of these two estimated vectors and to build, then, a Mahalanobis distance. In order to have a common ground to compare all the observations, it seems more useful to choose M as the covariance matrix of the parameters assuming no outliers (see Cook and Peña (1984)), then

$$D_2^*(T) = \frac{(\hat{\pi}_0 - \hat{\pi})'(X_z' X_z)(\hat{\pi}_0 - \hat{\pi})}{h\,\hat{\sigma}_a^2} \qquad (3.7)$$

where we have divided the distance by the dimension of the vectors involved, h, to have a proper standardization.

The statistic (3.7) can also be interpreted as measuring the change in the vector of one step ahead forecasts. Using the estimated parameters assuming no outliers this vector is

$$\hat{Z}_0 = X_z \hat{\pi}_0$$

and using the parameters estimated assuming an additive outlier at T:

$$\hat{Z} = X_z \hat{\pi} \quad .$$

The Euclidean distance between both vectors of forecasts is:

$$(\hat{Z}_0 - \hat{Z})'(\hat{Z}_0 - \hat{Z}) = (\hat{\pi}_0 - \hat{\pi})' X_z' X_z (\hat{\pi}_0 - \hat{\pi})$$

and so $D_2^*(T)$ can also be interpreted as a standardized measure of the Euclidean distance between the vectors of one step-ahead forecast built with $\hat{\pi}_0$ and $\hat{\pi}$.

Using (3.6), the statistic can be written as

$$D_2^*(T) = \frac{\hat{w}^2}{h \, \hat{\sigma}_a^2} \cdot \hat{E}_T'(X_z' X_z)^{-1} \hat{E}_T \quad .$$

The problem in computing $D_2^*(T)$ is that the estimates $\hat{w}$ and $\hat{E}_T$ require the nonlinear estimation of the intervention. A solution first suggested by Fox (1972) is to substitute $\hat{w}$ for another consistent estimator of w easier to compute. He suggested using the vector $\hat{\pi}_0$ instead of $\hat{\pi}$ to compute

$$\tilde{w}_0 = z_t - \sum_{i=1}^{h} \hat{\delta}_{0,i}(z_{T+i} + z_{T-1})$$

where

$$\hat{\delta}_{0,i} = (\hat{\pi}_{0,i} - \sum_{\ell=1}^{h-i} \hat{\pi}_{0,\ell} \hat{\pi}_{0,\ell-i}) / (\Sigma \hat{\pi}_{0,\ell}^2)$$

and $\hat{\pi}_{0,j}$ is the j^{th} component of $\hat{\pi}_0$. Fox (1972) verified using simulation that the approximation was good for moderate sample sizes. In the same spirit we suggest using $\hat{\pi}_0$ instead of $\hat{\pi}$ to compute $\hat{E}_T$. Calling

$$\tilde{E}_T = S_T - A_T \hat{\pi}_0$$

the statistics we obtain is

$$D_2(T) = \frac{1}{h} \frac{\tilde{w}_0^2}{\hat{\sigma}_a^2} \cdot \tilde{E}_T'(X_z' X_z)^{-1} \tilde{E}_T \quad . \tag{3.8}$$

The likelihood ratio test to check for additive outliers is asymptotically equivalent (see Chang and Tiao (1983)) to

$$\lambda_{2,T}^{2} = \frac{\tilde{w}_{0}^{2}}{\hat{\sigma}_{a}^{2}(\Sigma\, \hat{\pi}_{0,\ell}^{2})^{-1}}$$

and so, D_2 can be written

$$D_2(T) = \frac{\lambda_{2,T}^{2}}{h(\sum_{\ell=0}^{h} \hat{\pi}_{0,\ell}^{2})} \tilde{E}_T'(X_z' X_z)^{-1} \tilde{E}_T \quad . \tag{3.9}$$

Equation (3.9) shows the difference between detecting outliers and studying influence. The influence of a particular value can be decomposed into two terms. The first, $\lambda_{2,T}^{2} h^{-1}(\sum_{\ell=0}^{h} \hat{\pi}_{0,\ell}^{2})^{-1}$ is mainly a function of the size of the outlier relative to the model. The second, $\tilde{E}_T'(X_z X_z)^{-1}\tilde{E}_T$ is a measure of the relative importance of the observations in the series around the point in which the outlier happens. If we call, for $\ell = 1,\ldots h$

$$\hat{a}_{T+\ell} = Z_{T+\ell} - \hat{\pi}_1 Z_{T+\ell-1} - \ldots - \hat{\pi}_{\ell}(Z_T - \hat{w}) - \ldots - \hat{\pi}_h Z_{T+\ell-h}$$

the residuals in the model estimated allowing for outliers, and

$$\hat{b}_{T-\ell} = Z_{T-\ell} - \hat{\pi}_1 Z_{T-\ell+1} - \ldots - \hat{\pi}_{\ell} Z_T - \ldots - \hat{\pi}_h Z_{T-\ell+h}$$

the backwards residuals, then it is straightforward to show using the definition of $\hat{E}_T$ that:

$$\hat{E}_T' = [\hat{a}_{T+1} + \hat{b}_{T-1}, \ldots, \hat{a}_{T+h} + \hat{b}_{T-h}] \quad .$$

So the quadratic form $\tilde{E}_T'(X_z X_z)^{-1}\tilde{E}_T$ is taking into account that the importance of the outlier in the parameter estimates depends on the previous and posterior h observations.

This statistic should be computed as a routine diagnostic check for time series models because, as we have shown, the presence of additive outliers can have a strong influence on the estimated parameters of the model.

