Journal of Multivariate Analysis 101 (2010) 2168–2183

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Robust estimation of periodic autoregressive processes in the presence of additive outliers

A.J.Q. Sarnaglia [a], V.A. Reisen [a], C. Lévy-Leduc [b,*]

[a] *Departamento de Estatística, Universidade Federal do Espírito Santo, Vitória/ES, Brazil*

[b] *CNRS/LTCI/TelecomParisTech - 46, rue Barrault, 75634 Paris Cédex 13, France*

A R T I C L E I N F O

Article history:
Received 20 August 2009
Available online 4 June 2010

AMS subject classifications:
62M10
62E20
62F12
62F35

Keywords:
Additive outliers
PAR model
Periodicity
Robustness
Influence function

A B S T R A C T

This paper suggests a robust estimation procedure for the parameters of the periodic AR (PAR) models when the data contains additive outliers. The proposed robust methodology is an extension of the robust scale and covariance functions given in, respectively, Rousseeuw and Croux (1993) [28], and Ma and Genton (2000) [23] to accommodate periodicity. These periodic robust functions are used in the Yule–Walker equations to obtain robust parameter estimates. The asymptotic central limit theorems of the estimators are established, and an extensive Monte Carlo experiment is conducted to evaluate the performance of the robust methodology for periodic time series with finite sample sizes. The quarterly Fraser River data was used as an example of application of the proposed robust methodology. All the results presented here give strong motivation to use the methodology in practical situations in which periodically correlated time series contain additive outliers.

1. Introduction

Often, time series drawn from economics, hydrology, climatology and air pollution among others areas exhibit periodically varying covariances (also known as cyclostationary time series) in the sense that the mean and covariances are periodic functions of the time. An important linear time series model whose parameters vary as functions of time is called the periodic autoregressive moving average (PARMA) model, and, in recent years, it has been widely used for modeling periodically correlated time series. Hurd and Gerr [16], Vecchia and Ballerini [37] and Hurd and Miamee [17] discussed the procedures to detect periodic correlations in time series. Lund and Basawa [20] showed that PAR models can be represented in a vector form through the Vector Autoregressive (VAR) models. The condition for the VAR process to be stationary is described in detail by Lütkepohl [22]. Ursu and Duchesne [33] derived the asymptotical distributions of the least square estimators of the parameters of the vector periodic AR time series models (PVAR), and they also introduced a portmanteau test statistic for diagnosing the adequacy of the PVAR models. For a recent review of the PAR models, see the introduction section of Ursu and Duchesne [33]. In the context of applied frameworks, Franses and Paap [13] used the PAR processes to model several quarterly UK macroeconomic series. McLeod [24] made some comparisons between some river flow series, see, also, [3,21,15] among others.

There are several ways to estimate the autoregressive parameters of PAR models. McLeod [25] suggested the use of the sample periodic autocorrelation function in the periodic Yule–Walker equations to obtain the estimates. The Yule–Walker estimation method was also considered in [1]. In this paper, the authors also proposed an innovation algorithm to obtain the parameter estimates of the PARMA model, and the asymptotic results of the estimates were established. The asymptotic

* Corresponding author.
E-mail address: celine.levy-leduc@telecom-paristech.fr (C. Lévy-Leduc).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.05.006

properties of least-square estimators were discussed in [31]. Vecchia [35,36] studied maximum likelihood estimation in univariate PARMA models, and also discussed the exact likelihood function for Gaussian processes.

Atypical observations (outliers) are quite common in practical situations, and, more recently, many researchers have paid attention to model time series with outliers. There are several types of outliers which cause different effects on the estimates. However, in general, the following three types are usually considered [10]: innovation outliers (IO), which affect all subsequent observations; additive outliers (AO) or replacement outliers (RO), which have no effect on subsequent observations. AO outliers affect the parameter estimates more than IO, and they have the same effect as RO [23].

Several authors have investigated the outliers' effect on estimating the parameters of autoregressive integrated moving-average (ARIMA) models. For example, Ledolter [18] showed that the forecast intervals are very sensitive to the presence of additive outliers; Chang et al. [8] and Chen and Liu [9] showed that the estimated parameters of the ARMA model are more biased when the data contain outliers; Deutsch et al [11] and Chan [6,7] derived the autocorrelation function bias induced by the presence of outliers. The effects of outliers on parameter estimation have also been studied in the case of processes with a strong correlation structure, see [12] where a robust memory parameter estimator was introduced in the case of the Fractionally ARMA (ARFIMA) model.

The study of robustness in periodic processes is still in its infancy. As far as the authors of this paper are aware, this topic has not been much explored yet in the literature. Shao [30] discussed a robust estimation procedure for PAR models.

In spite of the restricted literature in this topic, periodic time series with atypical observations are quite common in many fields of application. Hence, it is an interesting research area in which to work. Here, PAR models are studied in the case of periodic time series with additive outliers. The paper proposes a robust estimator for the autoregressive parameters of PAR models. The robust estimation procedure is a generalization of the robust scale estimator and autocovariance functions proposed by, respectively, Rousseeuw and Croux [28] and Ma and Genton [23], to accommodate periodicity. Based on these methodologies, robust estimators of the periodic autocovariance functions are derived, and they are used to obtain consistent periodic Yule–Walker parameter estimates. Asymptotic central limit theorems of the robust estimators are established. Based on Monte Carlo experiments, an extensive simulation exercise is provided to show the effect of additive outliers in the order identification AIC and BIC criterions and also to analyze the empirical robustness properties of the proposed methodology.

The article is structured as follows: In Section 2, some preliminaries of the periodic processes under outliers are presented. Section 3 deals with the estimation and model selection tools. The impact of outliers on the correlation structure of the processes is the motivation of Section 4. In Section 5 the robust model estimation is given, and Monte Carlo experiments are in Section 6. The quarterly Fraser River data is analyzed with the robust and classical correlation approaches, and it is discussed in Section 7. Finally, some remarks and conclusions are stated in Section 8.

2. Periodic correlated processes

Let $\{Y_t\}_{t\geq 1}$ be a stochastic process with seasonal characteristics of period s. The index t can be written, by integer division, as $t = t(r, m) = (r-1)s + m$, where $m = 1, \ldots, s$ and $r = 1, 2, \ldots$. For example, in the case of monthly data, $s = 12$, m and r are, respectively, the month and the year.

In what follows, definitions and properties of the process $\{Y_{t(r,m)}\}$ are discussed. Let $\gamma_Y^{(r,m)}$ be the autocovariance of the process $\{Y_{t(r,m)}\}$, that is,

$$\gamma_Y^{(r,m)}(h) = \mathrm{Cov}\left[Y_{t(r,m)}, Y_{t(r,m)-h}\right],$$

for some non negative integer h.

Definition 2.1. $\{Y_{t(r,m)}\}_{r\geq 1,\ m=1,\ldots,s}$ is said to be a periodic stationary process or a periodically correlated process if, for all $r \geq 1$ and m in $\{1, \ldots, s\}$,

(i) $\mathbb{E}[Y^2_{t(r,m)}] < \infty$,
(ii) $\mu_{r,m} = \mathbb{E}[Y_{t(r,m)}] = \mu_m$,
(iii) $\gamma_Y^{(r,m)}(h) = \gamma_Y^{(m)}(h)$, for all non negative integer h,

that is, if $\mu_{r,m}$ and $\gamma_Y^{(r,m)}(.)$ are finite and if they do not depend on r.

Based on the above definition, the autocorrelation function of the process $\{Y_{t(r,m)}\}$ is straightforwardly defined by

$$\rho_Y^{(r,m)}(h) = \frac{\gamma_Y^{(r,m)}(h)}{\sqrt{\gamma_Y^{(r,m)}(0)\gamma_Y^{(r,m-h)}(0)}}, \quad h \geq 0.$$

Remark 1. Note that, if $\{Y_{t(r,m)}\}_{r\geq 1, m=1,\ldots,s}$ is a periodically correlated process then $\gamma_Y^{(m)}(h)$ and $\rho_Y^{(m)}(h)$ are periodic functions with period s, that is, $\gamma_Y^{(m)}(h) = \gamma_Y^{(m+ks)}(h)$ and $\rho_Y^{(m)}(h) = \rho_Y^{(m+ks)}(h)$, for all non negative integer k. If $s = 1$, the condition of periodic stationarity is equivalent to the usual condition for homogeneous processes [32].

Hereafter, $\gamma_Y^{(m)}(.)$ and $\rho_Y^{(m)}(.)$ are, respectively, denoted by PeACV and PeACF.

As previously mentioned, the main purpose of this paper is to study the influence of additive outliers in PAR models and also to propose a robust procedure to estimate model parameters. The following definition introduces the model for periodically stationary processes with additive outliers.

Definition 2.2. $\{Z_{t(r,m)}\}_{r\geq 1, m=1,\ldots,s}$ is a periodic stationary process with additive outliers if it satisfies

$$Z_{t(r,m)} = Y_{t(r,m)} + \sum_{i=1}^{\ell} \left(\omega_i V_i^{t(r,m)}\right) \tag{1}$$

where $\{Y_{t(r,m)}\}_{r\geq 1, m=1,\ldots,s}$ is a periodic stationary process and, for all $r \geq 1$ and $m = 1, \ldots, s$, $(V_i^{t(r,m)})_{1\leq i\leq \ell}$ are independent random variables with $\mathbb{P}(V_i^{t(r,m)} = -1) = \mathbb{P}(V_i^{t(r,m)} = 1) = p_i/2$ and $\mathbb{P}(V_i^{t(r,m)} = 0) = 1 - p_i$, $p_i \in (0, 1)$, for all i. Also, for all r and m, $(V_i^{t(r,m)})_{1\leq i\leq l}$ and $\{Y_{t(r,m)}\}_{r\geq 1, m=1,\ldots,s}$ are independent variables and ℓ and ω_i are, respectively, the maximal number of outliers at time $t(r, m)$ and the magnitude of the ith outlier.

Note that, according to Definition 2.2, the presence of outliers is random. This assumption is quite plausible since, in general, the occurrence of atypical observations in time series is not deterministic.

