# Empirical study of robust estimation methods for PAR models with application to the air quality area

Carlo Correa Solci, Valdério A. Reisen, Alessandro Jose Queiroz Sarnaglia, Pascal Bondon

# Empirical Study of Robust Estimation Methods for PAR Models with Application to the Air Quality Area

Carlo Correa Solci* Valdério Anselmo Reisen*‡†

Alessandro Jose Queiroz Sarnaglia‡ Pascal Bondon§

**Abstract**

This paper compares three estimators for periodic autoregressive (PAR) models. The first is the classical periodic Yule-Walker estimator (YWE) [McLeod, 1994]. The second is a robust version of YWE (RYWE) which uses the robust autocovariance function [Ma and Genton, 2000] in the periodic Yule-Walker equations [Sarnaglia et al., 2010], and the third is the robust least squares estimator (RLSE) based on iterative least squares with robust versions of the original time series [Shao, 2008]. The daily mean particulate matter concentration ($PM_{10}$) data is used to illustrate the methodologies in a real application, that is, in the Air Quality area.

Keywords: Robust estimation; PAR models; Outliers; $PM_{10}$ pollutant.

## 1 Introduction

Classical time series analysis generally relies on stationarity assumptions, see, e.g., Fuller [1976], Priestley [1981], Brockwell and Davis [1991], Shumway and Stoffer [2017]. Despite the broad use of stationary tools, in some cases, this requirement is too restrictive. Examples of non-stationary phenomena are unit roots, deterministic trends, heteroskedasticity, among others.

The periodic correlation (PC) or cyclostationarity property, introduced by the seminal paper of Gladyshev [1961], deserves special attention due to the fact that this phenomenon is not revealed by usual stationary tools, which may lead to a model misspecification [Tiao and Grupe, 1980]. Due to this fact, special methods to identify the presence of PC in time series have been proposed, see, for example, Hurd and Gerr [1991] and Bloomfield et al. [1994]. PC appears in many areas of application: Gardner and Franks [1975] investigate cyclostationarity in electrical engineering; Lund et al. [1995] have found PC in climatological time series; Noakes et al. [1985] have discovered PC in time series of monthly river flows, among others.

One of the most popular models for PC is the PAR model, which is a generalization of the well-known AR model introduced by Box and Jenkins [1970] where the coefficients and orders vary periodically in time. Estimation methods for the PAR model parameters have been studied by many authors among, Basawa and Lund [2001], Anderson and Meerschaert [2005], Sarnaglia et al. [2015].

Although the PAR model has been applied in several fields, to the best of our knowledge, it is still relatively unexplored in the air quality research field, especially in the context of contaminated data. Among the air pollutants, the particulate matter with diameter smaller than 10 µm ($PM_{10}$), is recognized for its effects on human health and is one of the most common and important pollution variables collected by an air quality monitoring network, see, for example, Reisen et al. [2014], Souza et al. [2018] and references therein. Note that, the air quality data usually present

*PPGEA/UFES,Vitória-ES-Brazil

†CentraleSupélec

‡DEST/UFES-Vitória-ES-Brazil

§Laboratoire des Signaux et Systèmes (L2S), CNRS-CentraleSupélec-Université Paris-Sud, 3 rue Joliot Curie, 91192, Gif-sur-Yvette, France

asymmetric distributions and large peeks of concentrations. Classical estimators such the sample mean, variance and autocovariance functions are affected by these observations. This suggests to consider robust estimators for PAR parameters, like the ones proposed by Shao [2008], Sarnaglia et al. [2010, 2016]. One issue of this paper is to fit robust PAR models to $PM_{10}$ concentrations.

To the best our knowledge, there are no empirical studies in the literature of cyclostationary processes investigating the behaviour of robust estimators under asymmetric errors and observations which can be identified as atypical or outliers. This paper aims to investigate finite sample properties of such estimators under asymmetric errors and atypical observations through a Monte Carlo study. In addition, a pollutant $PM_{10}$ concentration data set is used as an example of application since it may present observations with high levels of pollutant concentrations that may produce sample distributions with heavy tails.

The rest of the paper is organized as follows: Section 2 introduces the well-known PAR model and describes three estimation methods; Section 3 presents and discusses the results of the Monte Carlo experiment; Section 4 illustrates the use of the estimation methodologies with an application to fit a regression model with PAR errors to $PM_{10}$ concentrations; Section 5 concludes the paper.

