Chapter 15
Finding Outliers in Linear and Nonlinear Time Series

Pedro Galeano and Daniel Peña

15.1 Introduction

Outliers, or discordant observations, can have a strong effect on the model building process for a given time series. First, outliers introduce bias in the model parameter estimates, and then, distort the power of statistical tests based on biased estimates. Second, outliers may increase the confidence intervals for the model parameters. Third, as a consequence of the previous points, outliers strongly influence predictions. There are two main alternatives to analyze and treat outliers in time series. First, robust procedures can be applied to obtain parameter estimates not affected by the presence of outliers. These robust estimates are then used to identify outliers by using the residuals of the fit. Second, diagnostic methods are useful to detect the presence of outliers by analyzing the residuals of the model fit through iterative testing procedures. Once the outliers have been found, their effects are jointly estimated with the model parameters, obtaining, as a by-product, robust model parameter estimates. In this paper we focus on diagnostic methods and refer to Chap. 8 of Maronna et al. (2006) for a detailed review of robust procedures for ARMA models and Muler and Yohai (2008) and Muler et al. (2009) for two recent references.

For linear models, Fox (1972) introduced additive outliers (AO), which affect a single observation, and innovative outliers (IO), which affect a single innovation, and proposed the use of likelihood ratio test statistics for testing for outliers in autoregressive models. Tsay (1986) proposed an iterative procedure to identify outliers, to remove their effects, and to specify a tentative model for the underlying process. Chang et al. (1988) derived likelihood ratio criteria for testing the existence of outliers of both types and criteria for distinguishing between them and proposed

P. Galeano (✉) · D. Peña
Departamento de Estadística, Universidad Carlos III de Madrid, C/Madrid, 126, 28903 Getafe, Madrid, Spain
e-mail: pedro.galeano@uc3m.es

D. Peña
e-mail: daniel.pena@uc3m.es

C. Becker et al. (eds.), *Robustness and Complex Data Structures*,
DOI 10.1007/978-3-642-35494-6_15,

an iterative procedure for estimating the time series parameters in ARIMA models. Tsay (1988) extended the previous findings to two new types of outliers: the level shift (LS), which is a change in the level of the series, and the temporary change (TC), which is a change in the level of the series that decreases exponentially. Chen and Liu (1993a) proposed an iterative outlier detection to obtain joint estimates of model parameters and outlier effects that leads to more accurate model parameter estimates than previous ones. Luceño (1998) developed a multiple outlier detection method in time series generated by ARMA models based on reweighed maximum likelihood estimation. Gather et al. (2002) proposed a partially graphical procedure based on mapping the time series into a multivariate Euclidean space which can be applied online. Sánchez and Peña (2010) proposed a procedure that keeps the powerful features of previous methods but improves the initial parameter estimate, avoids confusion between innovative outliers and level shifts and includes joint tests for sequences of additive outliers in order to solve the masking problem. Finally, papers dealing with seasonal ARIMA models are Perron and Rodríguez (2003), Haldrup et al. (2011) and Galeano and Peña (2012), among others.

Recently, the focus has moved to outliers in nonlinear time series models. For instance, Chen (1997) proposed a method for detecting additive outliers in bilinear time series. Battaglia and Orfei (2005) proposed a model-based method for detecting the presence of outliers when the series is generated by a general nonlinear model that includes as particular cases the bilinear, the self-exciting threshold autoregressive (SETAR) model and the exponential autoregressive model, among others. In financial time series modeling, Doornik and Ooms (2005) presented a procedure for detecting multiple AO's in generalized autoregressive conditional heteroskedasticity (GARCH) models at unknown dates based on likelihood ratio test statistics. Carnero et al. (2007) studied the effect of outliers in the identification and estimation of GARCH models. Grané and Veiga (2010) proposed a general detection and correction method based on wavelets that can be applied to a large class of volatility models. Hotta and Tsay (2012) introduced two types of outlier in GARCH models: the level outlier (LO) corresponds to the situation in which a gross error affects a single observation that does not enter into the volatility equation, while the volatility outlier (VO) corresponds to the previous situation but the outlier enters into the volatility affecting all the remaining observations in the time series. Finally, Fokianos and Fried (2010) introduced three different outliers for the particular case of integer-valued GARCH (INGARCH) models and proposed a multiple outlier detection procedure for such outliers.

The literature on outliers in multivariate time series is brief. Tsay et al. (2000) generalized the four types of outliers usually considered in ARIMA models to the case of vector autoregressive moving average (VARMA) models and highlighted the differences between univariate and multivariate outliers. Importantly, the effect of a multivariate outlier not only depends on the model and the outlier size, as in the univariate case, but on the interaction between the model and size. These authors also proposed an iterative procedure for estimating the location, type and size of multivariate outliers. Galeano et al. (2006) proposed a method based on projections for identifying outliers without requiring initial specification of the multivariate model.

These authors showed that a multivariate outlier produces at least a univariate outlier in almost every projected series, and by detecting the univariate outliers, it is possible to identify the multivariate ones. Baragona and Battaglia (2007a) and Baragona and Battaglia (2007b) have proposed methods to discover outliers in dynamic factor models and in multiple time series by means of an independent component approach, respectively. Finally, Pankratz (1993) has considered outliers in dynamic regression models.

Other interesting issues related with outliers have been analyzed in the literature. For instance, detection of outliers in online monitoring data have been developed in Davies et al. (2004) and Gelper et al. (2009), among others. The effects of outliers in exponential smoothing techniques have been considered by Kirkendall (1992) and Koehler et al. (2012). The relationship between outliers, missing observations and interpolation techniques have been analyzed in Peña and Maravall (1991), Battaglia and Baragona (1992), Ljung (1993) and Baragona (1998). Forecasting time series with outliers have been addressed by Chen and Liu (1993b), for ARMA models, Franses and Ghijsels (1999), for GARCH models, and Gagné and Duchesne (2008), for dynamic vector time series models.

The rest of this contribution is organized as follows. In Sect. 15.2, we review outliers in univariate ARIMA models and discuss procedures for outlier detection and robust estimation. In Sect. 15.3, we consider outliers in non-linear time series models. Section 15.4 is devoted to outliers in multivariate time series models. Finally, Sect. 15.5 concludes the paper.

15.2 Outliers in ARIMA Models

This section reviews outliers in ARIMA time series models. We first introduce the four types of outliers usually considered in these models: additive outlier, innovative outlier, level shift and temporary change. Another type of unexpected events can be considered in the framework of an intervention event in the time series data, such as the ramp shift. Then, we describe procedures for outlier identification and estimation.