Several models have been proposed and studied to adjust time series generated from periodic stationary processes, see [3,21,15] for application of these models to real data. The Periodic Autoregressive (PAR) process, defined below, is a generalization of the Autoregressive (AR) process [4,5], and it is widely used for modeling periodic time series.

Definition 2.3. The periodic stationary process $\{Y_{t(r,m)}\}_{r\geq 1, m=1,\ldots,s}$ is a PAR process if and only if $\{Y_{t(r,m)}\}$ is a solution of the difference equations given by

$$Y_{t(r,m)} = \mu_m + \sum_{i=1}^{p_m} \left\{\phi_i^{(m)} \left[Y_{t(r,m)-i} - \mu_{m-i}\right]\right\} + \epsilon_{t(r,m)} \tag{2}$$

where, $(\epsilon_{t(r,m)})_{r\geq 1} \sim iid(0, \sigma^2_{(m)})$ and for all $r, r' \geq 1$ and $m, m' = 1, \ldots, s$, $\epsilon_{t(r,m)}$ and $\epsilon_{t(r',m')}$ are independent, p_m is the order of the autoregressive polynomial at m, and $\phi_j^{(m)}, j = 1, \ldots, p_m$, are the autoregressive parameters of the process.

Note that, the parameters $(\phi_j^{(m)})_{1\leq j\leq p_m}$ in Eq. (2) are also periodic functions of period s, that is, $\phi_j^{(m)} = \phi_j^{(m+ks)}$, for all j in $\{1, \ldots, p_m\}$ and for all non negative k.

Under the causality assumption, which is stronger than the usual stationarity condition, the solution of Eq. (2) can be also written in the following way

$$Y_{t(r,m)} - \mu_m = \sum_{j\geq 0} \psi_j(m)\epsilon_{t(r,m)-j}, \quad \text{where } \max_{1\leq m\leq s} \sum_{j\geq 0} |\psi_j(m)| < \infty.$$

In what follows, a sufficient condition is given to ensure the causality property of the PAR process. Following [20], the PAR process presented can be written in a s-variate AR form

$$\left(\Phi_0 - \Phi_1 B - \cdots - \Phi_P B^P\right)\left(\overrightarrow{Y}_r - \overrightarrow{\mu}\right) = \overrightarrow{\Upsilon}_t$$

where B is the backshift operator ($B^k Y_t = Y_{t-k}$), $\overrightarrow{Y}_r = (Y_{t(r,1)}, \ldots, Y_{t(r,s)})'$, $\overrightarrow{\mu} = (\mu_1, \ldots, \mu_s)'$ and $\overrightarrow{\Upsilon}_t = (\epsilon_{t(r,1)}, \ldots, \epsilon_{t(r,s)})'$ ($'$ means the transpose matrix). The order $P = \lceil p/s\rceil$, $p = \max_{1\leq m\leq s} p_m$ and $\lceil x\rceil$ denotes the smallest integer greater than or equal to x. The AR coefficients (matrices with dimension $s \times s$) have entries given by

$$(\Phi_0)_{i,j} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i < j, \\ \phi_{i-j}^{(i)} & \text{if } i > j \end{cases} \qquad (\Phi_k)_{i,j} = \phi_{ks+i-j}^{(i)}, \quad \text{for } 1 \leq k \leq P.$$

A sufficient condition to guarantee the causality property of the PAR process is

$$\det\left(\Phi_0 - \sum_{k=1}^{P} \Phi_k z^k\right) \neq 0, \tag{3}$$

for all complex z satisfying $|z| \leq 1$. In the PAR(1) model, this condition is reduced to

$$\lambda = \left|\prod_{m=1}^{s} \phi_1^{(m)}\right| < 1. \tag{4}$$

3. Usual procedure for parameter estimation and model selection

The adjustment of a time series model to a data set should follow three important steps: estimation, diagnostic checking and model selection. This paper addresses the steps of identifying and estimating the PAR process, and these are discussed below. An empirical investigation of the methodology is given in Section 6.

3.1. Estimation of PeACV and PeACF

As mentioned in the previous section, the functions PeACV and PeACF are very important tools in the model estimation and identification steps. However, both functions, in general, are not known and need to be estimated. Let $\{y_t\}_{1\leq t\leq n}$ be a sample of n observations from the periodic stationary process $\{Y_t\}_{t\geq 1}$. For simplicity, it will be assumed that $n/s = N \in \mathbb{N}^\star$, that is, the series has s full periods. Using the same notation as previously, the sample of n observations can thus be written as $\{y_{t(r,m)}\}_{r=1,\ldots,N,m=1,\ldots,s}$.

McLeod [25] suggested estimating the functions PeACF and PeACV using the sample PeACF and PeACV given, respectively by

$$\widehat{\rho}_y^{(m)}(h) = \frac{\widehat{\gamma}_y^{(m)}(h)}{\sqrt{\widehat{\gamma}_y^{(m)}(0)\widehat{\gamma}_y^{(m-h)}(0)}}$$

and

$$\widehat{\gamma}_y^{(m)}(h) = \frac{1}{N}\sum_{r=r^*}^{N}\left\{\left(y_{t(r,m)} - \overline{y}_{(m)}\right)\left(y_{t(r,m)-h} - \overline{y}_{(m-h)}\right)\right\}, \quad h \geq 0,$$

where $\overline{y}_{(m)} = \frac{1}{N}\sum_{r=1}^{N} y_{t(r,m)}$ is the sample mean at m, and r^* is the smallest r such that $t(r, m) - h > 0$. These estimators have the same properties as the sample autocovariance and autocorrelation functions for homogeneous stationary time series [26].

3.2. Model estimation

Suppose the orders $(p_m)_{m=1,\ldots,s}$ of the PAR model (2) are known. Here, following the procedure given in [25], the parameters $(\phi_i^{(m)})_{1\leq i\leq p_m}$ are estimated by using the Yule–Walker equations

$$\sum_{i=1}^{p_m} \phi_i^{(m)}\gamma_Y^{(m-i)}(k-i) = \gamma_Y^{(m)}(k), \quad k = 1, \ldots, p_m, \tag{5}$$

where the functions $\gamma_Y^{(m)}(k)$, $1 \leq k \leq p_m$, $1 \leq m \leq s$, are replaced by their estimates previously defined.

It is well-known that for the AR(p) processes, the Yule–Walker estimators are consistent and asymptotically Gaussian distributed, see [5]. The equivalent property also holds for PAR processes and it is given in the following proposition.

Proposition 1. *Assume that $\{Y_t\}_{t\geq 1}$ is a causal mean-zero PAR process with a finite fourth moment satisfying Definition 2.3 with $\mu_m = 0$, for all $m = 1, \ldots, s$. Let $\hat{\boldsymbol{\phi}}$ be the Yule–Walker estimators of $\boldsymbol{\phi} = (\phi_1^{(1)}, \ldots, \phi_p^{(1)}, \ldots, \phi_1^{(s)}, \ldots, \phi_p^{(s)})'$. Then, as N tends to infinity,*

$$\sqrt{N}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}) \xrightarrow{d} \mathcal{N}(0, F^{-1}),$$

where $F = (F_{m_1,m_2,k,l})_{1\leq m_1,m_2\leq s,1\leq k,l\leq p}$, with $F_{m_1,m_2,k,l} = \mathbb{E}\left[\left(\frac{\partial\varepsilon_i^\star}{\partial\phi_k^{(m_1)}}\right)\left(\frac{\partial\varepsilon_i^\star}{\partial\phi_l^{(m_2)}}\right)\right]$, $p = \max_{1\leq m\leq s} p_m$ and $\varepsilon_i^\star = (\varepsilon_{t(i,1)}/\sigma_{(1)}, \ldots, \varepsilon_{t(i,s)}/\sigma_{(s)})$ is the mean zero periodic standardized white noise process.

The proof of this proposition is postponed to Section 9.

3.3. Model selection

In general, the order of the PAR model is not known. One way to identify the order is to use the periodic Akaike Information Criteria (AIC) and the periodic Bayesian Information Criteria (BIC) [25,29]. These quantities are computed as follows

$$\mathrm{AIC} = \sum_{m=1}^{s}\mathrm{AIC}_m, \qquad \mathrm{AIC}_m = N\ln\left(\widehat{\sigma}_m^2\right) + 2p, \quad m = 1, \ldots, s$$

and

$$\mathrm{BIC} = \sum_{m=1}^{s}\mathrm{BIC}_m, \qquad \mathrm{BIC}_m = N\ln\left(\widehat{\sigma}_m^2\right) + p\ln(N), \quad m = 1, \ldots, s$$

where $\widehat{\sigma}_m^2 = RSS/N$ and RSS is the residual sum of squares. Other approaches are discussed in [14].

4. Impact of additive outliers on the periodic correlation structure

As described in the introduction, outliers can affect the dependence structure of a time series and also the estimation of the parameters and the identification of the model. In this section, some results related to the effects of additive outliers on the correlation structure of a periodic correlated process are derived.

Proposition 2. *Assume that* $\{Z_{t(r,m)}\}_{r\geq 1, m=1,\ldots,s}$ *is a periodic stationary process with additive outliers as described in Definition 2.2. The PeACV of* $\{Z_{t(r,m)}\}$ *is given, for all m in* $\{1,\ldots,s\}$, *by*

$$\gamma_Z^{(m)}(h) = \begin{cases} \gamma_Y^{(m)}(0) + \sum_{i=1}^{l} \omega_i^2 p_i, & \text{if } h = 0; \\ \gamma_Y^{(m)}(h), & \text{if } h \in \mathbb{N}^\star. \end{cases}$$

A straightforward consequence of Proposition 2 is given in the following corollary.