# 2 The PAR Model and its Estimation Methods

Let $\{Y_t\}$, $t \in \mathbb{Z}$, be a stochastic process with $\mathsf{E}(Y_t^2) < \infty$, $\mu_t = \mathsf{E}(Y_t)$ and autocovariance $\gamma_t(h) = \mathsf{Cov}(Y_t, Y_{t-h})$. Process $\{Y_t\}$ has PC or is a periodically stationary process with period $\mathcal{S} \in \mathbb{N}^+$ ($\text{PS}_\mathcal{S}$), if

$$\mu_{t+\mathcal{S}} = \mu_t \quad \text{and} \quad \gamma_{t+\mathcal{S}}(h) = \gamma_t(h), \quad t, h \in \mathbb{Z}, \tag{1}$$

$\mathcal{S}$ being the smallest integer satisfying (1). Now, let $t = r\mathcal{S} + \nu$, where $r \in \mathbb{Z}$ and $\nu = 1, \ldots, \mathcal{S}$. It follows from (1) that $\mu_{r\mathcal{S}+\nu} = \mu_\nu$ and $\gamma_{r\mathcal{S}+\nu}(h) = \gamma_\nu(h)$. Therefore, the autocorrelation $\rho_t(h) = \mathsf{Corr}(Y_t, Y_{t-h})$ satisfies $\rho_{r\mathcal{S}+\nu}(h) = \rho_\nu(h)$. In addition, the partial autocorrelation (PACF) defined as

$$\alpha_t(h) = \mathsf{Corr}(Y_t, Y_{t-h} | Y_{t-1}, \ldots, Y_{t-h+1}) \quad t \in \mathbb{Z}, h \in \mathbb{N}^+,$$

see, e.g., Brockwell and Davis [1991], is also periodic in time, i.e., $\alpha_{r\mathcal{S}+\nu}(h) = \alpha_\nu(h)$. Clearly, the above functions only depend on the period $\nu$ and the lag $h$. When they do not depend on $\nu$, $\{Y_t\}$ is a standard stationary time series in the terminology of Box-Jenkins.For more details, see for example, McLeod [1994], Sarnaglia et al. [2010] and references therein.

The standard stationary linear model can be extended to the $\text{PS}_\mathcal{S}$ process $\{Y_{r\mathcal{S}+\nu}\}$ via

$$Y_{r\mathcal{S}+\nu} = \sum_{j \in \mathbb{Z}} \psi_j(\nu) \epsilon_{r\mathcal{S}+\nu-j},$$

where $\sum_{j \in \mathbb{Z}} |\psi_j(\nu)| < \infty$ for $\nu = 1, \ldots, \mathcal{S}$. The model is causal when $\psi_j(\nu) = 0$ for $j < 0$. In the same way, the model is invertible when

$$\sum_{j \geq 0} \pi_j(\nu) Y_{r\mathcal{S}+\nu-j} = \epsilon_{r\mathcal{S}+\nu},$$

where $\sum_{j \geq 0} |\pi_j(\nu)| < \infty$ for $\nu = 1, \ldots, \mathcal{S}$, see Sarnaglia et al. [2010] and references therein.

The PAR model is a generalization of the well-known AR process and is one of the most used models to fit a $\text{PS}_\mathcal{S}$ time series. The PAR model is given in the following definition.

**Definition 1.** A zero-mean $\text{PS}_\mathcal{S}$ process $\{Y_{r\mathcal{S}+\nu}\}$ follows a PAR($p_\nu$) model if it satisfies the difference equation

$$Y_{r\mathcal{S}+\nu} - \sum_{i=1}^{p_\nu} \phi_i(\nu) Y_{r\mathcal{S}+\nu-i} = \sigma_\nu \epsilon_{r\mathcal{S}+\nu}, \tag{2}$$

where $\{\epsilon_t\}$ is a sequence of uncorrelated random variables with $\mathsf{E}(\epsilon_t) = 0$, $\mathsf{E}(\epsilon_t^2) = 1$ and, for each cycle $\nu = 1, \ldots, \mathcal{S}$, $\boldsymbol{\phi}_\nu = (\phi_1(\nu), \ldots, \phi_{p_\nu}(\nu))'$ is the AR coefficient vector with order $p_\nu$ and $\sigma_\nu^2$ is the error variance.

Conditions to ensure causality of a PAR model can be derived from its vector AR representation, see, e.g., Sarnaglia et al. [2010]. In particular, for a PAR(1) model, the causality condition is

$$\left| \prod_{\nu=1}^{\mathcal{S}} \phi_1(\nu) \right| < 1. \tag{3}$$

It is assumed here that $p_1 = \cdots = p_{\mathcal{S}} = p$ and the following notation $\boldsymbol{\phi} = (\boldsymbol{\phi}_1', \ldots, \boldsymbol{\phi}_{\mathcal{S}}')' = (\phi_1(1), \ldots, \phi_p(1), \ldots, \phi_1(\mathcal{S}), \ldots, \phi_p(\mathcal{S}))'$.