15.2.1 Types of Outliers in ARIMA Models

15.2.1.1 The ARIMA Model

We say that x_t follows an ARIMA(p,d,q) model if x_t can be written as:

$$\phi(B)(1-B)^d x_t = c + \theta(B)e_t, \tag{15.1}$$

where c is a constant, B is the backshift operator such that $Bx_t = x_{t-1}$, $\phi(B)$ and $\theta(B)$ are polynomials in B of orders p and q given by $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$

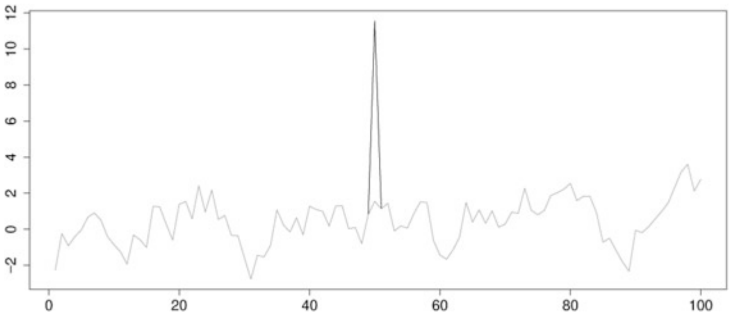

Fig. 15.1 Stationary series with and without an AO

and $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$, respectively, d is the number of unit roots and e_t is a white noise sequence of independent and identically distributed (i.i.d.) Gaussian with mean zero and variance σ_e^2. It is further assumed that the roots of $\phi(B)$ and $\theta(B)$ are outside the unit circle and have no common roots. The autoregressive representation of the ARIMA model in (15.1) is given by $\pi(B)x_t = c_\pi + e_t$, where $c_\pi = \theta^{-1}(B)c$ and $\pi(B) = \theta(B)^{-1}\phi(B)(1-B)^d$, while the moving average representation reduces to $x_t = c_\psi + \psi(B)e_t$, where $c_\psi = \phi^{-1}(B)(1-B)^{-d}c$ and $\psi(B) = \phi(B)^{-1}(1-B)^{-d}\theta(B)$.

15.2.1.2 Additive Outliers

An additive outlier (AO) corresponds to an exogenous change of a single observation of the time series and is usually associated with isolated incidents like measurement errors or impulse effects due to external causes. A time series $y_1, \ldots, y_T$ affected by the presence of an AO at $t = k$ is given by:

$$y_t = x_t + w I_t^{(k)}$$

for $t = 1, \ldots, T$, where x_t follows an ARIMA model in (15.1), w is the outlier size and $I_t^{(k)}$ is an indicator variable such that $I_t^{(k)} = 1$, if $t = k$, and $I_t^{(k)} = 0$, otherwise.

Figure 15.1 shows a simulated series with sample size $T = 100$ following an AR(1) model with parameter $\phi = 0.8$ and innovation variance $\sigma_e^2 = 1$, and the same series with an AO of size $w = 10$ at $t = 50$. Note how only a single observation is affected. An AO can have pernicious effects in all the steps of the time series analysis, i.e., model identification, estimation and prediction. For instance, the autocorrelation and partial autocorrelation functions, that are frequently used for model identification, can be severely affected by the presence of an AO.

15.2.1.3 Innovative Outliers

An innovative outlier (IO) corresponds to an endogenous change of a single innovation of the time series and is usually associated with isolated incidents like impulse

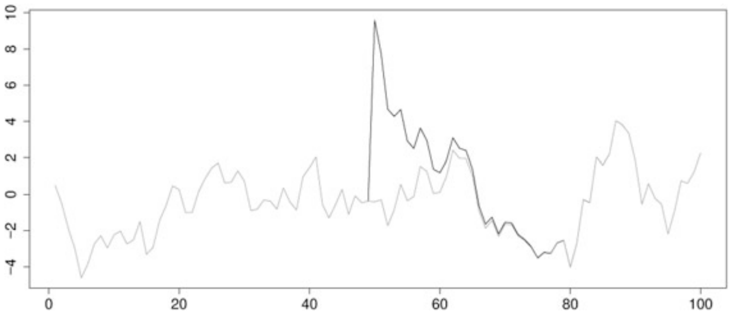

Fig. 15.2 Stationary series with and without an IO

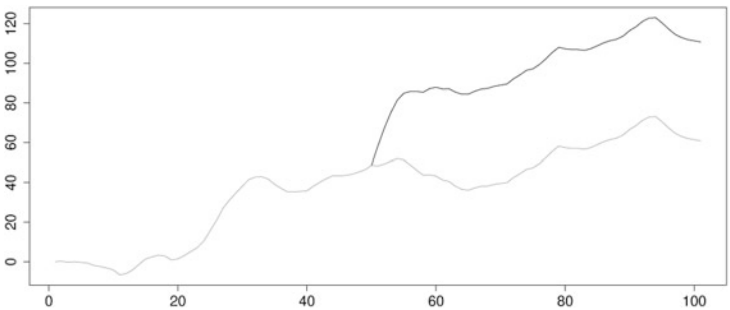

Fig. 15.3 Nonstationary series with and without an IO

effects due to internal causes. The innovations of a time series $y_1, \ldots, y_T$ affected by the presence of an IO at time point $t = k$ is given by:

$$a_t = e_t + wI_t^{(k)}, \tag{15.2}$$

where e_t are the innovations of the clean series x_t. Multiplying $\psi(B)$ in both sides of (15.2) leads to the equation for the observed series:

$$y_t = x_t + \psi(B)wI_t^{(k)}.$$

The effects of an IO on a series depend on the series being stationary or not. To see this point, Fig. 15.2 shows a simulated series with sample size $T = 100$ following an AR(1) model with parameter $\phi = 0.8$ and innovation variance $\sigma_e^2 = 1$, and the same series with an IO of size $w = 10$ at $t = 50$. Note how the IO modifies several observations of the series although its effect tends to disappear after a few observations. On the other hand, Fig. 15.3 shows a simulated series with sample size $T = 100$ following an ARIMA(1, 1, 0) model with parameter $\phi = 0.8$ and innovation variance $\sigma_e^2 = 1$, and the same series with an IO of size $w = 10$ at $t = 50$. Note how, in this case, the IO affects all the observations of the series starting from time point $t = 50$.