Corollary 3. *Assume that* $\{Z_{t(r,m)}\}_{r\geq 1, m=1,\ldots,s}$ *is a periodic stationary process with additive outliers as described in Definition 2.2. Then,*

1. *for all non negative integer h and for all m in* $\{1,\ldots,s\}$,
$$|\rho_Z^{(m)}(h)| \leq |\rho_Y^{(m)}(h)|$$
2. $\lim_{\omega_i\to\infty} \rho_Z^{(m)}(h) = 0$, *for any* $i = 1,\ldots,l$ *and all positive h.*

Corollary 3 shows that additive outliers introduce memory loss in the process $\{Z_{t(r,m)}\}$. This leads to estimates with significant negative bias, as shown in Section 6.

In the case of the sample PeACV, the effect of additive outliers is given in the following proposition.

Proposition 4. *Let* $\{z_{t(r,m)}\}_{r=1,\ldots,N, m=1,\ldots,s}$ *be a time series observed from the model given in Definition 2.2 where* $\{Y_{t(r,m)}\}$ *is a causal process with finite fourth moment. Then, for all m in* $\{1,\ldots,s\}$, *as N tends to infinity,*

$$\widehat{\gamma}_z^{(m)}(0) - \widehat{\gamma}_y^{(m)}(0) - \sum_{i=1}^{l} \omega_i^2 p_i = o_P(1),$$

and

$$\widehat{\gamma}_z^{(m)}(h) - \widehat{\gamma}_y^{(m)}(h) = o_P(1), \quad \text{if } h \neq 0,$$

where $\widehat{\gamma}_z^{(m)}(.)$ *and* $\widehat{\gamma}_y^{(m)}(.)$ *are, respectively the sample PeACV of* $\{z_{t(r,m)}\}$ *and* $\{y_{t(r,m)}\}$.

Propositions 2 and 4 show that the additive outliers can affect the statistical properties of the parameter estimates in periodic processes by the fact that these observations lead to an underestimation of the true correlation structure. In this context, it is necessary to use robust methods for estimating models of time series with outliers. This is the motivation of the next section.

5. Robust procedure to estimate parameters and select a model

In this section, a robust estimator of the PeACV is proposed when the PAR process has additive outliers. This robust estimator is based on the robust scale estimator proposed by Rousseeuw and Croux [28] which has already been widely used in different frameworks of non periodic correlated processes, see for instance [23,12,19]. The main contribution of this section is to provide central limit theorems for the robust scale and PeACV estimators when the underlying time series follows a PAR process.

5.1. Estimation of PeACV based on a robust scale function

Let $X_1,\ldots,X_n$ be random variables having a common cumulative distribution function F, and $\mathbf{x} = (X_1,\ldots,X_n)$. As explained in [19], a robust estimator of the autocovariance function of $\mathbf{x}$ can be achieved by using a robust scale estimator, hereafter denoted by $Q_n(\mathbf{x})$. Let the functionals T_1 and T_2 be defined as follows:

$$T_1 : \mathcal{M}(\mathbb{R}) \to \mathcal{F}(\mathbb{R})$$
$$F \mapsto \left\{ r \mapsto \iint \mathbb{1}_{\{|x-y|\leq r\}} \mathrm{d}F(x)\mathrm{d}F(y) \right\}, \tag{6}$$
$$T_2 : \mathcal{F}(\mathbb{R}) \to \mathbb{R}$$
$$f \mapsto f^{-1}(1/4). \tag{7}$$

In the above, $\mathcal{M}(\mathbb{R})$ denotes the set of cumulative distribution functions on $\mathbb{R}$ equipped with the topology of uniform convergence, and $\mathcal{F}(\mathbb{R})$ is the set of all increasing bounded cadlag functions, equipped with the topology of uniform convergence.

Let T_0 be the functional defined by

$$T_0 = T_2 \circ T_1. \tag{8}$$

In the following, F is assumed to belong to a location-scale family of the following type:

$$\{F_{\theta,\sigma}(.) = F_{0,1}((. - \theta)/\sigma), \theta \in \mathbb{R}, \sigma \in \mathbb{R}_+^*\},$$

where $F_{0,1}$ is a standard cumulative distribution function. Let F_n denote the empirical distribution function based on $(X_i)_{1\le i\le n}$ then $Q_n(\mathbf{x})$ is defined by

$$Q_n(\mathbf{x}) = cT_0(F_n), \tag{9}$$

where c is given by

$$c = 1/T_0(F_{0,1}). \tag{10}$$

The value of c ensures that the estimator Q_n is Fisher-consistent *i.e.* $cT_0(F_{\theta,\sigma}) = \sigma$, for all $\theta \in \mathbb{R}$ and $\sigma > 0$. In the case of Gaussian random variables [28] suggest to take $c = 2.2191$.

In what follows, under Assumption (A1), Proposition 5 establishes a central limit theorem for $Q_N^{(m)}(\mathbf{y}^{(m)})$ defined for any m in $\{1, \dots, s\}$ by

$$Q_N^{(m)}(\mathbf{y}^{(m)}) = Q_N(\{Y_{t(r,m)}\}_{1\le r\le N}), \tag{11}$$

where $\{Y_{t(r,m)}\}_{1\le r\le N}$ is a periodic correlated process given in Definition 2.1.

(A1) $(Y_{t(i,m)})_{i\ge 1}$ is a mean-zero Gaussian process with strong mixing coefficients α_N satisfying:

$$\exists a > 1 \quad \text{and} \quad C \ge 1, \qquad \alpha_N \le CN^{-a}.$$

The common cumulative distribution function of $\{Y_{t(i,m)}\}_{1\le i\le N}$ is denoted by $F^{(m)}$.

Proposition 5. *Under Assumption* (A1), $Q_N^{(m)}(.)$ *satisfies the following central limit theorem:*

$$\sqrt{N}(Q_N^{(m)} - Q^{(m)}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tilde{\sigma}^2),$$

where $Q^{(m)} = cT_0(F^{(m)})$, *the functional* T_0 *and the constant* c *are defined in* (8) *and* (10), *respectively. In the above* $\mathcal{D}$ *denotes the convergence in distribution,*

$$\tilde{\sigma}^2 = \mathbb{E}[\mathrm{IF}(Y_{t(1,m)}, Q^{(m)}, F^{(m)})^2] + 2\sum_{k\ge 1}\mathbb{E}[\mathrm{IF}(Y_{t(1,m)}, Q^m, F^{(m)})\mathrm{IF}(Y_{t(k+1,m)}, Q^m, F^{(m)})],$$

and

$$\mathrm{IF}(x, Q^{(m)}, F^{(m)}) = c\left(\frac{1/4 - F^{(m)}(x + T_0(F^{(m)})) + F^{(m)}(x - T_0(F^{(m)}))}{\int f^{(m)}(y)f^{(m)}(y + T_0(F^{(m)}))\mathrm{d}y}\right), \tag{12}$$

where $f^{(m)}$ *denotes the probability density function of* $F^{(m)}$.

See Section 9 for the proof of Proposition 5.

Following a similar idea as given in [23], the robust estimator of the autocovariance function is defined for a PAR process by

$$\hat{\gamma}_Q^{(m)}(h) = \frac{1}{4}\left[Q^2_{N-r^\star+1}(u^{(m)} + v^{(m)}) - Q^2_{N-r^\star+1}(u^{(m)} - v^{(m)})\right], \quad 0 \le h < N, \tag{13}$$

where $u^{(m)} = \{Y_{t(i,m)-h}\}_{r^\star\le i\le N}$, $v^{(m)} = \{Y_{t(i,m)}\}_{r^\star\le i\le N}$ and $r^\star$ is the smallest integer such that $t(r, m) - h > 0$.

A central limit theorem for $\hat{\gamma}_Q^{(m)}(.)$, $m = 1, \dots, s$, is established in the following proposition.

Proposition 6. *Let* h *be a non negative integer and* m *be an integer in* $\{1, \dots, s\}$. *Assume that* (A1) *holds for* $\{Y_{t(i,m)-h}\}_{i\ge r^\star}$ *and for* $\{Y_{t(i,m)}\}_{i\ge r^\star}$ *where* $r^\star$ *is the smallest integer such that* $t(r, m) - h > 0$. *Denote by* $F_+^{(m)}$ *the common c.d.f of* $\{Y_{t(i,m)-h} + Y_{t(i,m)}\}_{i\ge r^\star}$ *and by* $F_-^{(m)}$ *the common c.d.f of* $\{Y_{t(i,m)-h} - Y_{t(i,m)}\}_{i\ge r^\star}$. *Then,* $\hat{\gamma}_Q^{(m)}(h)$ *satisfies the following central limit theorem:*

$$\sqrt{N}\left(\hat{\gamma}_Q^{(m)}(h) - \gamma^{(m)}(h)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \check{\sigma}_h^2),$$

where $\gamma^{(m)}(h) = \mathrm{Cov}[Y_{t(r^\star,m)-h}, Y_{t(r^\star,m)}]$ *and*

$$\check{\sigma}_h^2 = \mathbb{E}[\psi(Y_{t(r^\star,m)-h}, Y_{t(r^\star,m)})^2] + 2\sum_{k\ge 1}\mathbb{E}[\psi(Y_{t(r^\star,m)-h}, Y_{t(r^\star,m)})\psi(Y_{t(r^\star+k,m)-h}, Y_{t(r^\star+k,m)})],$$

Table 1
Parameters of PAR(1) models used in the simulation.

Parameter	Model 1	Model 2	Model 3
$\phi_1^{(1)}$	0.25	0.9	1.5
$\phi_1^{(2)}$	−0.45	0.8	0.8
$\phi_1^{(3)}$	−0.15	0.7	1.2
$\phi_1^{(4)}$	0.35	0.6	0.5
λ	0.0059	0.3024	0.7200

where the function ψ is defined by

$$\psi : (x, y) \mapsto \frac{1}{2}\left\{Q(F_+^{(m)})\mathrm{IF}(x+y, Q^{(m)}, F_+^{(m)}) - Q(F_-^{(m)})\mathrm{IF}(x-y, Q^{(m)}, F_-^{(m)})\right\}. \tag{14}$$

In the above, IF *is defined in* (12).

The proof of Proposition 6 is postponed to Section 9.