### 2.1 The Yule-Walker Estimator (YWE)

Let $Y_1, \ldots, Y_{n\mathcal{S}}$ be a sample from the process $\{Y_t\}$. The estimates of the periodic mean $\mu_\nu$ and autocovariance function $\gamma_\nu(h)$ for $\nu = 1, \ldots, \mathcal{S}$ are, respectively, $\bar{Y}_\nu = \frac{1}{n}\sum_{r=0}^{n-1} Y_{r\mathcal{S}+\nu}$, and

$$\hat{\gamma}_\nu(h) = \frac{1}{n}\sum_{r=0}^{n-1}(Y_{r\mathcal{S}+\nu} - \bar{Y}_\nu)(Y_{r\mathcal{S}+\nu-h} - \bar{Y}_{\nu-h}),$$

where $Y_{r\mathcal{S}+\nu-h}$ is set to zero whenever $r\mathcal{S} + \nu - h < 1$ or $r\mathcal{S} + \nu - h > n\mathcal{S}$. Therefore, the sample ACF is

$$\hat{\rho}_\nu(h) = \frac{\hat{\gamma}_\nu(h)}{[\hat{\gamma}_\nu(0)\hat{\gamma}_{\nu-h}(0)]^{\frac{1}{2}}}.$$

The sample PACF can be obtained as in Sakai [1982] and Shao and Lund [2004].

The YWE of the PAR coefficients are obtained through the linear equations system

$$\sum_{i=1}^{p} \phi_i(\nu)\gamma_{\nu-i}(h-i) = \gamma_\nu(h), \quad h = 1, \ldots, p, \tag{4}$$

in which $\gamma_\nu(h)$ is replaced by $\hat{\gamma}_\nu(h)$. The YWE of $\boldsymbol{\phi}$ is defined as $\hat{\boldsymbol{\phi}} = (\hat{\boldsymbol{\phi}}_1', \ldots, \hat{\boldsymbol{\phi}}_{\mathcal{S}}')'$ where, for each $\nu$, $\hat{\boldsymbol{\phi}}_\nu = (\hat{\phi}_1(\nu), \ldots, \hat{\phi}_p(\nu))'$. Asymptotics results for YWE can be derived under the following assumption.

**Assumption 1.** $\{Y_t\}$ is a zero-mean causal PAR process and $\mathsf{E}(Y_t^4) < \infty$, $t \in \mathbb{Z}$.

The following result is due to Sarnaglia et al. [2010]: Under Assumption 1, the YWE estimator $\hat{\boldsymbol{\phi}}$ satisfies

$$\sqrt{n}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}) \rightsquigarrow \mathcal{N}_{p\mathcal{S}}(0, G), \quad n \to \infty,$$

where the $\rightsquigarrow$ symbol denotes convergence in distribution.

Other asymptotically equivalent estimators of PAR coefficients are the Least Squares Estimator (LSE) [Basawa and Lund, 2001] and the Maximum Likelihood Estimator (MLE) [Lund and Basawa, 1999]. Due to this equivalence, these estimators will not be considered in this paper.

As well known, the estimators YWE, LSE and MLE, are not resistant in the presence of atypical observations (outliers) see, for example, Shao [2008], Sarnaglia et al. [2010]. The lack of robustness of classical estimators has motivated the investigation of robust approaches in the literature. In the following subsections the robust methods introduced by Sarnaglia et al. [2010] and Shao [2008] are summarized and the empirical comparison with YWE estimator is presented in the Simulation Section.

## 2.2 The Robust Yule-Walker Estimator (RYWE)

Sarnaglia et al. [2010] have proposed a Robust Yule-Walker Estimator (RYWE) which is based on the autocovariance function proposed in Ma and Genton [2000]. Let $\mathsf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, the robust scale estimator of $\mathsf{y}$ proposed by Rousseeuw and Croux [1993], $Q_n(\mathsf{y})$, is defined as the following $k$th order statistic

$$Q_n(\mathsf{y}) = d\left\{|y_i - y_j|\,; 1 \leq i < j \leq n\right\}_{(k)}, \tag{5}$$

where $d$ is a constant factor to ensure Fisher-consistency and $k = \binom{c}{2} \approx 0.25\binom{n}{2}$, where $c = [n/2]+1$ is half of the size $n$ of the vector $\mathsf{y}$. For Gaussian random variables, $d = 2.2191$. Given a PS$_{\mathcal{S}}$ time series, $Y_1, \ldots, Y_{n\mathcal{S}}$, based on (5) Sarnaglia et al. [2010] define the robust sample (periodic) ACV function by

$$\tilde{\gamma}_\nu(h) = \frac{1}{4}[Q^2_{n-r+1}(\mathsf{u}_\nu + \mathsf{v}_\nu) - Q^2_{n-r+1}(\mathsf{u}_\nu - \mathsf{v}_\nu)], \quad 0 \leq h < [(n-1)\mathcal{S} + \nu], \tag{6}$$

where $\mathsf{u}_\nu = (Y_{r\mathcal{S}+\nu-h}, \ldots, Y_{(n-1)\mathcal{S}+\nu-h})$, $\mathsf{v}_\nu = (Y_{r\mathcal{S}+\nu}, \ldots, Y_{(n-1)\mathcal{S}+\nu})$. Note that, $\hat{\gamma}_\nu(h)$ does not posses the positive definite property. The RYWE is defined similarly to the YWE. For each $\nu = 1, \ldots, \mathcal{S}$, replace, in (4), the theoretical ACV, $\gamma_\nu(h)$, by its sample robust estimator in Equation 6, $\tilde{\gamma}_\nu(h)$, and solve the resulting linear equations system