4. A Measure of influence for innovational outliers

4.1. The effect on the parameter estimates

Suppose that the observed series z_t follows the model

$$\pi(B)Z_t = w(B)\zeta_t^{(T)} + a_t$$

where now $\pi(B) = \phi(B)(1-B)^d\theta(B)^{-1}$, and $w(B) = w_0 + w_1B + \ldots + w_{S-1}B^{S-1}$ represents a general dynamic intervention at time T, with transfer function $\nu(B) = w(B)\pi(B)^{-1}$. The case of an innovational outlier corresponds to making $w(B) = w_0$. We assume that the process is invertible and so the π coefficients will become practically zero for some finite lag h. Then, the model can be written:

$$Z_t = \sum_{i=1}^{h} \pi_i Z_{t-i} + \sum_{j=0}^{s-1} w_j \zeta_{t-j}^{(T)} + a_t \tag{4.1}$$

where $\zeta_\ell^{(T)} = 0$ if $\ell \neq T$ and $\zeta_T^{(T)} = 1$. We assume as before that the mean level of the process Z_t has been removed. Then, the least squares estimates of the parameters, that are equal to the conditional maximum likelihood estimators, are:

$$\begin{bmatrix} \hat{\pi} \\ \hline \hat{w} \end{bmatrix} = \left[\begin{array}{c|c} X_z'X_z & X_z'D \\ \hline D'X_z & D'D \end{array}\right]^{-1} \begin{bmatrix} X_z'Z \\ \hline D'Z \end{bmatrix}$$

where X_z and Z were defined in section 3 and

$$D' = [O_{s\times(T-h)} \quad I_{s\times s} \quad O_{s\times(n-h-T)}]$$

and so $D'D = I$. Using now the expression for the inverse of a partitioned matrix;

$$\left[\begin{array}{c|c} X_z'X_z & X'D \\ \hline D'X_z & DD' \end{array}\right]^{-1} = \left[\begin{array}{c|c} (X_z'X_z)^{-1} & 0 \\ \hline 0 & 0 \end{array}\right] + \begin{bmatrix} -(X_z'X_z)^{-1}X_z'D \\ \hline I \end{bmatrix} \left[(I-D'HD)^{-1}\right] \left[\begin{array}{c|c} -D'X_z(X_z'X_z)^{-1} & I \end{array}\right]$$

where H is the idempotent matrix $X_z(X_z'X_z)^{-1}X_z'$. Then, after some straight-forward algebra:

$$\hat{\pi}_0 = \hat{\pi} + (X_z'X_z)^{-1}X_{T,s}'\hat{w} \tag{4.2}$$

$$\hat{w} = (I - D'HD)^{-1}[Z_{T,s} - X_{T,s}\hat{\pi}_0] \tag{4.3}$$

where $\hat{\pi}_0 = (X_z'X_z)^{-1}X_z'Z$ is, as in section 3, the vector of estimated parameters assuming no outliers and:

$$X_{T,s} = \begin{bmatrix} Z_{T-1} & \cdots\cdots & Z_{T-h} \\ Z_T & & Z_{T-h+1} \\ \vdots & & \vdots \\ Z_{T+s-2} & \cdots\cdots & Z_{T-h+s-1} \end{bmatrix}, \quad Z_{T,s} = \begin{bmatrix} Z_T \\ Z_{T+1} \\ \vdots \\ Z_{T+s+1} \end{bmatrix} .$$

Let us call

$$E_{T,s} = Z_{T,s} - X_{T,s}\hat{\pi}_0$$

the vector of estimated residuals for the relevant observations assuming no outliers, and let $H_{T,S} = D'HD$ be the square symmetric matrix of dimension s that contains the distances from the vectors $(Z_{t-1},\ldots,Z_{t-h})$ to the origin; then

$$\hat{w} = (I - H_{T,s})^{-1}E_{T,s} \quad . \tag{4.4}$$

In the particular case of an innovational outlier $w(B) = w$ and (4.2) and (4.4) reduce to

$$\hat{\pi}_0 = \hat{\pi} + \hat{w}(X_z'X_z)^{-1}x_T \tag{4.5}$$

$$\hat{w} = e_{T,0}(1-d_T)^{-1} \tag{4.6}$$

where $e_{T,0} = Z_T - \hat{\pi}_0'X_T$ is the residual at point T from the model without intervention, and $d_T = X_T'(X'X)^{-1}X_T$ is the distance from the vector $X_T' = (Z_{T-1},\ldots,Z_{T-h})$ to the origin.

As an example, consider the AR(1) case. Calling $\hat{\phi}$ the parameter, (4.5) reduces to

$$\hat{\phi}_0 = \hat{\phi} + \frac{\hat{w}\; Z_{T-1}}{\Sigma\; Z_t^2}$$

and $\hat{\phi}_0$ can be greater or smaller than $\hat{\phi}$ depending on the sign of Z_{T-1}. It is clear that when $n \to \infty$, $\hat{\phi}_0 \to \hat{\phi}$, the least square estimator, and so $\hat{\phi}_0$ is a consistent estimator of ϕ.

The fact that under innovational outliers the least squares estimators of the parameters of an autoregressive process are consistent was first obtained by Mann and Wald (1943). Martin and Jong (1977) have studied the efficiency of the estimators in the first-order autoregressive process and show that although consistent the estimator can be quite inefficient.

These results explain the well known flexibility and adaptability of univariate ARMA models for forecasting purposes. As we noticed earlier, this type of outlier can be interpreted as a sudden change in one or more of the unknown series that are driving the observed series Z_t. It can be concluded that ARMA models are fairly robust at this kind of internal perturbation.