5.2. Robust model estimation

In this paper, robust estimators of the PAR parameters, denoted hereafter by $(\tilde{\phi}_i^{(m)})_{1\le i\le p_m, m=1,\ldots,s}$, are obtained by replacing the classical autocovariance in the Yule–Walker equations (5) by the robust autocovariance estimator $\hat{\gamma}_Q^{(m)}(.)$ as follows

$$\sum_{i=1}^{p_m} \phi_i^{(m)} \hat{\gamma}_Q^{(m-i)}(k-i) = \hat{\gamma}_Q^{(m)}(k), \quad k = 1, \ldots, p_m,\ m = 1, \ldots, s. \tag{15}$$

Note that this system of equations does not depend on the constant c since both sides are multiplied by c^2.

Proposition 7. *Under assumptions of Proposition 6, the robust Yule–Walker estimators $(\tilde{\phi}_i^{(m)})_{1\le i\le p_m, m=1,\ldots,s}$ satisfy*

$$\tilde{\phi}_i^{(m)} - \phi_i^{(m)} = O_P(N^{-1/2})$$

for all $i = 1, \ldots, p_m$ and m in $\{1, \ldots, s\}$.

The proof of this proposition is detailed in Section 9.

The model identification is also carried out using the AIC or the BIC criterion, however the residual sum of squares (RSS) is now computed using the robust estimators obtained from (15).

The empirical properties of the theory previously presented are the main motivation of the next section.

6. Monte Carlo studies

This section reports on the results of several Monte Carlo experiments to analyze the finite sample properties of the robust methodologies discussed previously, and the results are presented in Sections 6.1 and 6.2. In the former, the identification criterions AIC and BIC are investigated to verify the empirical proportion of correct model selection. The performance of model estimation procedures is the motivation of the second subsection. In all cases investigated, the sample autocovariance based on the moments, discussed in Section 3, is also considered for comparison purposes.

The empirical results were obtained under 10,000 replications of series of size $n = 400$ of PAR(1) processes generated according to Definitions 2.2 and 2.3, with the following data generating specifications: $s = 4$, $\mu_m \equiv 0$, and the probability and the magnitudes of the outliers are $p = 1\%$ and $\omega = 0$, 4 and 7,respectively. $\epsilon_{t(r,m)}$ are i.i.d $N(0, 1)$, for all r and m.

All calculations were obtained using Fortran codes, and the code given in [28] was adapted to obtain the robust covariance estimates. The models and their parameter values are specified in Table 1. The parameter values were chosen to have examples of time series models with low (Model 1), moderate (Model 2) and strong (Model 3) correlation dependencies.

6.1. Frequency rates of the correct model selection using the AIC and the BIC criterions

The empirical investigation of the AIC and BIC performance methods are reported in Table 2, in which the cells display the frequency of the correct model order selection. To obtain these results, the simulation was carried out as follows. The generated data was a PAR(1) process, and PAR models were estimated with orders 1 to 4, for series with and without outliers. The criterions AIC and BIC were calculated using the PeACV estimators by classical (Section 3.1) and robust (Section 5.1) methods. In the case of the robust model identification procedure, two cases, denoted as Robust-case 1 and Robust-case 2, were considered. In the former, the robust autocovariance functions were only used to compute the estimates of the coefficients $\phi_j^{(m)}$, whereas to obtain the results of Robust-case 2, the robust autocovariance estimator was used to obtain the

Table 2
Frequencies of correct model selection using AIC and BIC criterions. The data generating process is a PAR(1). $\omega = 0$ means the series has no outliers.

Model	ω	Classic		Robust-case 1		Robust-case 2	
		AIC	BIC	AIC	BIC	AIC	BIC
1	$\omega = 0$	0.8763	0.9343	0.9881	0.9834	0.9126	0.9163
	$\omega = 4$	0.8712	0.9376	0.9363	0.8650	0.8780	0.8677
	$\omega = 7$	0.8257	0.6644	0.7311	0.4326	0.7978	0.7931
2	$\omega = 0$	0.8756	0.9781	0.9991	1.0000	0.9757	0.9854
	$\omega = 4$	0.6766	0.9307	0.9815	0.9961	0.9721	0.9876
	$\omega = 7$	0.2903	0.9102	0.8188	0.9810	0.9549	0.9918
3	$\omega = 0$	0.8837	0.9978	0.9984	1.0000	0.9747	0.9978
	$\omega = 4$	0.4175	0.9382	0.8715	0.9865	0.9863	0.9979
	$\omega = 7$	0.0657	0.8401	0.2988	0.9167	0.9865	0.9989

estimates of the coefficients $\phi_j^{(m)}$ and also the residual variances σ_m^2, that is, its estimate is given by

$$\hat{\sigma}^2_{(m),Q} = \hat{\gamma}_Q^{(m)}(0),$$

where, in Eq. (13), $u^{(m)} = \{\hat{\epsilon}_{t(i,m)-h}\}_{r^* \le i \le N}$ and $v^{(m)} = \{\hat{\epsilon}_{t(i,m)}\}_{r^* \le i \le N}$. In the previous equalities, the $\hat{\epsilon}_{t(i,m)}$'s correspond to the residuals.

For the purpose of the empirical investigation considered in this paper, the results discussed in this subsection give the first insight of the effect of outliers in the step procedures of modeling PAR models in finite sample data sets. Firstly, the case is discussed where the data is outlier free ($\omega = 0$). From Table 2, it can be seen, when using the classical PeACV, that the AIC criterion presents the smallest percentage of correct model selection. This is in accord with the results observed in the empirical investigation of Franses and Paap [14]. These authors explored the model selection criterion empirical properties in PAR models with an extensive simulation exercise. Their results also gave evidence that the AIC criterion, in general, is inclined to opt for a too high model order.

However, interesting empirical evidence of the AIC method is observed when the robust PeACV is used to estimate the parameters (Robust-case 1 and Robust-case 2). This method presents a significant improvement in the model identification order. The percentage of correct order increases significantly in all models, and the method displays similar results to the BIC criterion. The latter criterion performs best in all situations by presenting selection frequencies which have very high and stable rate values.

The effect of outliers in order identification is clearly observed from the above results. Both criterions are destroyed with the increase of the outlier magnitudes. The poor performance of the method is not surprising, and it is one of the consequences caused by outliers in the classical PeACV discussed in the previous sections. As was expected, a significant reduction of the correct order identification frequencies is observed in Model 3, where the root is close to the unit root condition, that is, λ defined in (4) is close to one. Notwithstanding this, the BIC method is less affected than the AIC method.

The robust performance of the methods is shown by the use of the robust methodology suggested in this paper (Robust-case 1 and Robust-case 2). The outliers do not significantly affect the order selection of the robust criterion methods, the BIC criterion being, in general, the method with the best performance. This method presents, in general, values which are reasonably stable over the data generating processes. In the case of Model 1 with outliers, it is somewhat surprising that the robust methodologies are outperformed by the classical one. In order to verify if this phenomenon was possibly caused by the small sample size used, simulations of Model 1 for $N = 700$ ($n = 2800$) were carried out. For this sample size, the methods performed similarly to in the other models, that is, the robust ones outperform the classical methodology (the results are available upon request). For example, in the case of $\omega = 7$, the BIC method presented the rates 0.9392 and 0.9754 for the classical and Robust-case 1 methodologies, respectively. From Table 2 (Robust-case 2), it is also interesting to see the performance of the methods when the robust methodology is used to estimate the parameters and the residual variances. A significant improvement of the behavior of the methods is observed when the data is contaminated by outliers with very large magnitude ($\omega = 7$).

As a general conclusion, the findings support the claim that, even for outliers with large magnitude, the robust criterions AIC and BIC showed themselves to be robust against additive outliers.

6.2. Parameter estimates

This section reveals the most important empirical investigation of this paper. The asymptotic theories of the robust autocovariance and the estimation procedures discussed in the previous sections are investigated in different scenarios for finite sample size from the data generating PAR(1) models. First of all, the asymptotic properties given in Propositions 5 and 7 are empirically investigated and the results are shown graphically in Figs. 1 and 2. As a second part of this simulation section, the model parameters are estimated with the estimation procedures based on the Yule–Walker equations, using the classical and the robust sample autovariance functions. The sample mean, the bias and the Mean Squared Error (MSE) of the estimates of Models 1 to 3 are reported in Tables 3–5, respectively.

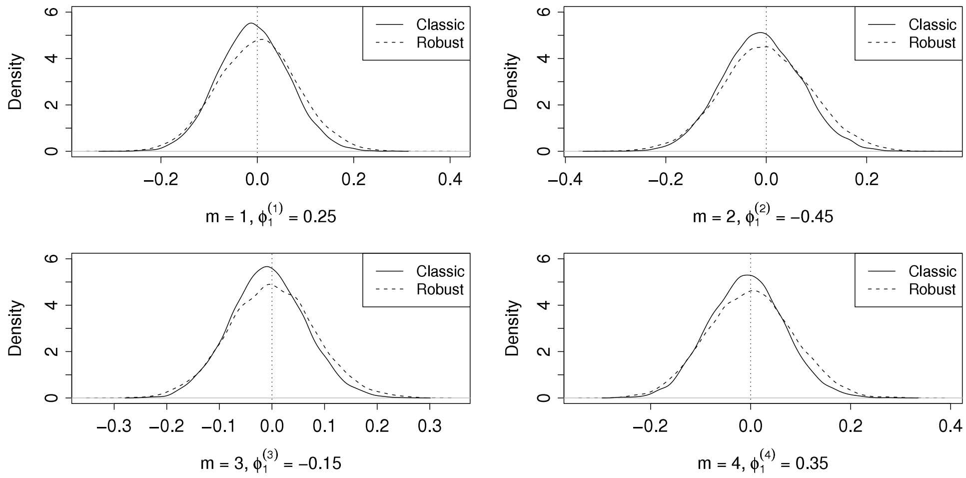

Fig. 1. Empirical densities of $\sqrt{N}(Q_N^{(m)} - Q^{(m)})$ and $\sqrt{n}(sd - Q^{(m)})$ when $\omega = 0$.