$$\sum_{i=1}^{p} \phi_i(\nu)\tilde{\gamma}_{\nu-i}(h-i) = \tilde{\gamma}_\nu(h), \quad k = 1, \ldots, p, \tag{7}$$

for $\phi_1(\nu), \ldots, \phi_p(\nu)$. The RYWE estimator is defined as $\tilde{\boldsymbol{\phi}} = (\tilde{\boldsymbol{\phi}}_1', \ldots, \tilde{\boldsymbol{\phi}}_{\mathcal{S}}')'$. The white noise variance $\sigma^2_\nu$ can also be robustly estimated by the same argument as in Equation 4 with $h = 0$.

**Assumption 2.** For any $\nu = 1, \ldots, \mathcal{S}$, $\{Y_{r\mathcal{S}+\nu};\ r \in \mathbb{Z}\}$ is a mean-zero Gaussian process with strong mixing coefficients $\alpha_n$ satisfying: $\alpha_n \leq Cn^{-a}$, for some $a > 1$ and $C \geq 1$.

The following result is due to Sarnaglia et al. [2010]. Under Assumptions 1 and 2, the RYWE estimator $\tilde{\phi}_i(\nu)$ satisfies

$$\tilde{\phi}_i(\nu) - \phi_i(\nu) = O_p(n^{-1/2}),$$

for $i = 1, \ldots, p$, $\nu = 1, \ldots, \mathcal{S}$.

## 2.3 The Robust Least Squares Estimator (RLSE)

Shao [2008] proposes an alternative to the conditional Least Squares Estimator (LSE). The LSE of $\boldsymbol{\phi}$ can be defined as the solution of the $p\mathcal{S}$-dimensional estimating equations

$$S_n(\boldsymbol{\phi}_\nu) = \frac{1}{\sigma_\nu}\sum_{r=0}^{n-1} \epsilon_{r\mathcal{S}+\nu} Y_{r\mathcal{S}+\nu-i} = 0, \quad 1 \leq i \leq p, \quad 1 \leq \nu \leq \mathcal{S}, \tag{8}$$

where the error terms are given by $\epsilon_{r\mathcal{S}+\nu} = \epsilon_{r\mathcal{S}+\nu}(\boldsymbol{\phi}_\nu) = (Y_{r\mathcal{S}+\nu} - \sum_{i=1}^{p}\phi_i(\nu)Y_{r\mathcal{S}+\nu-i})/\sigma_\nu$, $0 \leq r < n$, $1 \leq \nu \leq \mathcal{S}$. Asymptotic properties for the LSE have been studied by Basawa and Lund [2001].

Shao [2008] aims to achieve robustness by replacing, in Equation (8), $\epsilon_{r\mathcal{S}+\nu}$ and $Y_{r\mathcal{S}+\nu}$ by their robust versions $\breve{\epsilon}_{r\mathcal{S}+\nu}$ and $\breve{Y}_{r\mathcal{S}+\nu}$, respectively, which are defined as

$$\breve{\epsilon}_{r\mathcal{S}+\nu} = \psi(\epsilon_{r\mathcal{S}+\nu}) \tag{9}$$

and

$$\breve{Y}_{r\mathcal{S}+\nu} = \begin{cases} Y_{r\mathcal{S}+\nu}, & \text{if } \breve{\epsilon}_{r\mathcal{S}+\nu} = \epsilon_{r\mathcal{S}+\nu}, \\ \sum_{i=1}^{p}\phi_i(\nu)\breve{Y}_{r\mathcal{S}+\nu-i} + \sigma_\nu\breve{\epsilon}_{r\mathcal{S}+\nu}, & \text{if } \breve{\epsilon}_{r\mathcal{S}+\nu} \neq \epsilon_{r\mathcal{S}+\nu}. \end{cases} \tag{10}$$

Therefore, the Robust LSE (RLSE) is the solution of the robustified estimating equations

$$\breve{S}_n(\boldsymbol{\phi}_\nu) = \frac{1}{\sigma_\nu}\sum_{r=0}^{n-1}\psi\left(\frac{Y_{r\mathcal{S}+\nu} - \sum_{i=1}^{p}\phi_i(\nu)Y_{r\mathcal{S}+\nu-i}}{\sigma_\nu}\right)\breve{Y}_{r\mathcal{S}+\nu-i} = 0, \quad 1 \le i \le p, \quad 1 \le \nu \le \mathcal{S}. \quad (11)$$

Shao [2008] has considered $\psi(\cdot)$ as the Huber type function, defined by

$$\psi(x) = \psi_c(x) = \begin{cases} x, & \text{if } |x| \le c, \\ c\,\mathrm{sign}(x), & \text{if } |x| > c, \end{cases} \quad (12)$$

because it is monotonic, which ensures existence and uniqueness of solution to Equation 11. Nevertheless, any odd bounded and differentiable function can be a candidate for the $\psi(\cdot)$ function, including the so-called redescending functions, such as Bisquare, Hampel, Generalized Gauss-Weight, among others. However, they have many zeroes, which may lead to non-optimal solutions.