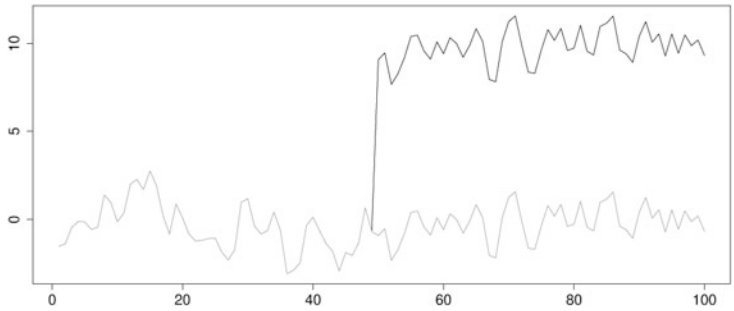

Fig. 15.4 Stationary series with and without a LS

15.2.1.4 Level Shifts

A level shift (LS) is a change in the mean level of the time series starting at $t = k$ and continuing until the end of the observed period. Therefore, a time series $y_1, \ldots, y_T$ affected by the presence of a LS at $t = k$ is given by:

$$y_t = x_t + wS_t^{(k)},$$

where $S_t^{(k)} = (1 - B)^{-1} I_t^{(k)}$ is a step function. Note that a LS serially affects the innovations as follows:

$$a_t = e_t + \pi(B) w S_t^{(k)}.$$

Figure 15.4 shows a simulated series with sample size $T = 100$ following an AR(1) model with parameter $\phi = 0.8$ and innovation variance $\sigma_e^2 = 1$, joint with the same series with a LS of size $w = 10$ at $t = 50$. Note how the LS affects all the observation of the series after $t = 50$. A LS has a strong effect in both identification and estimation of the observed series. Indeed, the effect of an LS is close to the effect of an IO on a nonstationary series.

15.2.1.5 Temporary Changes

A temporary change (TC) is a change with effect that decreases exponentially. Therefore, a time series $y_1, \ldots, y_T$ affected by the presence of a TC at $t = k$ is given by

$$y_t = x_t + \frac{1}{1 - \delta B} w I_t^{(k)},$$

where δ is the exponential decay parameter such that $0 < \delta < 1$. Note that if δ tends to 0, the TC reduces to an AO, whereas if δ tends to 1, the TC reduces to a LS. Under the presence of a TC, the innovations are affected as follows:

$$a_t = e_t + \frac{\pi(B)}{1 - \delta B} w I_t^{(k)}.$$

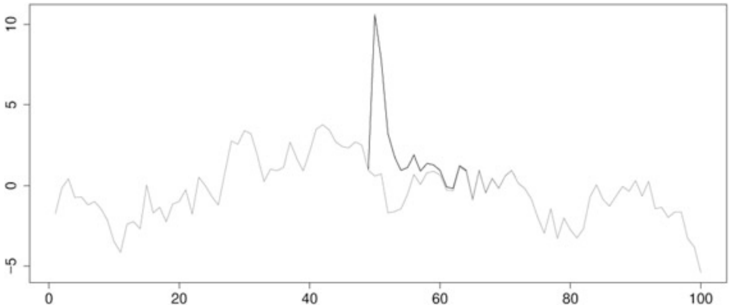

Fig. 15.5 Stationary series with and without a TC

Then, if $\pi(B)$ is close to $1-\delta B$, the effect of a TC on the innovations is very close to the effect of an IO. Otherwise, the TC can affect several innovations.

Figure 15.5 shows a simulated series with sample size $T = 100$ following an AR(1) model with parameter $\phi = 0.8$ and innovation variance $\sigma_e^2 = 1$, and the same series with an TC of size $w = 10$ and decay rate $\delta = 0.7$ at $t = 50$. Note how the TC have a decreasing effect in the observations of the series after $t = 50$.

15.2.1.6 Ramp Shifts

Finally, a ramp shift (RS) is a change in the trend of the time series in an ARIMA$(p,1,q)$ model starting at $t = k$ and continuing until the end of the observed period. Therefore, a time series $y_1,\dots,y_T$ affected by the presence of a RS at $t = k$ is given by

$$y_t = x_t + wR_t^{(k)},$$

where $R_t^{(k)} = (1-B)^{-1}S_t^{(k)}$ is a ramp function. Note that a RS on the $I(1)$ series y_t is a LS on the differenced series $(1-B)y_t$. Note also that a RS serially affects the innovations as follows:

$$a_t = e_t + \pi(B)wR_t^{(k)}.$$

Other types of unexpected events have been considered in the literature. For instance, variance changes have been considered in Tsay (1988), while patches of additive outliers have been studied in Justel et al. (2001) and Penzer (2007). Modeling alternative unexpected events is also possible as the intervention framework is flexible enough to model many different situations. For example, new effects can be defined using combinations of the outliers previously considered.

15.2.2 Outlier Identification and Estimation

In general, all the types of outliers that we have presented can be written in a general equation:

$$y_t = x_t + \nu(B) w I_t^{(k)}, \tag{15.3}$$

where $\nu(B) = 1$ for an AO, $\nu(B) = \psi(B)$ for an IO, $\nu(B) = (1 - B)^{-1}$ for a LS, $\nu(B) = (1 - \delta B)^{-1}$ for a TC and $\nu(B) = (1 - B)^{-2}$ for a RS, respectively. Therefore, outliers in time series can be seen as particular cases of interventions, introduced by Box and Tiao (1975), to model dynamic changes on a time series at known time points.

Assume that we observe a series, $y_1, \ldots, y_T$, following an ARIMA(p, d, q) model as in (15.1) with known parameters and with an outlier of known type at $t = k$. Multiplying by $\pi(B)$ in (15.3) leads to the equation for the innovations:

$$a_t = e_t + w_i z_{i,t}, \tag{15.4}$$

for $i =$ AO, IO, LS and TC, where w_{AO}, w_{IO}, w_{LS} and w_{TC} is the size of the outlier for AO, IO, LS and TC, respectively, and $z_{i,t} = \pi(B)\nu_i(B) I_t^{(k)}$, where $\nu_{\text{AO}}(B) = 1$, $\nu_{\text{IO}}(B) = \psi(B)$, $\nu_{\text{LS}}(B) = (1 - B)^{-1}$ and $\nu_{\text{TC}}(B) = (1 - \delta B)^{-1}$, respectively. From (15.4), for any particular case, one can easily estimate the size of the outlier by least-squares leading to:

$$\hat{w}_i = \frac{\sum_{t=1}^{T} z_{i,t} a_t}{\sum_{t=1}^{T} z_{i,t}^2}$$

with variance $\rho_i^2 \sigma_e^2$ where $\rho_i^2 = (\sum_{t=1}^{T} z_{i,t}^2)^{-1}$. Consequently, knowing the type and location of the outlier, it is easy to adjust the outlier effect on the observed series using the corresponding estimates, $\hat{w}_{\text{AO}}$, $\hat{w}_{\text{IO}}$, $\hat{w}_{\text{LS}}$ or $\hat{w}_{\text{TC}}$, respectively.

Also, the estimates of the outlier size can now be used to test whether one outlier of known type has occurred at $t = k$. Indeed, the likelihood ratio test statistic for the null hypothesis $H_0 : w_i = 0$ against the alternative $H_1 : w_i \neq 0$, is given by

$$\tau_{i,k} = \frac{\hat{w}_{i,k}}{\rho_i \sigma_e}.$$

The statistic $\tau_{i,k}$, under the null hypothesis, follows a Gaussian distribution.