4.2. Building a measure of influence for innovational outliers

We consider again model (4.1) that has a general dynamic outlier with transfer function $\nu(B) = w(B)\pi(B)^{-1}$. Then, a measure of influence can be built as before using

$$D(T,w) = \frac{(\hat{\pi}_0-\hat{\pi})'M(\hat{\pi}_0-\hat{\pi})}{h}$$

where M is a positive definite matrix that defines the metric. Choosing, as in section 3.2, $M = (X_z'X_z)\sigma_a^{-2}$ and using (4.2) and (4.4), then

$$D(T,w) = \frac{E_{T,s}'(I - H_{T,s})^{-1}H_{T,s}(I - H_{T,s})^{-1}E_{T,s}}{h\,\hat{\sigma}_a^2} \tag{4.7}$$

where we have used that:

$$H_{T,s} = X_{T,s}(X_z'X_z)^{-1}X_{T,s}' \quad .$$

The statistics $D(T,w)$ is similar to the one proposed by Cook and Weisberg (1982) to analyze the influence of a set of points on the parameter estimates of the regression model. The similarity is clear because of the linear structure of equation (4.1). If we now assume that $w(B) = w$, and so $\zeta_t^{(T)}$ is an innovational outlier, using (4.5) and (4.6):

$$D_1(T) = \frac{1}{h}\,\frac{e_T^2}{\sigma_a^2(1-d_T)}\left(\frac{d_T}{1-d_T}\right) \quad . \tag{4.8}$$

that is identical to the statistics suggested by Cook (1977) to measure the effect of an observation in the parameters of a regression model. The statistics can be interpreted as the product of two terms: The first, $e_T^2\,\sigma_a^{-2}(1-d_T)^{-1}$ is the standardized residual at the point of the intervention, and the second, $d_T(1-d_T)^{-1}$, represents the distance of x_T to the origin but with relation to a metric build without taking into account x_T. $D_1(T)$ can also be expressed as a function of the likelihood criteria advocated by Fox (1972) and Chang and Tiao (1983) to test for innovational outliers:

$$D_1(T) = \frac{\lambda_T^2}{h} \frac{d_T}{(1-d_T)^2} .$$

Note that the influence of the outlier observation depends now only on the relative values of the h observations before the intervention, as measured by d_T, and not on the h posterior observations, in contrast with what happened in the additive outliers case. In the next section we will compare both statistics.

5. Comparison of both statistics

5.1. Deleting versus forecasting

There are two basic theoretical reasons to recommend using the influence measure based on additive outlier (3.8) instead of the one derived for innovational outliers (4.8) as a routine checking device in univariate time series model. The first, is that additive outlier are expected to be much more influential than innovational outliers. The second, is that this measure provides a more reasonable generalization to dynamic problems of the measures derived in the context of independent observations.

The main justification for the influence measures suggested for the linear model is that these measures are standardized versions of a finite approximation to the influence curve introduced by Hampel (1974). (See for instance Cook and Weisberg (1982) and Welsch (1982)). However, all this theory relies on a sample of independent and identically distributed observations, which is obviously inadequate to deal with stochastic processes and, in particular, with time series.

In the regression model, for instance, the empirical influence curve for one observation can be expressed as the difference between the parameter estimated with and without this observation. This idea cannot be generalized to time series in which the deletion of one observation changes the dynamic of the sample. However, it is well known that in the regression situation the deletion procedure is equivalent to treat the observation as a missing value and estimate the model:

$$Y = X\beta + \zeta_{(T)} w + U \tag{5.1}$$

where Y is the vector of response, X the matrix of explicative variables, β the vector of parameters, $\zeta_{(T)}$ is a dummy or intervection variable (as defined in section 2) and U is the vector of perturbations. Then, the vector of estimated parameters from (5.1), $\hat{\beta}_{(T)}$, does not depend on the T observation (y_T, x_T).

This "missing value" procedure can be extended in a straightforward idea to any dynamic situation, because it leads to substitute the observation under investigation by its forecast using all the sample, instead of deleting it. Of course, in the regression case of independent observations both procedures are equivalent, but they are not in the time series context. The key point of the approach is to obtain an estimator of the parameters that does not depend on the data under investigation.

This is the approach used in the additive outlier case, and it is easy to prove that, in contrast with the innovational outliers model, the parameter estimated obtained for the additive outlier case does not depend on the observation under question. Note that $\hat{\pi}$ is given by (3.1) and as $\hat{y}_T = \hat{Z}_{T/n}$, that does not depend on the value Z_T, neither $\hat{X}_y$ nor $\hat{Y}$ depend on Z_T.

5.2. Some simulation results

To illustrate the behavior of these statistics in practice, figure 1 shows 50 simulated values of an AR(1) time series with $\phi = .7$ in which an additive outlier equal to 4 standard deviations of the series has been added in position 33. The estimated parameter value drops from .7 to .58 and figure 1 displays the behavior of D_1 and D_2.

The maximum value of the D_1 statistic occurs in observation 34, instead of in observation 33 as could be expected. To understand why, let us look at the AR(1) process as a regression equation, $Z_T = \phi x_T + a_T$, with $X_T = Z_{T-1}$. Then, the first component of D_1 is the square of the standardized residual

$$\frac{e_T^2}{\hat{\sigma}^2(1-d_T)}$$

and the second, $d_T(1-d_T)^{-1}$ is a measure of the distance of x_T to the center of gravity of the previous observations. As here $X_T = Z_{T-1}$, if Z_T is anomalous because of the presence of an additive outlier, then $x_{T+1} = Z_T$ will be far away from the rest, and the term $d_T(1-d_T)^{-1}$ will be very big and can dominate the product, as happens here.

The statistic D_2, however, shows clearly the 33 observation as atypical. Note that observations Z_{T-1} and Z_{T+1} are shown as influential too, although less that Z_T as expected. In general, if the process is AR(p), the influential effect of an additive outlier in T appears, although with decreasing effect, on observations Z_{T-p} to Z_{T+p} as well. This effect is symmetric around Z_T. For instance in the AR(1) case, it is easily seen that if there is an additive outlier w on time I,

$$E[\tilde{w}_{0,T-1} \mid \hat{\phi}_0^2] = E[\tilde{w}_{0,T+1}] = -\frac{w\hat{\phi}_0}{1+\hat{\phi}_0^2} .$$

This effect can be noticed in the plot of the $\tilde{w}_0$ values displayed in figure 1.