**Assumption 3.** The marginal density function $f_\epsilon(\cdot)$ of the error $\epsilon_{r\mathcal{S}+\nu}$ in Equation 2 is symmetric about the origin.

The following result has been derived by Shao [2008]: Under Assumptions 1 and 3, the RLSE $\breve{\boldsymbol{\phi}}$ satisfies

$$\sqrt{n}(\breve{\boldsymbol{\phi}} - \boldsymbol{\phi}) \rightsquigarrow \mathcal{N}_{p\mathcal{S}}(0, A), \quad n \to \infty,$$

where the covariance matrix $A$ is given in Equation 14 of Shao [2008].

In practice, the estimates are obtained using the following iterative procedure starting with an appropriate initial guess value for the RLSE. Suppose $\breve{\boldsymbol{\phi}}^{(l)}$ represents the vector of estimates at the $l$th iteration. Then, at the $(l+1)$th iteration, calculate the residuals

$$e_{r\mathcal{S}+\nu}^{(l)} = Y_{r\mathcal{S}+\nu} - \sum_{i=1}^{p}\breve{\phi}_i^{(l)}(\nu)Y_{r\mathcal{S}+\nu-i}, \quad 1 \le \nu \le \mathcal{S},$$

where $Y_t = 0$, $t \le 0$, estimate the white noise standard deviation at the period $\nu$, $\sigma_\nu$, by

$$\breve{\sigma}_\nu^{(l)} = \mathrm{Median}\left(\left|e_\nu^{(l)}\right|, \left|e_{\mathcal{S}+\nu}^{(l)}\right|, \ldots, \left|e_{(n-1)\mathcal{S}+\nu}^{(l)}\right|\right), \quad 1 \le \nu \le \mathcal{S},$$

calculate the robust version of $e_{r\mathcal{S}+\nu}^{(l)}$ through (9), $\breve{e}_{r\mathcal{S}+\nu}^{(l)} = \psi(e_{r\mathcal{S}+\nu}^{(l)})$, obtain $\breve{Y}_{r\mathcal{S}+\nu}^{(l)}$ by (10) with $\breve{\epsilon}_{r\mathcal{S}+\nu}$ substituted for the robust residual $\breve{e}_{r\mathcal{S}+\nu}^{(l)}$ and $\sigma_\nu$ replaced with $\breve{\sigma}_\nu^{(l)}$, and evaluate the solution $\breve{\boldsymbol{\phi}}^{(l+1)}$ of the robustified estimating equations in (11) replacing $\breve{Y}_{r\mathcal{S}+\nu}$ with $\breve{Y}_{r\mathcal{S}+\nu}^{(l)}$ and $\sigma_\nu$ with $\breve{\sigma}_\nu^{(l)}$. Stop the procedure according to some convergence criterion.

## 3 Monte Carlo Study

In order to investigate the impact of atypical observations on the estimates obtained from the methods discussed previously, series of periodically stationary processes were generated with and without additive outliers. Let $\{X_{r\mathcal{S}+\nu}\}$ be defined as follows

$$X_{r\mathcal{S}+\nu} = Y_{r\mathcal{S}+\nu} + \omega V_{r\mathcal{S}+\nu} \quad (13)$$

where $\{Y_{r\mathcal{S}+\nu}\}$ is a PAR model with $\mathcal{S} = 4$ and coefficients given in Table 1. The parameter values were chosen to have examples of time series models with low (Model 1) and strong (Model 2) correlation dependencies, that is, Model 2 is closer to the non-causality region than Model 1. $\{V_t\}$ is a sequence of independent random variables with $\mathsf{P}(V_t = -1) = \mathsf{P}(V_t = 1) = \xi/2$ and $\mathsf{P}(V_t = 0) = 1-\xi$, $0 \le \xi < 1$, $Y_t$ and $V_s$ are independent processes for all $t, s$ and $\omega$ is the magnitude of the outlier. The sample sizes were taken as small ($N = 400$) and large ($N = 1600$), .i.e., $n = 100$

and $n = 400$ cycles, respectively, which are common sample sizes in practical situation. The initial value for RLSE was taken as the true parameter vector.

The simulation study is divided in these two cases; uncontaminated and contaminated series. The contaminated series were generated from model in (13) with the following specifications: $\omega = 7$ and $\xi = 0.01$. The effect of the normality departure in the white noise sequence was also studied by generating the random variables $\epsilon_t$ such that $\epsilon_t \sim \mathcal{N}(0, 1)$ and $\sqrt{2}\epsilon_t + 1 \sim \chi^2_{(1)}$. In both cases, $\mathsf{E}(\epsilon_t) = 0$ and $\mathsf{E}(\epsilon_t^2) = 1$. For each of these scenarios, 1000 replicates of $\{Y_t\}$ were generated to compute the mean of the empirical Bias and Root Mean Squared Error (RMSE). For the RLSE, according to Shao [2008], $c = 3.06$ was fixed in the Huber function (Equation 12) such that, residuals greater than 3.06 (in absolute value) are regarded as outliers. Other model configurations were also considered in the simulation study such as heavy-tailed distributions, different period lengths, sample sizes and coefficient values and other outlier magnitudes. However, in general, the results leaded to similar conclusions and are not displayed here, but they are available upon request.