However, in practice, the number, location, type and size of the outliers are unknown. Several papers, including Chang et al. (1988), Tsay (1988), Chen and Liu (1993a) and Sánchez and Peña (2010), among others, have proposed iterative procedures in which the idea is to compute the likelihood ratio test statistics for all the observations of the series under the null hypothesis of no outliers. In particular, the procedure by Chen and Liu (1993a), which is standard nowadays, works as follows. In a first step, an ARIMA model is identified for the series and the parameters are estimated using maximum likelihood. Then, the likelihood ratio test statistics $\tau_{i,t}$, for

$i =$ AO, IO, LS and TC, are computed. If the maximum of all these statistics is significant, an outlier of the type that provides with the maximum statistic is detected. Then, the series is cleaned of the outlier effects and the parameters of the model are re-estimated. This step is repeated until no more outliers are found. In a second step, the outliers effects and the ARIMA model parameters are estimated jointly using a multiple regression model. If some outlier is not significant, it is removed from the outliers set. Then, the multiple regression model is re-estimated. This step is repeated until all the outlier effects are significant. In a final step, the two previous steps are repeated but initially using the ARIMA model parameters estimates obtained at the end of the second step. However, this procedure has three main drawbacks. First, when a level shift is present in the series, the procedure tends to identify an innovative outlier instead of the level shift. Second, the initial estimation of the model parameters usually leads to a very biased set of parameters that may produce the procedure to fail. Third, the masking and swamping effects, although mitigated with respect to previous procedures, are still present if a sequence of outlier patches is present in the time series. Sánchez and Peña (2010) proposed a procedure for multiple outlier detection and robust estimation that tries to avoid these three problems. In particular, to solve the first problem, it is proposed to compare AO versus IO and deal with LS alone. To solve the second problem, it is proposed to use influence measures to identify the observations that have a larger impact on estimation and estimate the parameters assuming that the most influential observations are missing. Finally, to solve the third problem, an influence measure for LS or sequences of patchy outliers is proposed that can be used to carry out the initial cleaning of the time series.

15.3 Outliers in Nonlinear Time Series Models

This section reviews outliers in some nonlinear time series models. In particular, we first consider the model-based method proposed by Battaglia and Orfei (2005) for detecting the presence of outliers when the series is generated by a general nonlinear model. Second, we summarize the effect of outliers in GARCH models following Carnero et al. (2007) and present a method proposed by Hotta and Tsay (2012) for detecting outliers in GARCH models. Finally, we describe the method proposed by Fokianos and Fried (2010) to detect outliers in INGARCH models.

15.3.1 Outliers in a General Nonlinear Model

Battaglia and Orfei (2005) assumed a time series x_t following the model:

$$x_t = f\left(\mathbf{x}^{(t-1)}, \mathbf{e}^{(t-1)}\right) + e_t, \tag{15.5}$$

where f is a nonlinear function also containing unknown parameters, $\mathbf{x}^{(t-1)} = (x_{t-1}, x_{t-2}, \ldots, x_{t-p})'$, $\mathbf{e}^{(t-1)} = (e_{t-1}, e_{t-2}, \ldots, e_{t-p})'$ and e_t is a white noise sequence of independent and identically distributed (i.i.d.) Gaussian with mean zero and variance σ_e^2. Note that the model in (15.5) covers several well known nonlinear models, such as the bilinear model, the self-exciting threshold autoregressive (SETAR) model and the exponential autoregressive model, among others.

For the model in (15.5), Battaglia and Orfei (2005) consider additive and innovative outliers. First, for an AO at $t = k$, the observed series is $y_1, \ldots, y_T$, given by $y_t = x_t + wI_t^{(k)}$, for $t = 1, \ldots, T$, where x_t follows the model in (15.5). Therefore, the observed series can be written as $y_t = f(\mathbf{y}^{(t-1)}, \mathbf{a}^{(t-1)}) + a_t$, where $\mathbf{y}^{(t-1)} = (y_{t-1}, y_{t-2}, \ldots, y_{t-p})'$ and $\mathbf{a}^{(t-1)} = (a_{t-1}, a_{t-2}, \ldots, a_{t-p})'$, respectively, for $t = 1, \ldots, T$. The innovations of the observed series can be obtained recursively from $a_t = y_t - f(\mathbf{y}^{(t-1)}, \mathbf{a}^{(t-1)})$. On the other hand, for an IO at $t = k$, the observed series is given by $y_t = f(\mathbf{y}^{(t-1)}, \mathbf{a}^{(t-1)}) + a_t$, where $a_t = e_t + wI_t^{(k)}$, where $\mathbf{y}^{(t-1)} = (y_{t-1}, y_{t-2}, \ldots, y_{t-p})'$ and $\mathbf{a}^{(t-1)} = (a_{t-1}, a_{t-2}, \ldots, a_{t-p})'$, respectively, for $t = 1, \ldots, T$.

Estimation of outlier effects can be done similarly to the ARIMA case through least squares. Battaglia and Orfei (2005) showed that the LS estimate of w for an IO is given by $\hat{w}_{\mathrm{IO}} = a_k$ with variance σ_e^2, and that the LS estimate of w for an AO is given by

$$\hat{w}_{\mathrm{AO}} = \frac{\sum_{j=0}^{T-k} c_j a_{k+j}}{\sum_{j=0}^{T-k} c_j^2}$$

with variance $(\sum_{j=0}^{T-k} c_j^2)\sigma_e^2$, where

$$c_j = -\left[\frac{\partial}{\partial y_{t-j}} f\left(y^{(k+j-1)}, a^{(k+j-1)}\right) + \sum_{i=1}^{j} c_{j-i} \frac{\partial}{\partial a_{t-j}} f\left(y^{(k+j-1)}, a^{(k+j-1)}\right) \right],$$

for $j = 1, \ldots, T - k$. Consequently, the likelihood ratio test statistics to test for the presence of an AO and a IO at $t = k$ are given by

$$\tau_{\mathrm{IO}} = \frac{a_k}{\sigma_e},$$

and

$$\tau_{\mathrm{AO}} = \frac{\sum_{j=0}^{T-k} c_j a_{k+j}}{\sqrt{\sum_{j=0}^{T-k} c_j^2}\sigma_e},$$

respectively. Under the null hypothesis of no outlier, τ_{AO} and τ_{IO} have a standard Gaussian distribution.

In order to detect the presence of several outliers in a nonlinear time series, Battaglia and Orfei (2005) considered a procedure similar to that used in Chen and Liu (1993a).