Figure 2 shows a second simulation of the same process, using differenct random numbers. However, the outlier, although of the same magnitude (4σ) as before, has now been added on observation 30. The estimated parameter turns out to be .4335, which means that the outlier is more influential in this second case. The plot of D_1 suggests that observations 30^{th} and 31^{th} are both influential and roughly in the same amount. However, the statistics D_2 indicates clearly to observation 30^{th} as the most influential.

Table 1 displays some relevant values of both statistics in these two simulations. The largest value of $D_2(T)$ in the second simulation shows that the decrease in the estimated value is greater, as we have seen. In both cases $D_1(T)$ shows observation $T + 1$ as the most influential, but other simulations have shown that this result does not hold in general.

In summary, for additive outliers statistic D_2 seems to have a very stable behavior and correctly identifies the outlier value as more or less influential. On the other hand, the behavior of D_1 is not so consistent which makes its interpretation and use less reliable.

The better behavior of D_2 in the case of additive outliers is not surprising, because this statistic has been built precisely to measure these effects. However, the simulations we have made seem to indicate that its behavior is very consistent for innovational outliers as well. For instance, figure 3a shows the result of simulating 50 observations from the model:

$$z_t = \frac{10}{1-.9B}\zeta_t^{(40)} + \frac{a_t}{1-.9B}$$

where $\zeta_t^{(4)} = 0$, $t \neq 40$, and $\zeta_{40}^{(40)} = 1$. The estimated parameter when fitting an AR(1) is $\hat{\phi} = .89$, and so in this case the effect of the outlier in the parameter estimate is very small. The maximum value of $D_1(T)$ is .04 for observation 40. However the behavior of $D_1(T)$ is clearly erratic, as shown in figure 3a. $D_2(T)$ pinpoints observation 39 as the most important in the parameter estimates. The reason is that, for the AR(1), we have shown that

$$\phi_0 = \phi + \frac{\hat{w}\, Z_{T-1}}{\sigma\, Z_t^2}$$

and so the change in the parameter estimates depend on $\hat{w}$ and on Z_{T-1} and either Z_t or Z_{T-1} can appear as more influential depending on their relative values. The maximum value of D_2, $D_2(39)$, is .2 which means that the effect of the outlier is small.

Figure 3b shows a second simulation of the model but with $\phi = .3$, as in

$$Z_t = \frac{10}{1-.3B}\,\zeta_t^{(4)} + \frac{a_t}{1-.3B} .$$

The estimated AR(1) parameter turns out to be .38 with that sample. The figure 3b shows that $D_2(T)$ indicates again observation 39 as more influential, and $D_2(39) = 1.5$, which means that now the change in the estimated parameter is larger. $D_1(T)$ has its maximum value at point 41 but their value is small, .125, indicating, (wrongly) a small effect on the estimated parameter.

We conclude that statistics $D_2(T)$ seems to have a better behavior not only in the case of additive outliers but for innovational outliers as well.

TABLE 1

	T	$D_1(T)$	$D_1(T+1)$	$D_2(T-1)$	$D_2(T)$	$D_2(T+1)$
simulation 1	33	.002	.736	.133	.641	.383
simulation 2	30	.366	.466	.049	1.482	.195

Figure 1

Simulation of an AR(1) with $\phi = .7$ and an additive outlier in position 33 (Z_T) and behavior of two statistics for influence observations.