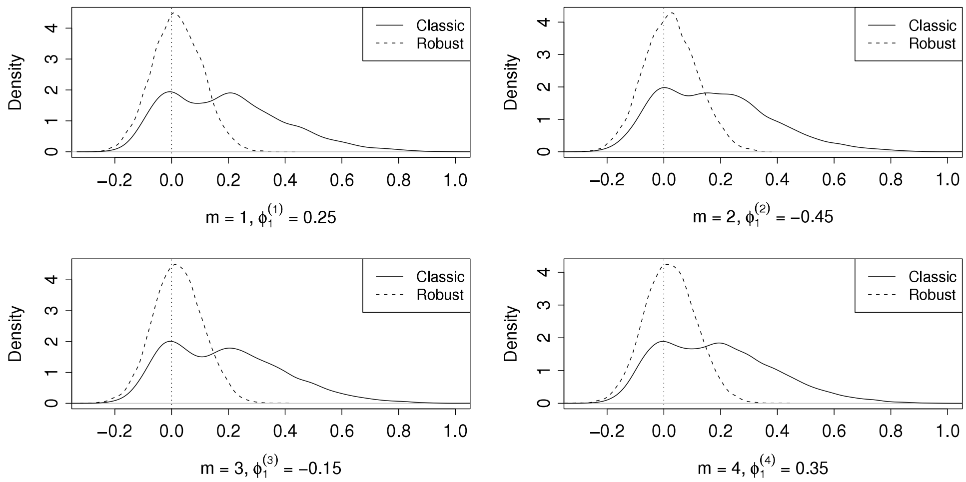

Fig. 2. Empirical densities of $\sqrt{N}(Q_N^{(m)} - Q^{(m)})$ and $\sqrt{n}(sd - Q^{(m)})$ when $\omega = 7$.

6.2.1. Empirical investigations of the asymptotic properties of the Q_N and the robust autocovariance

As a simulation example to verify the finite sample property of Proposition 5, 10,000 samples from Model 1, with and without outliers, were simulated to obtain the values of $Q_N^{(m)}$, $m = 1, \ldots, 4$, and the standard deviation *sd* of each series $Z_1, \ldots, Z_n$. The standard deviations of the PAR(1) model are $Q^{(1)} = 1.070$ ($\phi_1^{(1)} = 0.25$), $Q^{(2)} = 1.217$ ($\phi_1^{(2)} = -0.45$), $Q^{(3)} = 1.027$ ($\phi_1^{(3)} = -0.15$) and $Q^{(4)} = 1.125$ ($\phi_1^{(4)} = 0.35$). From these, empirical densities of the quantities $\sqrt{N}(Q_N^{(m)} - Q^{(m)})$ and $\sqrt{N}(sd - Q^{(m)})$ were calculated and the plots are in Figs. 1 and 2.

These figures give some insights relating to one of the main contributions of this paper, which is stated in Proposition 5. In the case of without outliers, the empirical densities of the quantities $\sqrt{N}(Q_N^{(m)} - Q^{(m)})$ and $\sqrt{N}(sd - Q^{(m)})$ have very similar behaviors; both present shapes close to the $N(0, 1)$ distribution. This is consistent with the result in Proposition 5. However, the presence of outliers radically alters the picture of $\sqrt{N}(sd - Q^{(m)})$, whereas the density of $\sqrt{N}(Q_N^{(m)} - Q^{(m)})$ remains unaffected by atypical observations; see Fig. 2.

6.2.2. Model parameter estimation results

Finally, to end this simulation section the performance of the estimation methods for the parameters of Models 1 to 3 are now presented and discussed. The estimates are reported in Tables 3–5 and some of these results are illustrated graphically in the box-plots given in Fig. 3.

Initially, the case where the series is not contaminated ($\omega = 0$) is analyzed. It can be seen that when the value of λ is close to zero (Models 1 and 2) all methods are very competitive, which is not a surprising result given the performance of Q_n and *sd* previously reported. However, when the value of λ is near to the unit root (Model 3), the classical estimator continues to provide estimates closer to the real values, while the robust one presents estimates with a slight increase in the *MSE*.

Table 3
Mean, bias and $\sqrt{MSE}$ of the parameter estimates of Model 1.

ω	Parameters	Classic			Robust		
		Mean	$\|bias\|$	$\sqrt{MSE}$	Mean	$\|bias\|$	$\sqrt{MSE}$
$\omega = 0$	$\phi_1^{(1)} = 0.25$	0.2473	0.0027	0.0954	0.2517	0.0017	0.1120
	$\phi_1^{(2)} = -0.45$	−0.4482	0.0018	0.0984	−0.4492	0.0008	0.1155
	$\phi_1^{(3)} = -0.15$	−0.1496	0.0004	0.0919	−0.1495	0.0005	0.1048
	$\phi_1^{(4)} = 0.35$	0.3478	0.0022	0.1002	0.3490	0.0010	0.1174
$\omega = 4$	$\phi_1^{(1)} = 0.25$	0.2200	0.0300	0.1047	0.2528	0.0028	0.1142
	$\phi_1^{(2)} = -0.45$	−0.3993	0.0507	0.1221	−0.4550	0.0050	0.1208
	$\phi_1^{(3)} = -0.15$	−0.1329	0.0171	0.0970	−0.1497	0.0003	0.1087
	$\phi_1^{(4)} = 0.35$	0.3074	0.0426	0.1168	0.3522	0.0022	0.1214
$\omega = 7$	$\phi_1^{(1)} = 0.25$	0.1856	0.0644	0.1304	0.2596	0.0096	0.1178
	$\phi_1^{(2)} = -0.45$	−0.3361	0.1139	0.1813	−0.4675	0.0175	0.1244
	$\phi_1^{(3)} = -0.15$	−0.1149	0.0351	0.1082	−0.1543	0.0043	0.1110
	$\phi_1^{(4)} = 0.35$	0.2609	0.0891	0.1564	0.3650	0.0150	0.1242

Table 4
Mean, bias and $\sqrt{MSE}$ of the parameter estimates of Model 2.

ω	Parameters	Classic			Robust		
		Mean	$\|bias\|$	$\sqrt{MSE}$	Mean	$\|bias\|$	$\sqrt{MSE}$
$\omega = 0$	$\phi_1^{(1)} = 0.9$	0.8877	0.0123	0.0772	0.9036	0.0036	0.1046
	$\phi_1^{(2)} = 0.8$	0.7961	0.0039	0.0651	0.7978	0.0022	0.0859
	$\phi_1^{(3)} = 0.7$	0.6969	0.0031	0.0628	0.6978	0.0022	0.0811
	$\phi_4^{(1)} = 0.6$	0.5962	0.0038	0.0685	0.5978	0.0022	0.0852
$\omega = 4$	$\phi_1^{(1)} = 0.9$	0.8187	0.0813	0.1311	0.8933	0.0067	0.1107
	$\phi_1^{(2)} = 0.8$	0.7513	0.0488	0.0967	0.7908	0.0092	0.0923
	$\phi_1^{(3)} = 0.7$	0.6581	0.0419	0.0875	0.6911	0.0089	0.0860
	$\phi_1^{(4)} = 0.6$	0.5598	0.0402	0.0892	0.5930	0.0070	0.0888
$\omega = 7$	$\phi_1^{(1)} = 0.9$	0.7266	0.1734	0.2401	0.9229	0.0229	0.1164
	$\phi_1^{(2)} = 0.8$	0.6785	0.1215	0.1770	0.8122	0.0122	0.0955
	$\phi_1^{(3)} = 0.7$	0.5982	0.1018	0.1539	0.7124	0.0124	0.0906
	$\phi_1^{(4)} = 0.6$	0.5026	0.0974	0.1496	0.6110	0.0110	0.0906

Table 5
Mean, bias and $\sqrt{MSE}$ of the parameter estimates of Model 3.

w	Parameters	Classic			Robust		
		Mean	$\|bias\|$	$\sqrt{MSE}$	Mean	$\|bias\|$	$\sqrt{MSE}$
$\omega = 0$	$\phi_1^{(1)} = 1.5$	1.4754	0.0246	0.0629	1.4981	0.0019	0.0962
	$\phi_1^{(2)} = 0.8$	0.7940	0.0060	0.0348	0.7931	0.0069	0.0512
	$\phi_1^{(3)} = 1.2$	1.1940	0.0060	0.0399	1.1940	0.0060	0.0657
	$\phi_1^{(4)} = 0.5$	0.4914	0.0086	0.0327	0.4897	0.0103	0.0433
$\omega = 4$	$\phi_1^{(1)} = 1.5$	1.4126	0.0874	0.1281	1.4681	0.0319	0.1143
	$\phi_1^{(2)} = 0.8$	0.7797	0.0203	0.0462	0.7869	0.0131	0.0554
	$\phi_1^{(3)} = 1.2$	1.1664	0.0336	0.0643	1.1817	0.0183	0.0757
	$\phi_1^{(4)} = 0.5$	0.4838	0.0162	0.0394	0.4858	0.0142	0.0470
$\omega = 7$	$\phi_1^{(1)} = 1.5$	1.3116	0.1884	0.2537	1.4840	0.0160	0.1153
	$\phi_1^{(2)} = 0.8$	0.7542	0.0458	0.0761	0.7894	0.0106	0.0583
	$\phi_1^{(3)} = 1.2$	1.1165	0.0835	0.1269	1.1809	0.0191	0.0827
	$\phi_1^{(4)} = 0.5$	0.4700	0.0300	0.0553	0.4874	0.0126	0.0484

In general, all methods present pretty similar results. Significant changes in the behavior of the classical estimation method are observed when the process has outliers, especially for large ω. As previously mentioned, the outliers provoke a memory reduction of the process. From the tables it can be seen that the classical estimates underestimate the real parameters by presenting negative large bias, and this empirical evidence is in close agreement with the memory loss property previously mentioned. In general, the performance of the robust estimator shows that its estimates seem to be unaffected by the