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

| Parameter | Model 1 | Model 2 |
|---|---|---|
| $\phi_1(1)$ | 0.9 | 1.5 |
| $\phi_1(2)$ | 0.8 | 0.8 |
| $\phi_1(3)$ | 0.7 | 1.2 |
| $\phi_1(4)$ | 0.6 | 0.5 |
| $\lambda$ | 0.3024 | 0.7200 |

Tables 2 and 3 display the Bias and RMSE for Models 1 and 2, respectively. For illustration purpose, the empirical distributions of $\sqrt{n}(\hat{\phi}_i(\nu) - \phi_i(\nu))$, $\sqrt{n}(\tilde{\phi}_i(\nu) - \phi_i(\nu))$ and $\sqrt{n}(\breve{\phi}_i(\nu) - \phi_i(\nu))$ for Model 1, with $n = 400$, under the uncontaminated and contaminated Gaussian scenarios, are presented in Figures 1 and 2, respectively. The same plots for the uncontaminated case with asymmetric errors are displayed in Figure 3. For Model 1 (Table 2), under the uncontaminated Gaussian case, the reduction of the Bias and RMSE when increasing the sample size suggests that all estimators are consistent. It is observed that, in this scenario, YWE and RLSE present better results. The findings for the asymmetric uncontaminated case show that the RYWE is extremely affected by skewness of the data, presenting a persistent Bias which does not seem to vanish by increasing the sample size. The results for heavy tailed errors are similar, which show that the RYWE method is very sensitive to departures from normality. This indicates that the normality requirement in Assumption 2 is crucial to ensure asymptotic properties of the RYWE. The YWE and RLSE do not seem to be affected by non-Gaussian errors (both asymmetric and heavy tailed). This gives empirical evidence that the technical requirement of symmetric errors (Assumption 3) may be over restrictive to ensure asymptotic normality of the RLSE. As expected, atypical observations increase the Bias and RMSE of the YWE. Under normally distributed errors, both RYWE and RLSE show robustness with Bias and RMSE almost unchanged with the presence of outliers. For asymmetric errors, the RLSE is the only one which presents good performance.

The conclusions for Model 2 (Table 3) are similar to the previous case. It is worth noting that in this stronger dependence scenario, there is a overall reduction of the Bias and RMSE. Another remarkable fact is that, in this case, the RYWE does not seem to be strongly affected by asymmetric errors in both uncontaminated and contaminated scenarios.

From Figure 1, it can be seen that the empirical distributions are virtually the same and they have shape very close to the N(0,1) distribution. This corroborates the asymptotic results of the standardized estimators even for a small sample size. However, the scale of the RYWE distribution is slightly greater than of those of the YWE and RLSE. Figure 2 shows the robustness to outliers of RYWE and RLSE methods under Gaussian errors, while the distribution of YWE is shifted to the left due to the well-known memory loss property. Its scale is also increased as a result of the contamination. Figure 3 illustrates the prominent shift to the right of the RYWE distribution

Table 2: Bias and RMSE for Model 1 and outliers with probability $\xi = 0.01$.