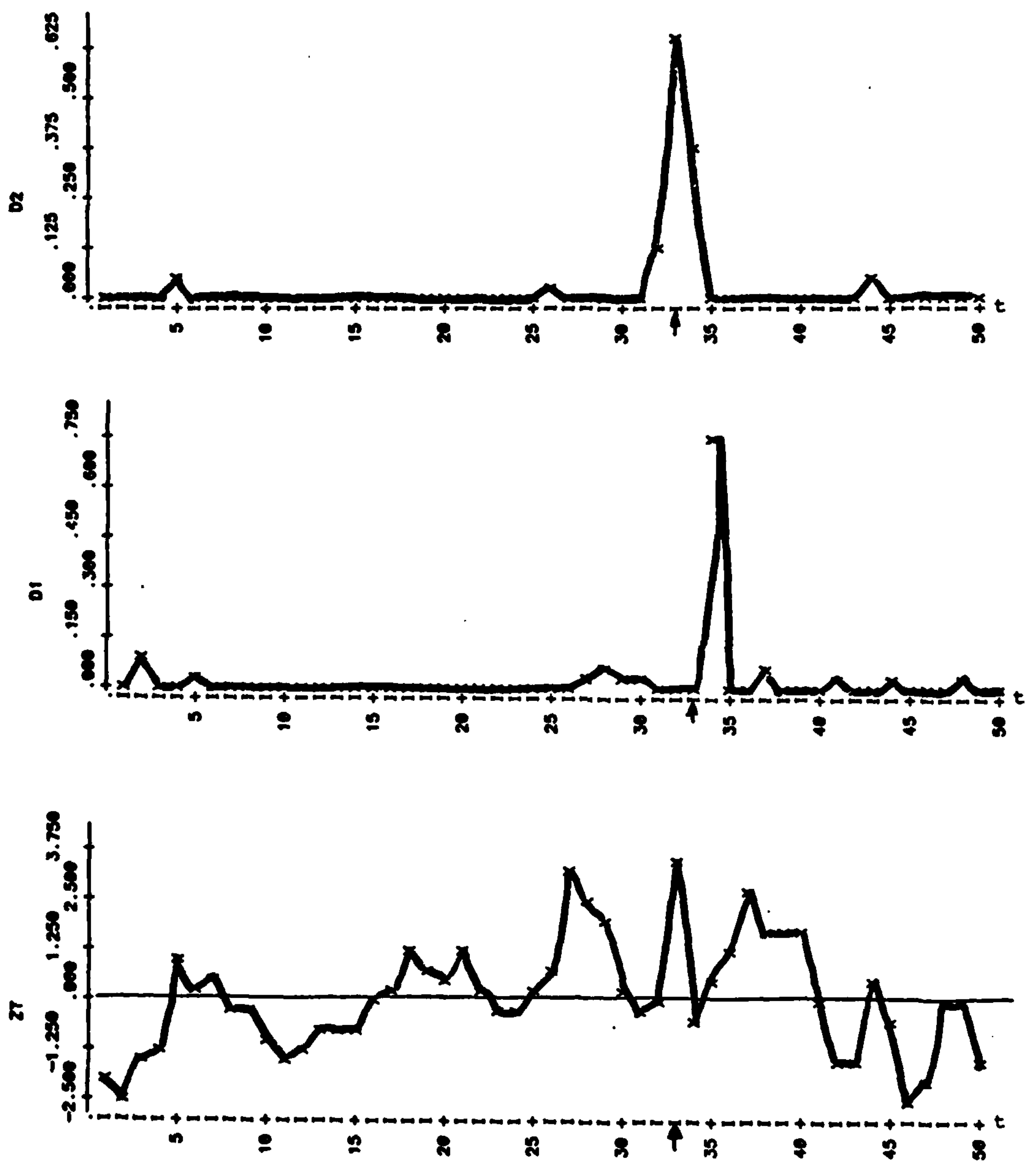


Figure 2

Simulation of an AR(1) with $\phi = .7$ and an additive outlier in position 30 (Z_T) and the behavior of two statistics for influence observations.