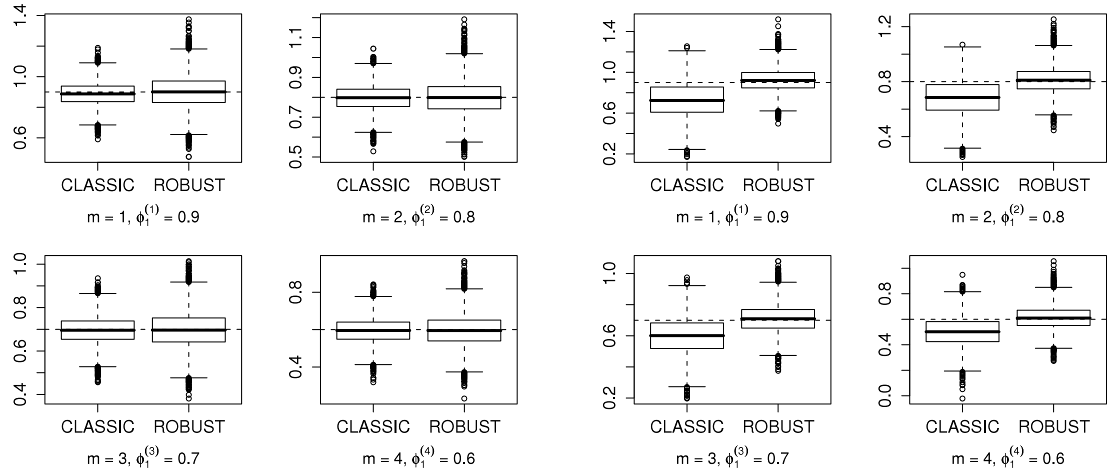

Fig. 3. Box-plots of the estimates from Model 2 when $\omega = 0$ (left) and when $\omega = 7$ (right).

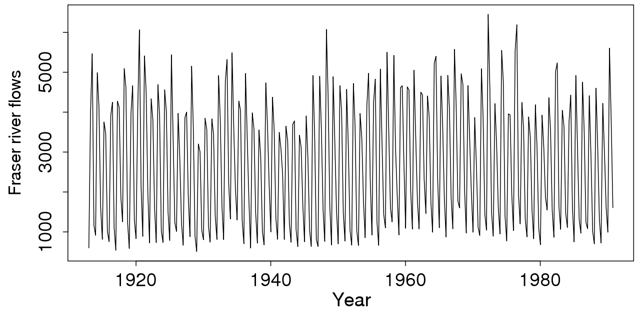

Fig. 4. Mean quarterly flows of Fraser River.

outliers. The visual variability of the estimates of Model 2 when $\omega = 0$ and $\omega = 7$ are presented in Figs. 3 and 4, respectively. These illustrate graphically the performance of the methods previously discussed.

The empirical investigation here presented suggests that, in general, when there is no evidence of atypical observations, Yule–Walker estimates based on the classical autocovariance function give satisfactory results, which confirms the asymptotic results given in Proposition 7. Nevertheless, the robust procedure presented very competitive estimates. The empirical study also provides evidence that special attention has to be paid when the data possibly exhibits atypical observations. The robust methodology showed that, in general, the estimates are essentially constant across different parameter values and outlier magnitudes.

7. An application

The main goal of this section is to apply the robust methodology, discussed in the previous section, to the quarterly mean flows (cm) of Fraser River. The series ranges from the first quarter in 1913 to the fourth quarter of 1991, and is available in the pear package (Periodic Autoregression Analysis) of the R software. The flows of Fraser River have also been studied in many works. For example, McLeod [25] analyzed the monthly mean of the series and adjusted a PAR model with different orders according to the periods. The quarterly mean flows (cm) are displayed in Fig. 4. The seasonality component is clearly evidenced.

The spikes and the troughs in the series (Fig. 4) and the box-plot pictures (Fig. 5(a)) indicate that the dynamic of the series varies across the four periods. This strongly suggests that the series has a periodic correlation structure. This phenomenon is also observed in the scatter plots of (Fig. 5(b)), from where it can be seen that there is evidence of correlations in period 1 versus period 4 and also in period 3 versus period 4.

Also, the plot of the series (Fig. 4) and the box-plots of periods 1 and 3 (Fig. 5(a)) show few observations that can possibly be identified as atypical observations, or outliers. From this arises the question whether or not the series has outliers, and if there any, how much is the impact of these observations on the parameter estimates. Because of the suspicion of outliers in the series, the data was centered by subtracting the seasonal medians, instead of the seasonal means, in the modeling step. As a part of the identification model, the sample classical and robust autocorrelations were calculated; the plots are in Fig. 6.

The autocorrelations present a positive decay behavior with a weak correlation dependence. These plots give a visual indication of a PAR model with order 1 for all periods. This was confirmed by the calculus of the BIC criterion. Among all

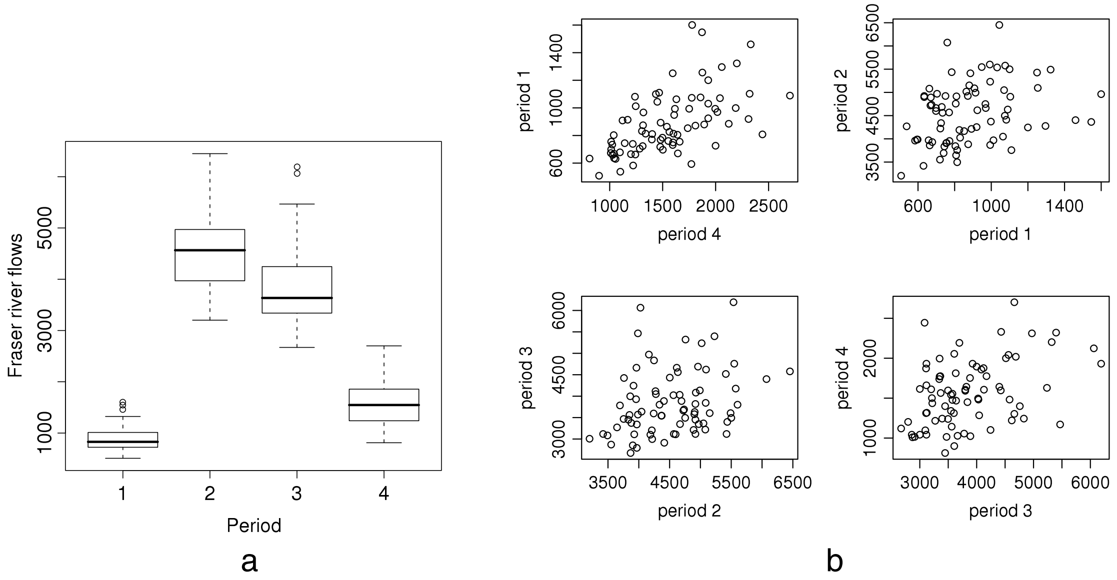

Fig. 5. Figures of the box-plots (a) and the scatter plots of the periods (b).

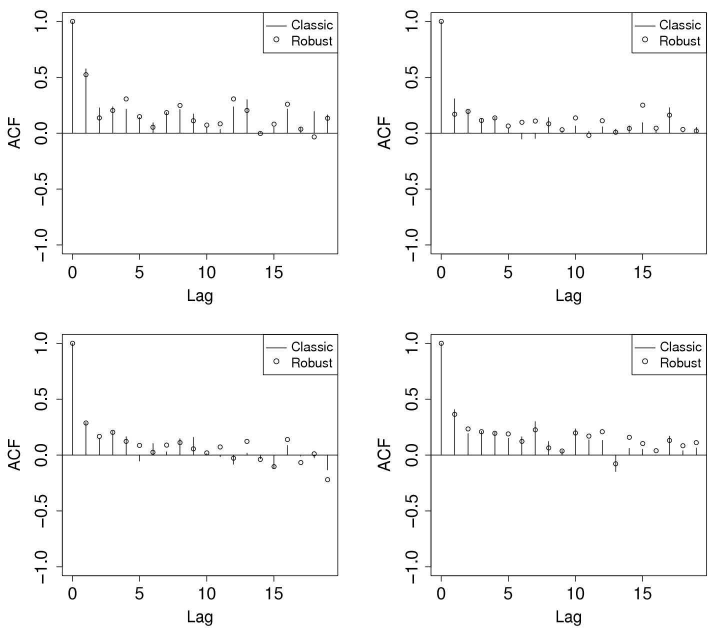

Fig. 6. Sample classical and robust autocorrelations from period 1 (top left) to period 4 (bottom right).

Table 6
Yule–Walker parameter estimates using the classical and the robust methodologies.

Parameters	Estimates	
	Classic	Robust
$\phi_1^{(1)}$	0.3249	0.3442
$\phi_1^{(2)}$	0.8243	1.0215
$\phi_1^{(3)}$	0.3265	0.3121
$\phi_1^{(4)}$	0.2078	0.2576

models considered, the criterion yielded a PAR(1, 1, 1, 1) model to fit the data. The robust BIC method also indicated the same model order. The estimates are in Table 6.

Table 7
Yule–Walker parameter estimates using the classical and the robust methodologies.

Parameters	Estimates	
	Classic	Robust
$\phi_1^{(1)}$	0.3214	0.3442
$\phi_1^{(2)}$	0.2944	1.1035
$\phi_1^{(3)}$	−0.0020	0.3147
$\phi_1^{(4)}$	0.2201	0.2576

Both models gave close estimates, except for the period 2. This is not a surprising result from the correlation plots (Fig. 6). It can be seen that for all periods, both correlations behave similarly and their curves cross each other. However, at lag 1 of period 2, there is a small discrepancy between the correlations. This may explain the difference between the estimates in the case of period 2. The standardized residuals of both models did not diagnose model inadequacy (these are available upon-request). Based on this analysis, there is no doubt that the outliers, if any, are not strong enough to have an effect on the estimates.

Since the previous investigation indicated that if the series has outliers they are not strong enough to destroy the classical parameter estimates, outliers with magnitude 10 were randomly distributed with $p = 0.01$ in the data. The contaminated series is not plotted here to save space, but available upon request. The PAR(1, 1, 1, 1) model was again estimated and the results are in Table 7.