| $\omega$ | $\epsilon_t$ | $n$ | $\phi_1(\nu)$ | YWE Bias | YWE RMSE | RYWE Bias | RYWE RMSE | RLSE Bias | RLSE RMSE |
|---|---|---|---|---|---|---|---|---|---|
| 0 | $\mathcal{N}(0,1)$ | 100 | 0.9 | -0.007 | 0.077 | -0.003 | 0.103 | 0.003 | 0.079 |
| | | | 0.8 | -0.002 | 0.065 | 0.004 | 0.084 | -0.002 | 0.065 |
| | | | 0.7 | 0.000 | 0.063 | -0.001 | 0.083 | -0.001 | 0.063 |
| | | | 0.6 | -0.005 | 0.066 | -0.003 | 0.083 | -0.005 | 0.067 |
| | | 400 | 0.9 | -0.001 | 0.037 | -0.001 | 0.047 | 0.002 | 0.038 |
| | | | 0.8 | -0.001 | 0.031 | 0.000 | 0.038 | -0.001 | 0.031 |
| | | | 0.7 | -0.001 | 0.032 | 0.001 | 0.038 | -0.001 | 0.032 |
| | | | 0.6 | 0.000 | 0.032 | 0.000 | 0.039 | 0.000 | 0.033 |
| | $\frac{\chi^2_{(1)}-1}{\sqrt{2}}$ | 100 | 0.9 | -0.006 | 0.076 | 0.178 | 0.209 | 0.006 | 0.063 |
| | | | 0.8 | -0.007 | 0.065 | 0.117 | 0.147 | -0.003 | 0.055 |
| | | | 0.7 | -0.004 | 0.065 | 0.088 | 0.119 | -0.003 | 0.052 |
| | | | 0.6 | -0.005 | 0.069 | 0.077 | 0.108 | -0.004 | 0.056 |
| | | 400 | 0.9 | -0.001 | 0.037 | 0.179 | 0.185 | 0.002 | 0.030 |
| | | | 0.8 | 0.000 | 0.033 | 0.115 | 0.122 | 0.000 | 0.026 |
| | | | 0.7 | -0.001 | 0.033 | 0.089 | 0.096 | -0.001 | 0.026 |
| | | | 0.6 | -0.001 | 0.034 | 0.085 | 0.093 | 0.000 | 0.028 |
| 7 | $\mathcal{N}(0,1)$ | 100 | 0.9 | -0.181 | 0.247 | 0.014 | 0.120 | -0.031 | 0.095 |
| | | | 0.8 | -0.118 | 0.176 | 0.012 | 0.096 | -0.027 | 0.076 |
| | | | 0.7 | -0.105 | 0.157 | 0.015 | 0.091 | -0.024 | 0.077 |
| | | | 0.6 | -0.097 | 0.151 | 0.012 | 0.091 | -0.027 | 0.081 |
| | | 400 | 0.9 | -0.183 | 0.203 | 0.017 | 0.055 | -0.027 | 0.050 |
| | | | 0.8 | -0.129 | 0.144 | 0.012 | 0.046 | -0.021 | 0.041 |
| | | | 0.7 | -0.108 | 0.124 | 0.013 | 0.044 | -0.019 | 0.041 |
| | | | 0.6 | -0.103 | 0.119 | 0.014 | 0.043 | -0.022 | 0.043 |
| | $\frac{\chi^2_{(1)}-1}{\sqrt{2}}$ | 100 | 0.9 | -0.172 | 0.243 | 0.213 | 0.251 | -0.019 | 0.076 |
| | | | 0.8 | -0.126 | 0.180 | 0.142 | 0.175 | -0.021 | 0.063 |
| | | | 0.7 | -0.105 | 0.158 | 0.112 | 0.144 | -0.018 | 0.061 |
| | | | 0.6 | -0.096 | 0.151 | 0.103 | 0.134 | -0.023 | 0.067 |
| | | 400 | 0.9 | -0.182 | 0.202 | 0.211 | 0.219 | -0.021 | 0.041 |
| | | | 0.8 | -0.129 | 0.145 | 0.142 | 0.149 | -0.017 | 0.033 |
| | | | 0.7 | -0.112 | 0.127 | 0.111 | 0.119 | -0.017 | 0.033 |
| | | | 0.6 | -0.106 | 0.121 | 0.106 | 0.114 | -0.019 | 0.037 |

caused by the skewness of the errors.

# 4 An Application to the Air Quality Area (the $PM_{10}$ Data)

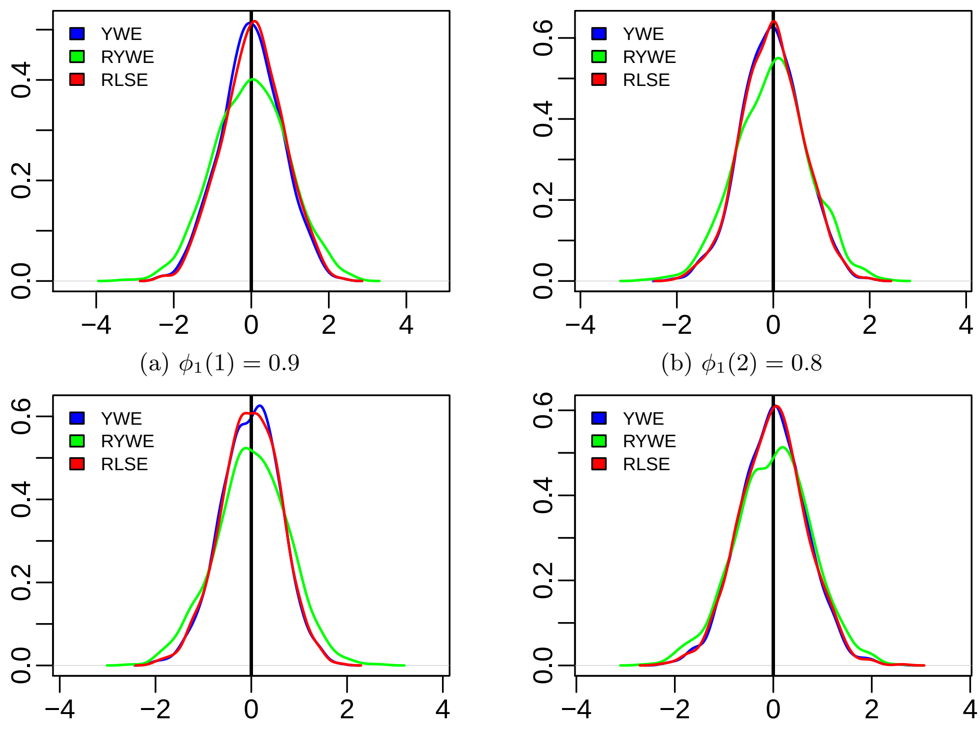
The application is based on a data set (air pollutant variables) collected at Automatic Air Quality Monitoring Network (RAMQAr) in the Great Vitória Region GVR-ES, Brazil, which is composed by nine monitoring stations placed in strategic locations and accounts for the measuring of several atmospheric pollutants and meteorological variables in the area. GVR is comprised of seven cities with a population of approximately 1.9 million inhabitants in an area of $2319\,\text{km}^2$. The region is situated along the South Atlantic coast of Brazil (latitude 20°19′15″S, longitude 40°20′10″W) and has a tropical humid climate, with average temperatures ranging from 24 °C to 30 °C. This data set has been previously investigated in different contexts of time series modeling, such as periodic models, robustness in long-memory models, heteroskedastic long memory process, time series regression with principal component analysis, among others. See, for instance, Sarnaglia et al. [2015], Reisen et al. [2014], Sarnaglia et al. [2010], Souza et al. [2018], Fajardo et al. [2018],