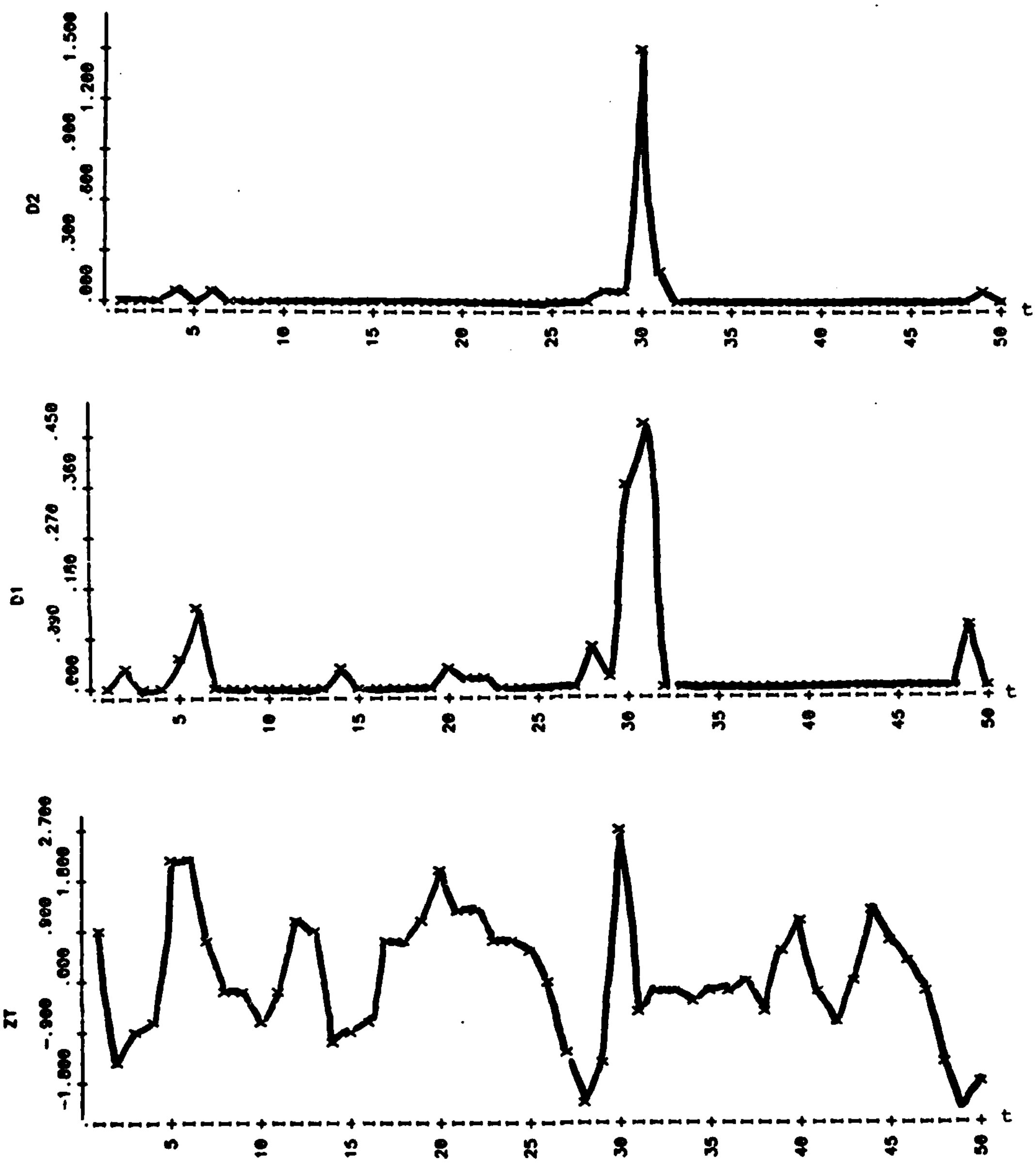


Figure 3a

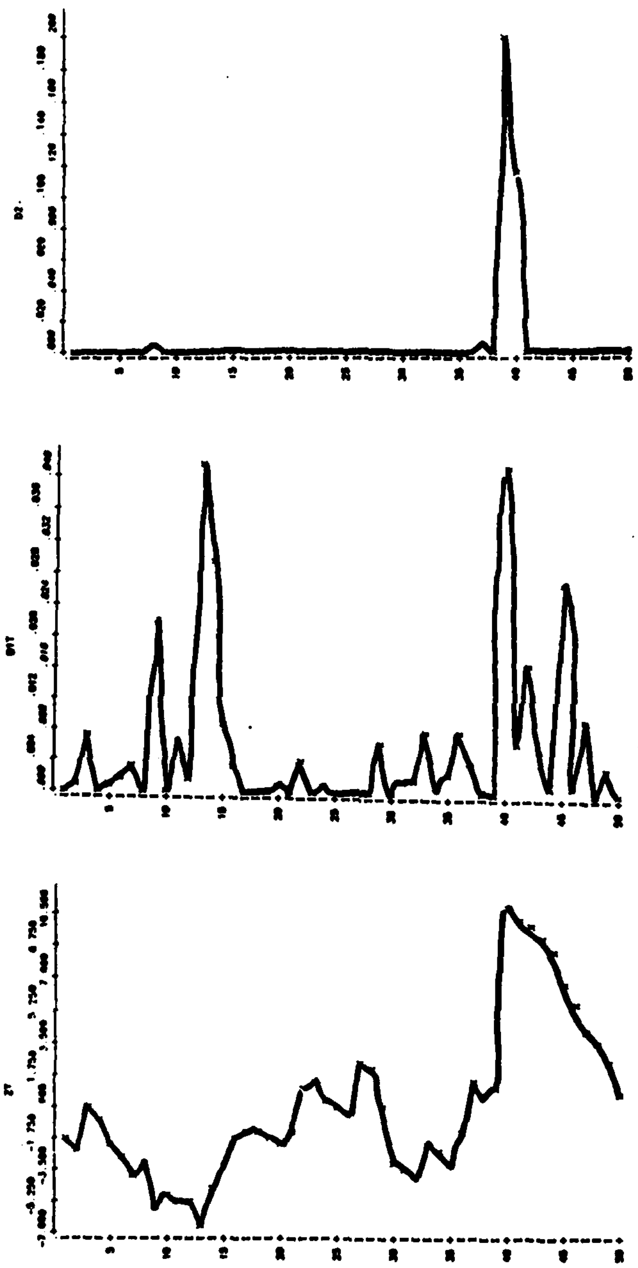
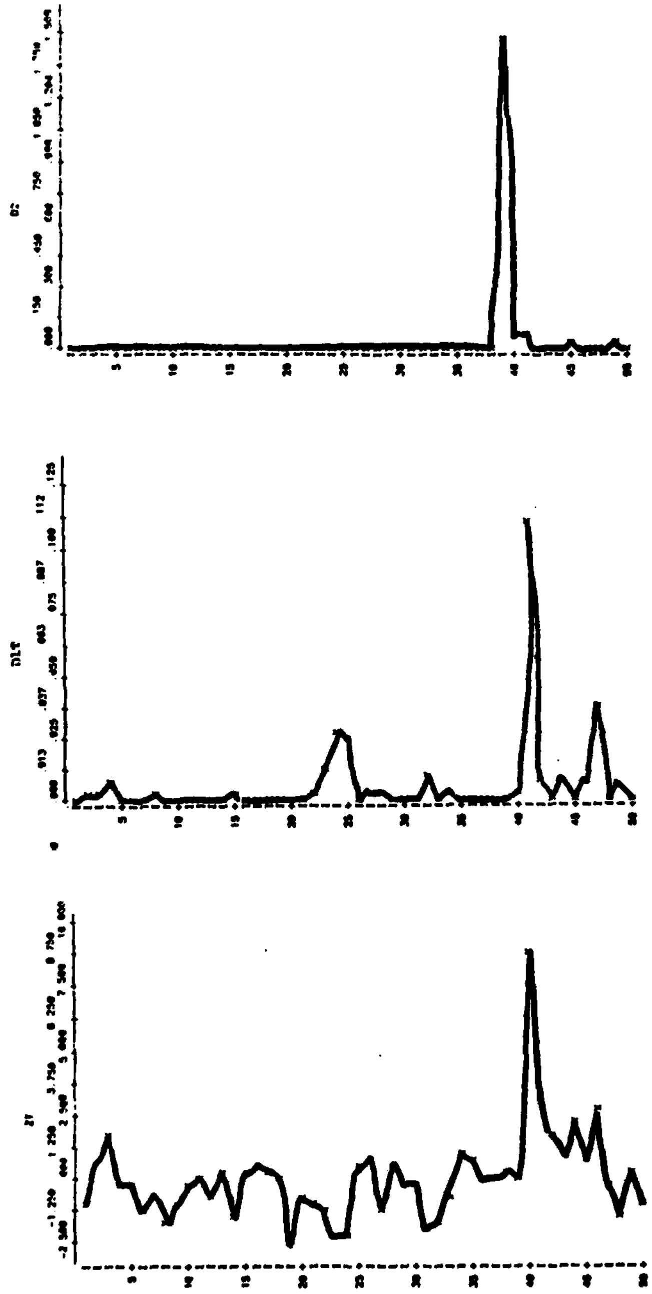


Figure 3b



6. An Application

The data that will be used to illustrate the use of the previous statistics is the series of extinctions of marine animals over the past 250 million years and is displayed in figure 4. These data have been studied by Raup and Sepkoski (1984) and show periodicity in the peaks of extinctions that they attribute to deterministic extraterrestrial causes. Kitchell and Peña (1984) show that the observed pseudo-periodicity can be explained by a fifth order non stationary autoregressive process with one root equal to 1 and four others complex roots.

Table 2 displays the original death rate series (D), the residuals of the best estimated model (E) and the values of the D_1 and D_2 statistics. Both of those influence statistics are plotted in figure 4.

Again D_1 fails to indicate clearly the influential points and shows peaks in the 32^{th} and 34^{th} observation. However, D_2 pinpoints without any doubt observation 30. The atypical value of this observation is clear from figure 3 and the residual at this point is outstanding and bigger than 3 standard deviations. However, the small value of D_2 for this point (.389) suggests that this observation is not very influential as far as the parameter values are concerned. So, although there are only 39 observations the autoregressive model is very robust to the effect of a single outlier.