From this table, it can be seen that robust estimates remained unchanged whereas the estimate from the classical methodology for period 3 was totally destroyed. The two outliers were randomly located in period 2 of years 1916 and 1964. These outlier positions explained why the parameter of period 3 was the only one significantly affected.

8. Conclusions

This paper investigates the impact of additive outliers on estimating and identifying the PAR model. The paper proposes a robust estimation procedure which is based on the robust scale and covariance functions given in, respectively, Rousseeuw and Croux [28] and Ma and Genton [23]. The asymptotic properties of the estimator are established and its finite sample performance is investigated through Monte Carlo experiments. The empirical investigation is carried out in different scenarios, such as order identification and estimation in series with and without outliers. The quarterly Fraser River data was analyzed as an application of the methodology studied here. The behavior of the sample autocorrelations did not exhibit the presence of outliers in the data. Based on all these investigations, the robust methodology proved to be an alternative estimation procedure in the context of modeling PAR processes with additive outliers.

9. Proofs

Proof of Proposition 1. Let $\boldsymbol{\phi}^\star$ denote the least-squares estimator of $\boldsymbol{\phi}$ as defined in [31] for a zero-mean PAR(p) process. According to the theorem on page 289 in [31],

$$\sqrt{N}(\boldsymbol{\phi}^\star - \boldsymbol{\phi}) \xrightarrow{d} \mathcal{N}(0, F^{-1}),$$

where F is defined in Proposition 1. Let us now prove that, for all $m \in \{1, \dots, s\}$,

$$N^{1/2}(\hat{\phi}^{(m)} - \phi^{\star(m)}) = o_P(1).$$

Using the same line of reasoning as in [5], page 263, we obtain

$$N^{1/2}(\hat{\phi}^{(m)} - \phi^{\star(m)}) = N^{1/2}(\hat{\Gamma}_p^{(m)})^{-1}\left(\hat{\gamma}_p^{(m)} - N^{-1}X^{(m)\prime}Y^{(m)}\right) + N^{1/2}\left((\hat{\Gamma}_p^{(m)})^{-1} - N(X^{(m)\prime}X^{(m)})^{-1}\right)N^{-1}X^{(m)\prime}Y^{(m)},$$

where $X^{(m)} = (Y_{t(i,m)-j})_{1\le i\le N, 1\le j\le p}$, $Y^{(m)} = (Y_{t(1,m)}, \dots, Y_{t(N,m)})'$, $\hat{\gamma}_p^{(m)} = (\hat{\gamma}_Y^{(m)}(k))_{1\le k\le p}$ and $\hat{\Gamma}_p^{(m)} = (\hat{\gamma}_Y^{(m-j)}(i-j))_{1\le i,j\le p}$. The ith component of $N^{1/2}(\hat{\gamma}_p^{(m)} - N^{-1}X^{(m)\prime}Y^{(m)})$ is equal to

$$-N^{-1/2}\sum_{k=1-i}^{0} Y_{t(k,m)}Y_{t(k,m)+i} + N^{1/2}\bar{Y}^{(m)}\left((1 - N^{-1}i)\bar{Y}^{(m)} - N^{-1}\sum_{k=1}^{N-i}(Y_{t(k,m)} + Y_{t(k,m)+i})\right),$$

where $\bar{Y}^{(m)} = N^{-1}\sum_{r=1}^{N} Y_{t(r,m)}$. Since $(Y_{t(r,m)})$ is a causal PAR(p) process, Theorem 7.1.2 of Brockwell and Davis [5] gives $N^{1/2}\bar{Y}^{(m)} = O_P(1)$. From this, we deduce that $(N^{1/2}(\hat{\gamma}_p^{(m)} - N^{-1}X^{(m)\prime}Y^{(m)}))_i = o_P(1)$, for all $i \in \{1, \dots, p\}$. Using the same arguments as those given in [5] on page 264,

$$N^{1/2}\|(\hat{\Gamma}_p^{(m)})^{-1} - N(X^{(m)\prime}X^{(m)})^{-1}\| = o_P(1).$$

Given that $N^{-1}X^{(m)\prime}Y^{(m)} \xrightarrow{P} \gamma_p^{(m)}$ and $\hat{\Gamma}_p^{(m)} \xrightarrow{P} \Gamma_p^{(m)}$, as N tends to infinity, the proof of the proposition is concluded. □

Proof of Proposition 4. Let m be in $\{1,\ldots,s\}$, then, for all non negative h,

$$\widehat{\gamma}_z^{(m)}(h)-\widehat{\gamma}_y^{(m)}(h)-(\gamma_Z^{(m)}(h)-\gamma_Y^{(m)}(h))=(\widehat{\gamma}_z^{(m)}(h)-\gamma_Z^{(m)}(h))-(\widehat{\gamma}_y^{(m)}(h)-\gamma_Y^{(m)}(h)).$$

Using Theorem 7.2.1 of Brockwell and Davis [5], we get that $\widehat{\gamma}_y^{(m)}(h)-\gamma_Y^{(m)}(h)=o_P(1)$, for all m in $\{1,\ldots,s\}$ and all non negative h. Since $\{\sum_{i=1}^{l}\omega_i V_i^{t(r,m)}\}_{r\geq 1,m=1,\ldots,s}$ are i.i.d. random variables, the process $\{Z_{t(r,m)}\}$ is also causal and satisfies the previous theorem. Thus, $\widehat{\gamma}_z^{(m)}(h)-\gamma_Z^{(m)}(h)=o_P(1)$, which concludes the proof. □

Proof of Proposition 5. Since $(Y_{t(r,m)})_{r\geq 1,m=1,\ldots,s}$ is a periodic stationary process, $(Y_{t(i,m)})_{i\geq 1}$ is also a stationary process with autocovariance function at lag k equal to $\gamma_Y^{(m)}(ks)$, for a given m in $\{1,\ldots,s\}$.

Using (A1), Theorem 7.2, (P. 96) of Rio [27] thus implies that $\sqrt{N}(F_N^{(m)}-F^{(m)})$ converges in distribution to a Gaussian process in the space of cadlag functions equipped with the topology of uniform convergence. The notation $F_N^{(m)}$ corresponds to the empirical distribution function based on $(Y_{t(i,m)})_{1\leq i\leq N}$.

Under (A1), we also have that: $(T_1(F^{(m)}))'(T_0(F^{(m)}))=2\int f^{(m)}(y)f^{(m)}(y+T_0(F^{(m)}))\mathrm{d}y$. Thus, Lemma 1 established in [19] can be applied. The asymptotic expansion of $a_N(Q_N^{(m)}-Q^{(m)})$ obtained in (19) is thus valid with $a_N=\sqrt{N}$. Then, a CLT for $N^{-1/2}\sum_{i=1}^{N}\mathrm{IF}(Y_{t(i,m)},Q^{(m)},F^{(m)})$ has to be proved. Lemma 12 established in [19] can easily be extended to hold even for Gaussian random variables non necessarily standard. Applying this lemma, we obtain that $\tau=2$ in Theorem 4 (P. 2256) of Arcones [2]. Thus, by Assumption (A1), Condition (2.40) of this theorem is satisfied with $\tau=2$. This concludes the proof of the proposition. □

Proof of Proposition 6. Let us denote by $F_{+,N-r^\star+1}^{(m)}$ and $F_{-,N-r^\star+1}^{(m)}$ the empirical c.d.f of $(Y_{t(i,m)-h}+Y_{t(i,m)})_{r^\star\leq i\leq N}$ and $(Y_{t(i,m)-h}-Y_{t(i,m)})_{r^\star\leq i\leq N}$ respectively. Since $(Y_{t(i,m)-h})_{i\geq r^\star}$ and for $(Y_{t(i,m)})_{i\geq r^\star}$ Theorem 7.2 (P. 96) of Rio [27] can be applied. Thus, we obtain that $(N-r^\star+1)^{1/2}(F_{+,N-r^\star+1}^{(m)}-F_{+}^{(m)})$ converges in distribution to a Gaussian process in the space of cadlag functions equipped with the topology of uniform convergence and that the same holds for $(N-r^\star+1)^{1/2}(F_{-,N-r^\star+1}^{(m)}-F_{-}^{(m)})$. As a consequence, the expansion (19) is valid for $Q_{N-r^\star+1}(\{Y_{t(i,m)-h}+Y_{t(i,m)}\}_{r^\star\leq i\leq N})$ and $Q_{N-r^\star+1}(\{Y_{t(i,m)-h}-Y_{t(i,m)}\}_{r^\star\leq i\leq N})$ with $a_{N-r^\star+1}=(N-r^\star+1)^{1/2}$ and IF defined in (12). Then, using the Delta method (Theorem 3.1 (P. 26) in [34]), $\hat{\gamma}_Q(h)$ satisfies the following asymptotic expansion:

$$(N-r^\star+1)^{1/2}\left(\hat{\gamma}_Q^{(m)}(h)-\left\{Q^2(F_+^{(m)})-Q^2(F_-^{(m)})\right\}/4\right)=(N-r^\star+1)^{-1/2}\sum_{i=r^\star}^{N}\psi(Y_{t(i,m)-h},Y_{t(i,m)})+o_P(1),\qquad(16)$$

where ψ is defined in (14). Hence, we have to prove a CLT for

$$(N-r^\star+1)^{-1/2}\sum_{i=r^\star}^{N}\psi(Y_{t(i,m)-h},Y_{t(i,m)}).$$

Using Lemma 13, established in [19], using the definition of the Hermite rank given on P. 2245 in [2] and Assumption (A1), we obtain that Condition (2.40) of Theorem 4 (P. 2256) in [2] is satisfied with $\tau=2$. This concludes the proof of the proposition by observing that $\{Q^2(F_+^{(m)})-Q^2(F_-^{(m)})\}/4=\mathrm{Cov}(Y_{t(r^\star,m)-h},Y_{t(r^\star,m)})=\gamma^{(m)}(h)$. □