15.3.2 Outliers in GARCH Models

Carnero et al. (2007) have analyzed the effects of outliers on the identification and estimation of GARCH models. Regarding identification, Carnero et al. (2007) derived the asymptotic biases caused by outliers on the sample autocorrelations of squared observations generated by stationary processes and obtained the asymptotic biases of the ordinary least squares (OLS) estimator of the parameters of ARCH(p) models. Finally, these authors also studied the effects of outliers on the estimated asymptotic standard deviations of the estimators considered and showed that they are biased estimates of the sample standard deviations.

Recently, Hotta and Tsay (2012) have distinguished two types of outliers in GARCH models and have proposed a method for their detection. For simplicity of presentation, we consider the ARCH(1) model given by

$$\begin{aligned} x_t &= \sqrt{h_t} e_t, \\ h_t &= \alpha_0 + \alpha_1 x_{t-1}^2, \end{aligned}$$

where $\alpha_0 > 0$, $0 < \alpha_1 < 1$, and e_t are independent and identically distributed standard Gaussian random variables. Outliers in an ARCH(1) model encounter two different scenarios because an outlier can affect the level of x_t or the volatility h_t. Therefore, a volatility outlier, denoted by VO, and defined as follows

$$\begin{aligned} y_t &= \sqrt{h_t} e_t + w I_t^{(k)}, \\ h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 \end{aligned}$$

affects the volatility of the series, while a level outlier, denoted by LO, and given by

$$\begin{aligned} y_t &= \sqrt{h_t} e_t + w I_t^{(k)}, \\ h_t &= \alpha_0 + \alpha_1 \left(y_{t-1} - w I_{t-1}^{(k)} \right)^2 \end{aligned}$$

only affects the level of the series at the observation where it occurs.

Hotta and Tsay (2012) estimated w by means of the ML estimation method. These authors showed that the ML estimate of w for a VO is given by

$$\hat{w}_{\mathrm{VO}} = y_k,$$

and that there are two ML estimates of w for a LO: the first one is $\hat{w}_{\mathrm{LO}} = y_k$ and the second one is $\hat{w}_{\mathrm{LO}} = y_k - \hat{x}_k$, where $\hat{x}_k =$ is the square root of the positive solution of the second order equation:

$$\begin{aligned} g(x) = {} & \alpha_1^2 x^2 + \left(2\alpha_0\alpha_1 + \alpha_1^2 \left(\alpha_0 + \alpha_1 y_{k-1}^2 \right) \right) x \\ & + \left(\alpha_0^2 + \alpha_0\alpha_1 \left(\alpha_0 + \alpha_1 y_{k-1}^2 \right) - \alpha_1 y_{k+1}^2 \left(\alpha_0 + \alpha_1 y_{k-1}^2 \right) \right), \end{aligned}$$

provided that such a solution exists.

In order to test for the presence of a VO or a LO, Hotta and Tsay (2012) propose the use of the Lagrange multiplier (LM) test statistic that, for a VO, is defined as

$$\mathrm{LM}_k^{\mathrm{VO}} = \frac{y_k^2}{\alpha_0 + \alpha_1 y_{k-1}^2},$$

while for a LO, it is given by

$$\mathrm{LM}_k^{\mathrm{LO}} = \mathrm{LM}_k^{\mathrm{VO}} \left\{ 1 + \alpha_1 h_k \left(\frac{1}{h_{k+1}} - \frac{y_{k+1}^2}{h_{k+1}^2} \right) \right\}^2 \left(1 + 2\alpha_1^2 h_k \frac{y_k^2}{h_{k+1}^2} \right)^{-1}.$$

Note that $\mathrm{LM}_k^{\mathrm{LO}} = \mathrm{LM}_k^{\mathrm{VO}}$ if $\alpha_1 = 0$. Therefore, the two test statistics should not differ substantially when α_1 is close to 0. To detect multiple outliers, Hotta and Tsay (2012) thus propose to compute the maximum LM statistics

$$\mathrm{LM}_{\max}^{\mathrm{VO}} = \max_{2 \le t \le n} \mathrm{LM}_k^{\mathrm{VO}}, \qquad \mathrm{LM}_{\max}^{\mathrm{LO}} = \max_{2 \le t \le n} \mathrm{LM}_k^{\mathrm{LO}},$$

for which it is easy to compute critical values via simulation. If both statistics are significant, one may choose the outlier that gives the smaller p-value.

15.3.3 Outliers in INGARCH Models

Fokianos and Fried (2010) consider outliers in the integer-valued GARCH (INGARCH) model given by

$$\begin{aligned} x_t \mid \mathcal{F}_{t-1}^x &\sim \mathrm{Poisson}(\lambda_t), \\ \lambda_t &= \alpha_0 + \sum_{j=1}^{p} \alpha_j \lambda_{t-j} + \sum_{i=1}^{q} \beta_i x_{t-i}, \end{aligned} \tag{15.6}$$

for $t \ge 1$, where λ_t is the Poisson intensity of the process x_t, $\mathcal{F}_{t-1}^x$ stands for the σ-algebra generated by $\{x_{t-1}, \ldots, x_{1-q}, \lambda_{t-1}, \ldots, \lambda_0\}$, α_0 is an intercept, $\alpha_j > 0$, for $j = 1, \ldots, p$ and $\beta_i > 0$, for $i = 1, \ldots, q$ and $\sum_{j=1}^p \alpha_j + \sum_{i=1}^q \beta_i < 1$ to get covariance stationarity. Outliers in the INGARCH model (15.6) can be written as

$$\begin{aligned} y_t \mid \mathcal{F}_{t-1}^y &\sim \mathrm{Poisson}(\kappa_t), \\ \kappa_t &= \alpha_0 + \sum_{j=1}^{p} \alpha_j \kappa_{t-j} + \sum_{i=1}^{q} \beta_i y_{t-i} + w(1 - \delta B)^{-1} I_t^{(k)}, \end{aligned} \tag{15.7}$$

for $t \ge 1$, where κ_t is the Poisson intensity of the process y_t, $\mathcal{F}_{t-1}^y$ stands for the σ-algebra generated by $\{y_{t-1}, \ldots, y_{1-q}, \kappa_{t-1}, \ldots, \kappa_0\}$, w is the size of the outlier, and $0 \leq \delta \leq 1$ is a parameter that controls the outlier effect. In particular, $\delta = 0$ corresponds to an spiky outlier (SO) that influences the process from time k on, but to a rapidly decaying extent provided that α_1 is not close to unity, $0 < \delta < 1$ corresponds to a transient shift (TS) that affects several consecutive observations although its effect becomes gradually smaller as time grows and, finally, $\delta = 1$ corresponds to a level shift (LS) that affects permanently the mean and the variance of the observed series.