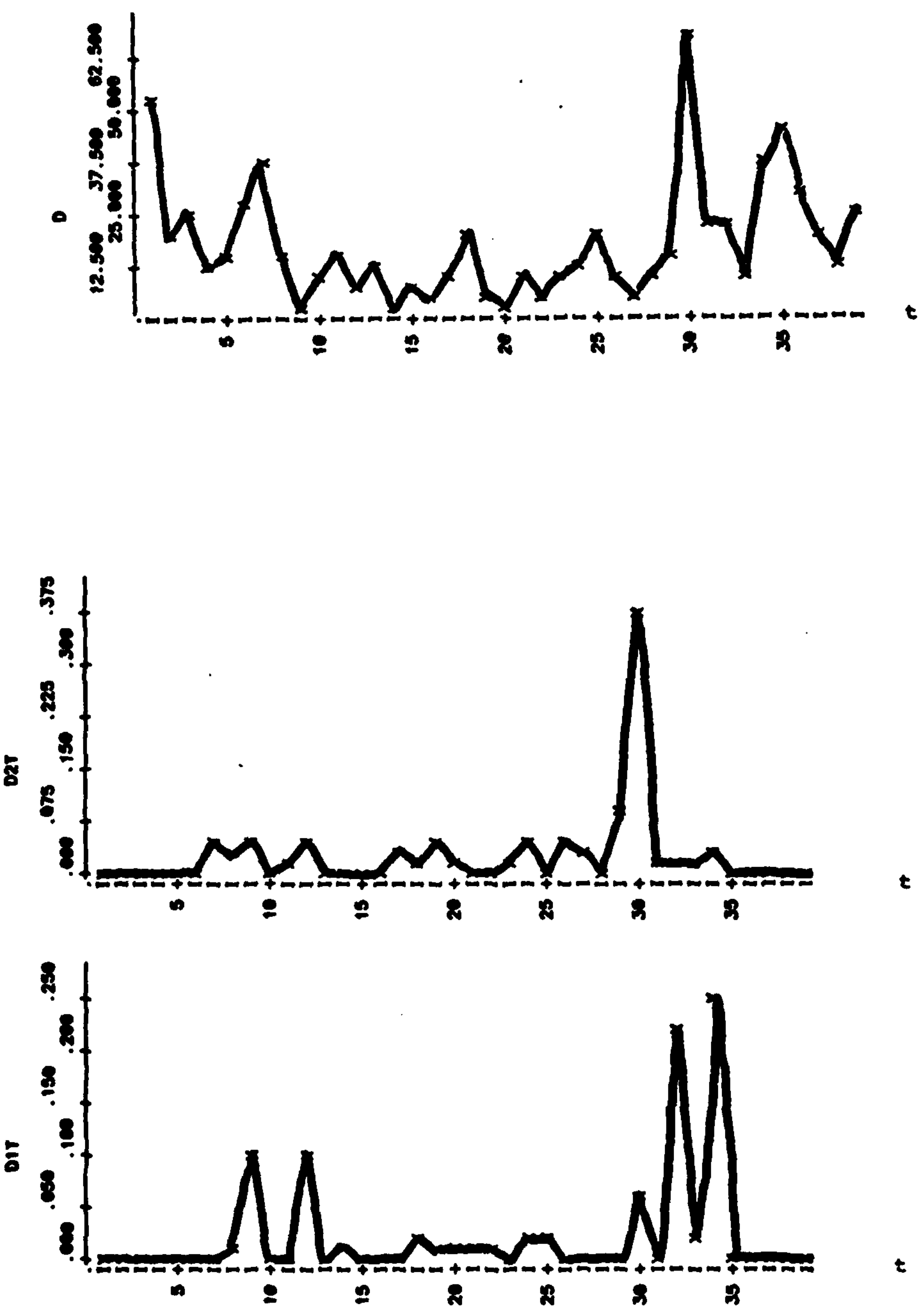
Table 3a presents the estimated autoregressive model with and without outliers. As the data are proportions different transformations has been used to test the sensitivity of the conclusion to the metric of the data. Table 3b presents the results with the logit transformation $y_t = \ln \frac{z_t}{1-z_t}$. It can be seen that the results are very similar, and the same hold for other possible transformations that have been applied.

Table 2

Original series (D), estimated residual from the AR(5) model (E), and statistics $D_1(T)$, and $D_2(T)$.

VARIABLE			$D_1(T)$	$D_2(T)$
COLUMN	D	E		
ROW	1	1	1	1
1	52.500	***	.000	.000
2	21.000	***	.000	.000
3	24.000	***	.000	.000
4	12.800	***	.000	.000
5	15.900	***	.000	.000
6	26.400	-.142	.002	.005
7	38.600	.703	.005	.051
8	15.900	-.643	.011	.030
9	2.600	-1.867	.104	.044
10	10.100	.140	.001	.000
11	15.200	-.306	.004	.021
12	7.100	-1.429	.097	.047
13	11.600	.361	.003	.004
14	3.500	-.567	.006	.003
15	7.600	.014	.000	.000
16	6.000	-.527	.005	.006
17	9.800	.289	.002	.028
18	19.500	.949	.016	.012
19	3.800	-.823	.014	.047
20	3.600	-.479	.005	.010
21	9.500	.650	.011	.000
22	6.000	-.679	.010	.003
23	10.200	.087	.000	.008
24	12.000	.859	.016	.051
25	18.900	1.111	.021	.000
26	9.900	-.167	.000	.047
27	5.800	-.374	.001	.031
28	9.200	.107	.000	.001
29	14.700	.257	.001	.087
30	66.300	2.384	.056	.381
31	22,200	-.054	.001	.011
32	21.900	.760	.216	.022
33	11.100	-.230	.020	.019
34	36.700	.851	.250	.033
35	45.800	.031	.000	.000
36	29.400	-.080	.000	.000
37	20.000	-.114	.000	.000
38	12.500	-.417	.000	.000
39	25.000	-.076	.000	.000