Proof of Proposition 7. Let m in $\{1,\ldots,s\}$ then

$$N^{1/2}(\tilde{\phi}^{(m)}-\phi^{(m)})=N^{1/2}(\tilde{\phi}^{(m)}-\hat{\phi}^{(m)})+N^{1/2}(\hat{\phi}^{(m)}-\phi^{(m)}),\qquad(17)$$

where $\hat{\phi}^{(m)}$ is the Yule–Walker estimator of $\phi^{(m)}=(\phi_1^{(m)},\ldots,\phi_{pm}^{(m)})'$ defined from the classical autocovariance estimator. By Proposition 1, the second term in the r.h.s of (17) is equal to $O_P(1)$. The first term in the r.h.s of (17) can be rewritten as follows

$$\begin{aligned}N^{1/2}(\tilde{\phi}^{(m)}-\hat{\phi}^{(m)})&=N^{1/2}\left((\tilde{\Gamma}_{pm}^{(m)})^{-1}\tilde{\gamma}_{pm}-(\hat{\Gamma}_{pm}^{(m)})^{-1}\hat{\gamma}_{pm}\right)\\&=N^{1/2}(\tilde{\Gamma}_{pm}^{(m)})^{-1}\left(\tilde{\gamma}_{pm}-\hat{\gamma}_{pm}\right)+N^{1/2}\left((\tilde{\Gamma}_{pm}^{(m)})^{-1}-(\hat{\Gamma}_{pm}^{(m)})^{-1}\right)\hat{\gamma}_{pm},\end{aligned}\qquad(18)$$

where $\hat{\gamma}_{pm}^{(m)}=(\hat{\gamma}_Y^{(m)}(k))_{1\leq k\leq pm}$, $\hat{\Gamma}_{pm}^{(m)}=(\hat{\gamma}_Y^{(m-j)}(i-j))_{1\leq i,j\leq pm}$, $\tilde{\gamma}_{pm}^{(m)}=(\hat{\gamma}_Q^{(m)}(k))_{1\leq k\leq pm}$ and $\tilde{\Gamma}_{pm}^{(m)}=(\hat{\gamma}_Q^{(m-j)}(i-j))_{1\leq i,j\leq pm}$. Using Theorem 7.2.1 of Brockwell and Davis [5] and Proposition 6, both terms in the r.h.s of (18) are equal to $O_P(1)$, which concludes the proof. □

Acknowledgments

We thank PIBIC-CNPq for their financial support.

Appendix

Lemma 1. *Let $(X_i)_{i\geq 1}$ be a stationary sequence of random variables having a common cumulative distribution function F. Let F_n be the empirical distribution function based on $(X_i)_{1\leq i\leq n}$. Assume that $a_n(F_n - F)$ converges in distribution in the space of cadlag functions equipped with the uniform norm for some positive sequence (a_n). Assume also that $T_1(F)$ is differentiable at $T_0(F)$ where $T_0(.)$ and $T_1(.)$ are defined in (6) and (8) respectively. Then, Q_n defined by (9) satisfies the following asymptotic expansion:*

$$a_n(Q_n - Q) = \frac{a_n}{n}\sum_{i=1}^{n} \mathrm{IF}(X_i, Q, F) + o_P(1), \tag{19}$$

where $Q = c\, T_0(F)$, the functional T_0 being defined in (8), c being defined in (10) and

$$\mathrm{IF}(x, Q, F) = 2c\left(\frac{1/4 - F(x + T_0(F)) + F(x - T_0(F))}{(T_1(F))'(T_0(F))}\right). \tag{20}$$

Lemma 2. *Let X be a standard Gaussian random variable. The influence function* IF *of Q_n defined by (20) satisfies the following properties:*

$$\mathbb{E}[\mathrm{IF}(X, Q, \Phi)] = 0, \tag{21}$$

$$\mathbb{E}[X\mathrm{IF}(X, Q, \Phi)] = 0, \tag{22}$$

$$\mathbb{E}[X^2\mathrm{IF}(X, Q, \Phi)] = (2\sqrt{\pi}\beta)^{-1}\exp(-1/(4c^2)) \neq 0, \tag{23}$$

where Φ is the c.d.f of a standard Gaussian random variable, β is defined in (10) and $\beta = \int \varphi(y)\varphi(y + 1/c)\mathrm{d}y$, φ being the p.d.f associated to Φ.

Lemma 3. *Let (X, Y) be a standard Gaussian random vector such that $\mathrm{Cov}(X, Y) = 0$. The influence function ψ defined by (14) satisfies the following properties:*

$$\mathbb{E}[\psi(X, Y)] = 0, \tag{24}$$

$$\mathbb{E}[X\psi(X, Y)] = \mathbb{E}[Y\psi(X, Y)] = 0, \tag{25}$$

$$\mathbb{E}[XY\psi(X, Y)] \neq 0. \tag{26}$$

References

[1] P.L. Anderson, M.M. Meerschaert, Parameter estimation for periodically stationary time series, Journal of Time Series Analysis 26 (14) (2005) 489–518.
[2] M. Arcones, Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors, Annals of Probability 22 (4) (1994) 2242–2274.
[3] P. Bloomfield, H.L. Hurd, R.B. Lund, Periodic correlation in stratospheric ozone data, Journal of Time Series Analysis 15 (1994) 127–150.
[4] G.E.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden Day, 1976.
[5] P.J. Brockwell, R.A. Davis, Time Series: Theory and Methods, Springer Verlag, 1991.
[6] W. Chan, A note on time series model specification in the presence of outliers, Journal of Applied Statistics 19 (1992) 117–124.
[7] W. Chan, Outliers and financial time series modelling: a cautionary note, Mathematics and Computers in Simulation 39 (1995) 425–430.
[8] I. Chang, G.C. Tiao, C. Chen, Estimation of time series parameters in the presence of outliers, Technometrics 30 (1988) 193–204.
[9] C. Chen, L. Liu, Joint estimation of model parameters and outlier effects in time series, Journal of the American Statistical Association 88 (1993) 284–297.
[10] L. Denby, R.D. Martin, Robust estimation of the first order autoregressive parameter, Journal of the American Statistical Association 74 (1979) 140–146.
[11] S.J. Deutsch, J.E. Richards, J. Swain, Effects of a single outlier on ARMA identification, Communications in Statistics: Theory and Methods 19 (1990) 2207–2227.
[12] F. Fajardo, V.A. Reisen, F. Cribari-Neto, Robust estimation in long-memory processes under additive outliers, Journal of Statistical Planning and Inference 139 (2009) 2511–2525.
[13] P.H. Franses, R. Paap, Forecasting with periodic autoregressive time series models, Econometric Institute Report 9927/A, 1999.
[14] P.H. Franses, R. Paap, Model selection in periodic autoregressions, Oxford Bulletin of Economics and Statistics 56 (1994) 421–439.
[15] W.A. Gardner, L.E. Franks, Intervention analysis with applications to economic and enviromental problems, IEEE Transactions on Information Theory 21 (1975) 4–14.
[16] H.L. Hurd, N.L. Gerr, Graphical methods for determining the presence of periodic correlation, Journal of Time Series Analysis 12 (4) (1991) 337–350.
[17] H.L. Hurd, A. Miamee, Periodically Correlated Random Sequences: Spectral Theory and Practice, in: Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2007, URL: http://dx.doi.org/10.1002/9780470182833.
[18] J. Ledolter, The effect of additive outliers on the forecast from ARMA models, International Journal of Forecasting 5 (1989) 231–240.
[19] C. Lévy-Leduc, H. Boistard, E. Moulines, M. Taqqu, V.A. Reisen, Robust estimation of the scale and of the autocovariance function of Gaussian short and long-range dependent processes, Journal of Time Series Analysis (2009) (in press) ArXiv:0912.4686.
[20] R.B. Lund, I.V. Basawa, Recursive prediction and likelihood evaluation for periodic ARMA models, Journal of Time Series Analysis 21 (2000) 75–93.
[21] R.B. Lund, H.L. Hurd, P. Bloomfield, R.L. Smith, Climatological time series with periodic correlation, Journal of Climate 8 (1995) 2787–2809.
[22] H. Lütkepohl, New Introduction to Multiple Time Series Analysis, Springer Verlag, 2005.
[23] Y. Ma, M. Genton, Highly robust estimation of the autocovariance function, Journal of Time Series Analysis 21 (2000) 663–684.
[24] A.I. McLeod, Parsimony, model adequacy and periodic correlation in time series forecasting, International Statistical Review 61 (1993) 387–393.
[25] A.I. McLeod, Diagnostic checking periodic autoregression models with application, Journal of Time Series Analysis 15 (1994) 221–233.
[26] M.B. Priestley, Spectral Analysis and Time Series, Academic Press, 1981.
[27] E. Rio, Théorie asymptotique des processus aléatoires faiblement dépendants, Springer, 2000.

[28] P.J. Rousseeuw, C. Croux, Alternatives to the median absolute deviation, Journal of the American Statistical Association 88 (1993) 1273–1283.
[29] G. Schwarz, Estimating the dimension of a model, The Annals of Statistics 6 (1978) 461–464.
[30] Q. Shao, Robust estimation for periodic autoregressive time series, Journal of Time Series Analysis 29 (2) (2008) 251–263.
[31] Q. Shao, P. Ni, Least-squares estimation and ANOVA for periodic autoregressive time series, Statistics and Probability Letters 69 (2004) 287–297.
[32] G.C. Tiao, M.R. Grupe, Hidden periodic autoregressive-moving average models in time series data, Biometrika 67 (1980) 365–373.
[33] E. Ursu, P. Duchesne, On modelling and diagnostic checking of vector periodic autoregressive time series models, Journal of Time Series Analysis 30 (1) (2008) 70–96.
[34] A.W. van der Vaart, Asymptotic statistics, Cambridge University Press, 2000.
[35] A.V. Vecchia, Periodic autoregressive-moving average (PARMA) modeling with applications to water resources, Journal of the American Water Resources Association 21 (1985a) 721–730.
[36] A.V. Vecchia, Maximum likelihood estimation for periodic autoregressive moving average models, Technometrics 27 (1985b) 375–384.
[37] A.V. Vecchia, R. Ballerini, Testing for periodic autocorrelations in seasonal time series data, Biometrika 78 (1991) 53–63.