Fokianos and Fried (2010) propose to estimate the outlier effect w via conditional maximum likelihood. Therefore, given the observed time series $y_1, \ldots, y_T$, the log-likelihood of the parameters of model (15.7), $\boldsymbol{\eta} = (\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, w)'$ conditional on $\mathcal{F}_0^y$ is given, up to a constant, by

$$\ell(\boldsymbol{\eta}) = \prod_{t=q+1}^{T} \big(y_t \log \kappa_t(\boldsymbol{\eta}) - \kappa_t(\boldsymbol{\eta}) \big) \tag{15.8}$$

with score function

$$\frac{\partial l(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \sum_{t=q+1}^{T} \left(\frac{y_t}{\kappa_t(\boldsymbol{\eta})} - 1 \right) \frac{\partial \kappa_t(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}.$$

In addition, the conditional information matrix for $\boldsymbol{\eta}$ is given by

$$\mathbf{G}(\boldsymbol{\eta}) = \sum_{t=q+1}^{T} \mathrm{Cov}\left[\frac{\partial l(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \,\middle|\, \mathcal{F}_{t-1}^y \right] = \sum_{t=1}^{T} \frac{1}{\kappa_t(\boldsymbol{\eta})} \left(\frac{\partial l(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \right) \left(\frac{\partial l(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \right)'.$$

Consequently, $\hat{w}$ is obtained from the ML estimate $\hat{\boldsymbol{\eta}}$ after maximizing the log-likelihood function (15.8). In order to test for the presence of an outlier at $t = k$, Fokianos and Fried (2010) propose to use the score test given by

$$T_k = \Delta' \mathbf{G}(\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_p, \tilde{\beta}_1, \ldots, \tilde{\beta}_q, 0)^{-1} \Delta,$$

where

$$\Delta = \frac{\partial l(\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_p, \tilde{\beta}_1, \ldots, \tilde{\beta}_q, 0)}{\partial \boldsymbol{\eta}},$$

and $(\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_p, \tilde{\beta}_1, \ldots, \tilde{\beta}_q, 0)'$ is the vector that contains the ML estimates of the parameters of the model (15.6) and the value $w = 0$. Under the null hypothesis of no outlier, T_k has an asymptotic χ_1^2 distribution. To detect an outlier of a certain type at an unknown time point, the idea is to obtain

$$T = \max_{q+1 \leq t \leq T} T_t,$$

and reject the null hypothesis of no outlier if T is large. The distribution of this statistic can be calibrated using bootstrap. Finally, to detect the presence of several outliers in a INGARCH time series, Fokianos and Fried (2010) proposed a procedure similar to that used in Chen and Liu (1993a).

15.4 Outliers in Multivariate Time Series Models

Outliers in multivariate time series has been much less analyzed than in the univariate case. Multivariate outliers were introduced in Tsay et al. (2000). These authors have also proposed a detection method based on individual and joint likelihood ratio statistics. Alternatively, Galeano et al. (2006) used projection pursuit methods to develop a procedure for detecting outliers. In particular, Galeano et al. (2006) showed that testing for outliers in certain projection directions can be more powerful than testing the multivariate series directly. In view of these findings, an iterative procedure to detect and handle multiple outliers based on a univariate search in these optimal directions were proposed. The main advantage of this procedure is that it can identify outliers without prespecifying a vector ARMA model for the data. An alternative method based on linear combinations of the components of the vector of time series can be found in Baragona and Battaglia (2007b) that considers an independent component approach. Finally, Baragona and Battaglia (2007a) have proposed a method to discover outliers in a dynamic factor model based on linear transforms of the observed time series. In this section, we briefly review the main findings in Tsay et al. (2000) and Galeano et al. (2006).

15.4.1 The Tsay, Peña and Pankratz Procedure

A r-dimensional vector time series $\mathbf{X}_t = (X_{1t}, \ldots, X_{rt})'$ follows a vector ARMA (VARMA) model if

$$\Phi(B)\mathbf{X}_t = \mathbf{C} + \Theta(B)\mathbf{E}_t, \quad t = 1, \ldots, T, \tag{15.9}$$

where $\Phi(B) = \mathbf{I} - \Phi_1 B - \cdots - \Phi_p B^p$ and $\Theta(B) = \mathbf{I} - \Theta_1 B - \cdots - \Theta_q B^q$ are $r \times r$ matrix polynomials of finite degrees p and q, $\mathbf{C}$ is a r-dimensional constant vector, and $\mathbf{E}_t = (E_{1t}, \ldots, E_{rt})'$ is a sequence of independent and identically distributed Gaussian random vectors with zero mean and positive-definite covariance matrix Σ. The autoregressive representation of the VARMA model in (15.9) is given by $\Pi(B)\mathbf{X}_t = \mathbf{C}_\Pi + \mathbf{E}_t$, where $\Phi(1)\mathbf{C}_\Pi = \mathbf{C}$ and $\Pi(B) = \Theta(B)^{-1}\Phi(B) = \mathbf{I} - \sum_{i=1}^{\infty} \Pi_i B^i$, while the moving-average representation of $\mathbf{X}_t$ is given by $\mathbf{X}_t = \mathbf{C}_\Psi + \Psi(B)\mathbf{E}_t$, where $\Phi(1)\mathbf{C}_\Psi = \mathbf{C}$ and $\Phi(B)\Psi(B) = \Theta(B)$ with $\Psi(B) = \mathbf{I} + \sum_{i=1}^{\infty} \Psi_i B^i$.

Tsay et al. (2000) generalize four types of univariate outliers to the vector case. Under the presence of a multivariate outlier, we observe a time series $\mathbf{Y} = (Y_1', \ldots, Y_T')'$, where $\mathbf{Y}_t = (Y_{1t}, \ldots, Y_{rt})'$, can be written as follows:

$$\mathbf{Y}_t = \mathbf{X}_t + \Lambda(B)\mathbf{w}I_t^{(k)}, \tag{15.10}$$

where $\mathbf{w} = (w_1, \ldots, w_r)'$ is the size of the outlier and $\mathbf{X}_t$ follows a VARMA model. The type of the outlier is defined by the matrix polynomial $\Lambda(B)$: for a multivariate innovational outlier (MIO), $\Lambda(B) = \Psi(B)$; for a multivariate additive outlier (MAO), $\Lambda(B) = \mathbf{I}$; for a multivariate level shift (MLS), $\Lambda(B) = (1 - B)^{-1}\mathbf{I}$;

and, finally, for a multivariate temporary (or transitory) change (MTC), $\Lambda(B) = (\mathbf{I} - \delta\mathbf{I}B)^{-1}$. In practice, an outlier may produce a complex effect, given by a linear combination of the previously discussed pure effects. Furthermore, different components of $\mathbf{X}_t$ may suffer different outlier effects. An example of this kind of mixed effects can be found in Galeano et al. (2006).