Figure 4



D
12.500
25.000
37.500
50.000
62.500
t
5
10
15
20
25
30
35
D2T
.000
.075
.150
.225
.300
.375
D1T
.000
.050
.100
.150
.200
.250

Table 3

MODEL	Q(11)	$\hat{\sigma}_a^2$
a) $(1 + \underset{(4.2)}{.66}B + \underset{(3.8)}{.56}B^2 + \underset{(4.6)}{.71}B^3 + \underset{(2.56)}{.38}B^4)\nabla Z_t = at$	6.2	122.8
$Z_t = \underset{(6.71)}{42.68}I_{30} + \dfrac{a_t}{\nabla(1 + \underset{(1.94)}{.32}B + \underset{(3.75)}{.56}B^2 + \underset{(2.71)}{.45}B^3 + \underset{(1.55)}{.22}B^4}$	4.1	63.65
b) $(1 + \underset{(4.0)}{.62}B + \underset{(3.8)}{.59}B^2 + \underset{(3.9)}{.62}B^3 + \underset{(2.7)}{.42}B^4)\nabla y_t = a_t$	7.0	.574
$y_t = \underset{(3.68)}{2.05}I_{30} + \dfrac{a_t}{\nabla(1 + \underset{(3.1)}{.49}B + \underset{(4.2)}{.65}B^2 + \underset{(2.5)}{.43}B^3 + \underset{(2.7)}{.40}B^4}$	5.0	.440

Z_t is the extinction rate series (series D in table 2) and Y_t is its logit transformation $Y_t = \ln Z_t/(100 - Z_t)$, B is the backshift operator, $\nabla = 1-B$, I_{30} is an indicator variable with $I(30) = 1$ and $I(i) = 0 \;\forall i \neq 30$, $Q(g)$ is the Ljung-Box statistics with g degrees of freedom and $\hat{\sigma}_a^2$ is the residual variance of the model.

ACKNOWLEDGEMENT

The author is very grateful to George E. P. Box and George C. Tiao for their comments and encouragement, and greatly indebted to Arthur B. Treadway who taught the author the importance of outliers in time series models.

REFERENCES

1. Abraham, B. and Box, G. E. P. (1979). "Bayesian analysis of some outliers problems in time series", Biometrika, 66, 2, 229-36.

2. Alba, E. and Zartman, D. L. (1980). "Testing outliers in time series: An application to remotely sensed temperatures inn cattle", in Analyzing Time Series, O. D. Anderson, editor, North-Holland.

3. Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). Regression Diagnostics, Wiley.

4. Box, G. E. P. and Tiao, G. G. (1975). "Intervention analysis with applications to economic and environmental problems", JASA, 70, 70-79.

5. Cook, R. D. (1977). "Detection of influential observations in linear regression", Technometrics, 19, 15-18.

6. Cook, R.. D. and Weisberg, S. (1982). Residuals and Influence in Regression, Chapman and Hall.

7. Cook, R. D. and Peña, D. (1984). "A Comparation of Influence Measure". Technical Summary Report, Mathematics Research Center, University of Wisconsin-Madison.

8. Chang, I. and Tiao, G. G. (1983). "Estimation of Time Series parameters in the presence of outliers", Technical Report No. 8, Statistics Research Center, University of Chicago.

9. Chernick, M. R., Downing, D. J. and Pike, D. H. (1982). "Detecting outliers in time series data". Journal of Americal Statistical Association, 77, 380, 743-747.

10. Denby, L. and Martin, R. D. (1979). "Robust estimation of the first-order autoregressive parameter". Journal of the American Statistical Association, 74, 365, 140-146.

11. Fox, A. J. (1972). "Outliers in Time Series", Journal of Royal Statistical Society, B, 3, 350-363.

12. Guttman, I. and Tiao, G. G. (1978). "Effect of correlation on the estimation of a mean in the presence of spurious observations". The Canadian Journal of Statistics, 6, 2, 229-247.

13. Hampel, F. R. (1974). "The influence curve and its Role in Robust estimation". Journal of the American Statistical Association, 62, 1179-1186.

14. Kitchell, J. A. and Peña, D. (1984). "Periodicity of extinctions in the geological past: Deterministic versus stochastic explanations", submitted to Science.

15. Mann, H. B. and Wald, A. (1943). "On the statistical treatment of linear stochastic difference equations", Econometrika, 11, 3 and 4, 173-220.

16. Martin, R. D. and Jong, J. (1977). "Asymptotic properties of Robust generalized M-estimates for the first order autoregressive parameter". Bell Laboratories memorandum, Murray Hill, NJ.

17. Peña, D. and Sanchez-Albornoz, D. (1983). Dependencia dinamica entre precios agricolas. El trigo en España, 1857-1890. Servicio de estudios del Banco de España.

18. Pregibon, D. (1981). "Logistic regression diagnostics". Annals of Statistics, 9, 705-724.

19. Raup, D. M. and Sepkoski, J. UJ. (1984). "Periodicity of extinctions in the geological part". Proc. of Nat. Acad. of Science, (to appear).

20. Treadway, A. B. (1978). Efectos sobre la economia espanola de una devaluacion de la peseta, Fundacion Ramon-Areces, Madrid.

21. Welsch, R. E. (1982). "Influence Functions and Regression Diagnostics" in Modern Data Analysis, Launer, R. L. and Siegel, A. F. editors. Academic Press.

DP/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2718	2. GOVT ACCESSION NO. AD A144 719	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Influential Observations in Time Series		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Daniel Peña		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics & Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709		12. REPORT DATE July 1984
		13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

influential observations, robustness of ARIMA models, diagnostic checks

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper studies how to identify influential observations in univariate ARIMA time series models and how to measure their effects on the estimated parameters of the model. The sensitivity of the parameters to the presence of either additive or innovational outliers is analyzed, and influence statistics based on the Mahalanobis distance are presented. The statistic linked to additive outliers is shown to be very useful to indicate the robustness of the fitted model to the given data set. Its application is illustrated using simulation results and a relevant set of historical data.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

10-84

DTIC