Given the parameters of the VARMA model for $\mathbf{X}_t$, the series of innovations are defined by $\mathbf{A}_t = \Pi(B)\mathbf{Y}_t - \mathbf{C}_\Pi$ and the relationship with the true innovations, $\mathbf{E}_t$, is given by

$$\mathbf{A}_t = \mathbf{E}_t + \Gamma(B)\mathbf{w}I_t^{(k)},$$

where $\Gamma(B) = \Pi(B)\Lambda(B) = \mathbf{I} - \sum_{i=1}^{\infty} \Gamma_i B^i$. Now, the least squares estimate of the size of an outlier of type i at time point k is given by

$$\mathbf{w}_{i,k} = -\left(\sum_{j=0}^{n-k} \Gamma_j' \Sigma^{-1} \Gamma_j\right)^{-1} \left(\sum_{j=0}^{n-k} \Gamma_j' \Sigma^{-1} \mathbf{A}_{k+j}\right),$$

where $\Gamma_0 = -\mathbf{I}$ and $i =$ MIO, MAO, MLS and MTC for subscripts, and has a covariance matrix given by $\Sigma_{i,k} = (\sum_{j=0}^{n-k} \Gamma_j' \Sigma^{-1} \Gamma_j)^{-1}$. The likelihood ratio test statistic for testing for the presence of a multivariate outlier of type i at $t = k$ is $J_{i,k} = \mathbf{w}_{i,k}' \Sigma_{i,k}^{-1} \mathbf{w}_{i,k}$. Under the null hypothesis of no outlier, $J_{i,k}$ has a χ_r^2 distribution. Tsay et al. (2000) also proposed a second statistic defined by $C_{i,k} = \max\{|w_{j,i,k}|/\sqrt{\sigma_{j,i,k}} : 1 \leq j \leq r\}$, where $w_{j,i,k}$ is the jth element of $\mathbf{w}_{i,k}$ and $\sigma_{j,i,k}$ is the jth element of the main diagonal of $\Sigma_{i,k}$, with the aim of look for outliers in individual components of the vector of series.

In practice, the parameter matrices are then substituted by their estimates and the following overall test statistics are defined:

$$J_{\max}(i, k_i) = \max_{1 \leq t \leq n} J_{i,t}, \qquad C_{\max}\left(i, k_i^*\right) = \max_{1 \leq t \leq n} C_{i,t},$$

where k_i and k_i^* denote respectively the time points at which the maximum of the joint test statistics and the maximum component statistics occur.

15.4.2 The Galeano, Peña and Tsay Procedure

Galeano et al. (2006) have proposed a method for detecting multivariate outliers in time series without requiring initial specification of the multivariate model. This is very important in these settings because model identification is quite complicated in the presence of outliers. The method is based on univariate outlier detection applied to some useful projections of the vector time series. The basic idea is simple: a multivariate outlier produces at least a univariate outlier in almost every projected series, and by detecting the univariate outliers we can identify the multivariate ones.

First, a non-zero linear combination of the components of the VARMA model in (15.9) follows a univariate ARMA model. Second, when the observed series $\mathbf{Y}_t$ is affected by an outlier, as in (15.10), the projected series $y_t = \mathbf{v}'\mathbf{Y}_t$ satisfies $y_t = x_t + \mathbf{v}'\Lambda(B)\mathbf{w}I_t^{(k)}$. Specifically, if $\mathbf{Y}_t$ has a MAO, the projected series is $y_t = x_t + \omega I_t^{(k)}$, so that it has an additive outlier of size $\omega = \mathbf{v}'\mathbf{w}$ at point $t = k$ provided that $\mathbf{v}'\mathbf{w} \neq 0$. Similarly, the projected series of a vector process with a MLS of size $\mathbf{w}$ will have a level shift with size $\omega = \mathbf{v}'\mathbf{w}$ at $t = k$. The same result also applies to MTC. A MIO can produce several effects. In particular, a MIO can lead to a patch of consecutive outliers with sizes $\mathbf{v}'\mathbf{w}, \mathbf{v}'\Psi_1\mathbf{w}, \ldots, \mathbf{v}'\Psi_{T-h}\mathbf{w}$, starting at $t = k$. Assuming that k is not close to T and because $\Psi_j \to 0$, the size of the outlier in the patch tends to zero. In the particular case that $\mathbf{v}'\Psi_i\mathbf{w} = \psi_i\mathbf{v}'\mathbf{w}$, for $i = 1, \ldots, T-k$, then y_t has an innovational outlier at $t = k$ with size $\beta = \mathbf{v}'\mathbf{w}$. However, if $\mathbf{v}'\Psi_i\mathbf{w} = 0$, for $i = 1, \ldots, T-k$, then y_t has an additive outlier at $t = k$ with size $\mathbf{v}'\mathbf{w}$, and if $\mathbf{v}'\Psi_i\mathbf{w} = \mathbf{v}'\mathbf{w}$, for $i = 0, \ldots, T-k$, then y_t has a level shift at $t = k$ with size $\beta = \mathbf{v}'\mathbf{w}$. Therefore, the univariate series y_t obtained by the projection can be affected by an additive outlier, a patch of outliers or a level shift.

Galeano et al. (2006) have shown that it is possible to identify multivariate outliers better by applying univariate test statistics to optimal projections than by using multivariate statistics on the original series. More precisely, it is possible to show that, in the presence of a multivariate outlier, the directions that maximize or minimize the kurtosis coefficient of the projected series include the direction of the outlier, that is, the direction that maximizes the ratio between the outlier size and the variance of the projected observations. Therefore, Galeano et al. (2006) proposed here a sequential procedure for outlier detection based on the directions that minimize and maximize the kurtosis coefficient of the projections. The procedure is divided into four steps: (1) obtain the optimal directions; (2) search for outliers in the projected univariate time series; (3) remove the effect of all detected outliers by using an approximated multivariate model; (4) iterate the previous steps applied to the cleaned series until no more outliers are found. It is important to note that in Step (2), the detection is carried out in two stages: first, MLS's are identified; second, MIO's, MAO's and MTC's are found. This is done in order to avoid confusions between multivariate innovational outliers and multivariate level shifts.

15.5 Conclusions

This chapter summarized outliers in both univariate and multivariate time series. Although many work has been done, more research is still needed in order to analyze outliers and unexpected events in time series. First, new effects in nonlinear time series models can be considered. For instance, level shifts in bilinear and SETAR models, transitory changes in GARCH models or additive outliers in INGARCH models would be of interest. Second, as far as we know, outlier detection in multivariate nonlinear time series models have been not considered yet. For instance, the extension of additive and volatility outliers to most of the available multivariate

GARCH models is almost straightforward. Likewise, outliers in multivariate bilinear and SETAR time series models is of interest. Finally, most of the existing literature on outlier detection focus on iterative testing procedures. Recently, Galeano and Peña (2012) have proposed a method to detect additive outliers by means of the use of a model selection criterion. The main advantage of the procedure is that all the outliers are detected in a single step. Although the computational cost of the procedure is high, the detection of outliers by means of model selection criteria is a promising line of research. Finally, this chapter is closely related with those by Barme–Delcroix (Chap. 3) who analyzes extreme events and Huskova (Chap. 11) who analyzes robust change point analysis.

Acknowledgements We gratefully acknowledge financial support by MCI grants SEJ2007/64500 and MTM2008-03010.

References

Baragona, R. (1998). Nonstationary time series, linear interpolators and outliers. *Statistica*, *58*, 375–394.

Baragona, R., & Battaglia, F. (2007a). Outliers in dynamic factor models. *Electronic Journal of Statistics*, *1*, 392–432.

Baragona, R., & Battaglia, F. (2007b). Outliers detection in multivariate time series by independent component analysis. *Neural Computation*, *19*, 1962–1984.

Battaglia, F., & Orfei, L. (2005). Outlier detection and estimation in nonlinear time series. *Journal of Time Series Analysis*, *26*, 107–121.

Battaglia, F., & Baragona, R. (1992). Linear interpolators and the outliers problem in time series. *Metron*, *50*, 79–97.

Box, G. E. P., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, *70*, 70–79.

Carnero, M. A., Peña, D., & Ruiz, E. (2007). Effects of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis*, *28*, 471–497.

Chang, I., Tiao, G. C., & Chen, C. (1988). Estimation of time series parameters in the presence of outliers. *Technometrics*, *30*, 193–204.

Chen, C. W. S. (1997). Detection of additive outliers in bilinear time series. *Computational Statistics & Data Analysis*, *24*, 283–294.

Chen, C., & Liu, L. M. (1993a). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, *88*, 284–297.

Chen, C., & Liu, L. M. (1993b). Forecasting time series with outliers. *Journal of Forecasting*, *12*, 13–35.

Davies, P. L., Fried, R., & Gather, U. (2004). Robust signal extraction for on-line monitoring data. *Journal of Statistical Planning and Inference*, *122*, 65–78.

Doornik, J. A., & Ooms, M. (2005). *Outlier detection in GARCH models*. 2005-W24 Nuffield economics working papers.

Fokianos, C., & Fried, R. (2010). Interventions in INGARCH processes. *Journal of Time Series Analysis*, *31*, 210–225.

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

Franses, P. H., & Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting*, *15*, 1–9.

Gagné, C., & Duchesne, P. (2008). On robust forecasting in dynamic vector time series models. *Journal of Statistical Planning and Inference*, *138*, 3927–3938.

Galeano, P., & Peña, D. (2012). Additive outlier detection in seasonal ARIMA models by a modified Bayesian information criterion. In W. R. Bell, S. H. Holan, & T. S. McElroy (Eds.), *Economic time series: modeling and seasonality* (pp. 317–336). Boca Raton: Chapman & Hall.

Galeano, P., Peña, D., & Tsay, R. S. (2006). Outlier detection in multivariate time series by projection pursuit. *Journal of the American Statistical Association*, *101*, 654–669.

Gather, U., Bauer, M., & Fried, R. (2002). The identification of multiple outliers in online monitoring data. *Estatistica*, *54*, 289–338.

Gelper, S., Schettlinger, K., Croux, C., & Gather, U. (2009). Robust online scale estimation in time series: a model-free approach. *Journal of Statistical Planning and Inference*, *139*, 335–349.

Grané, A., & Veiga, H. (2010). Wavelet-based detection of outliers in financial time series. *Computational Statistics & Data Analysis*, *54*, 2580–2593.

Haldrup, N., Montañes, A., & Sansó, A. (2011). Detection of additive outliers in seasonal time series. *Journal of Time Series Econometrics*, *3*, 2.

Hotta, L. K., & Tsay, R. S. (2012). Outliers in GARCH processes. In W. R. Bell, S. H. Holan, & T. S. McElroy (Eds.), *Economic time series: modeling and seasonality* (pp. 337–358). Boca Raton: Chapman & Hall.

Justel, A., Peña, D., & Tsay, R. S. (2001). Detection of outlier patches in autoregressive time series. *Statistica Sinica*, *11*, 651–673.

Kirkendall, N. J. (1992). Monitoring for outliers and level shifts in Kalman filter implementations of exponential smoothing. *Journal of Forecasting*, *11*, 543–560.

Koehler, A. B., Snyder, R. D., Ord, J. K., & Beaumont, A. (2012). A study of outliers in the exponential smoothing approach to forecasting. *International Journal of Forecasting*, *28*, 477–484.

Ljung, G. M. (1993). On outlier detection in time series. *Journal of the Royal Statistical Society. Series B. Methodological*, *55*, 559–567.

Luceño, A. (1998). Detecting possibly non-consecutive outliers in industrial time series. *Journal of the Royal Statistical Society. Series B. Methodological*, *60*, 295–310.

Maronna, R., Martin, R. D., & Yohai, V. (2006). *Robust statistics*. Chichester: Wiley.

Muler, N., Peña, D., & Yohai, V. (2009). Robust estimation for ARMA models. *The Annals of Statistics*, *37*, 816–840.

Muler, N., & Yohai, V. (2008). Robust estimates for GARCH models. *Journal of Statistical Planning and Inference*, *138*, 2918–2940.

Pankratz, A. (1993). Detecting and treating outliers in dynamic regression models. *Biometrika*, *80*, 847–854.

Penzer, J. (2007). State space models for time series with patches of unusual observations. *Journal of Time Series Analysis*, *28*, 629–645.

Peña, D., & Maravall, A. (1991). Interpolation, outliers and inverse autocorrelations. *Communications in Statistics. Theory and Methods*, *20*, 3175–3186.

Perron, P., & Rodríguez, G. (2003). Searching for additive outliers in nonstationary time series. *Journal of Time Series Analysis*, *24*, 193–220.

Sánchez, M. J., & Peña, D. (2010). The identification of multiple outliers in ARIMA models. *Communications in Statistics. Theory and Methods*, *32*, 1265–1287.

Tsay, R. S. (1986). Time series model specification in the presence of outliers. *Journal of the American Statistical Association*, *86*, 132–141.

Tsay, R. S. (1988). Outliers, level shifts and variance changes in time series. *Journal of Forecasting*, *7*, 1–20.

Tsay, R. S., Peña, D., & Pankratz, A. E. (2000). Outliers in multivariate time series. *Biometrika*, *87*, 789–804.