Table 3: Bias and RMSE for Model 2 and outliers with probability $\xi = 0.01$.

| $\omega$ | $\epsilon_t$ | $n$ | $\phi_1(\nu)$ | YWE Bias | YWE RMSE | RYWE Bias | RYWE RMSE | RLSE Bias | RLSE RMSE |
|---|---|---|---|---|---|---|---|---|---|
| 0 | $\mathcal{N}(0,1)$ | 100 | 1.5 | -0.009 | 0.055 | -0.009 | 0.100 | 0.000 | 0.055 |
| | | | 0.8 | -0.004 | 0.033 | -0.005 | 0.050 | -0.004 | 0.034 |
| | | | 1.2 | -0.006 | 0.040 | -0.008 | 0.066 | -0.006 | 0.040 |
| | | | 0.5 | -0.008 | 0.033 | -0.008 | 0.043 | -0.008 | 0.033 |
| | | 400 | 1.5 | -0.003 | 0.026 | -0.003 | 0.037 | -0.001 | 0.026 |
| | | | 0.8 | -0.001 | 0.016 | -0.001 | 0.021 | -0.001 | 0.016 |
| | | | 1.2 | -0.002 | 0.019 | -0.002 | 0.027 | -0.002 | 0.020 |
| | | | 0.5 | -0.002 | 0.015 | -0.003 | 0.018 | -0.003 | 0.015 |
| | $\frac{\chi^2_{(1)}-1}{\sqrt{2}}$ | 100 | 1.5 | -0.009 | 0.055 | 0.025 | 0.098 | 0.001 | 0.043 |
| | | | 0.8 | -0.007 | 0.036 | 0.011 | 0.048 | -0.005 | 0.028 |
| | | | 1.2 | -0.007 | 0.040 | 0.019 | 0.069 | -0.005 | 0.033 |
| | | | 0.5 | -0.009 | 0.033 | -0.004 | 0.039 | -0.006 | 0.027 |
| | | 400 | 1.5 | -0.002 | 0.026 | 0.032 | 0.050 | 0.000 | 0.021 |
| | | | 0.8 | -0.001 | 0.016 | 0.016 | 0.026 | -0.001 | 0.013 |
| | | | 1.2 | -0.002 | 0.019 | 0.023 | 0.035 | -0.002 | 0.015 |
| | | | 0.5 | -0.003 | 0.015 | 0.003 | 0.017 | -0.002 | 0.012 |
| 7 | $\mathcal{N}(0,1)$ | 100 | 1.5 | -0.174 | 0.246 | -0.023 | 0.122 | -0.014 | 0.061 |
| | | | 0.8 | -0.047 | 0.076 | -0.011 | 0.058 | -0.014 | 0.040 |
| | | | 1.2 | -0.079 | 0.125 | -0.013 | 0.079 | -0.013 | 0.048 |
| | | | 0.5 | -0.029 | 0.056 | -0.010 | 0.046 | -0.013 | 0.038 |
| | | 400 | 1.5 | -0.167 | 0.189 | -0.020 | 0.052 | -0.013 | 0.032 |
| | | | 0.8 | -0.042 | 0.050 | -0.007 | 0.025 | -0.007 | 0.019 |
| | | | 1.2 | -0.080 | 0.093 | -0.015 | 0.038 | -0.008 | 0.022 |
| | | | 0.5 | -0.023 | 0.032 | -0.006 | 0.021 | -0.007 | 0.018 |
| | $\frac{\chi^2_{(1)}-1}{\sqrt{2}}$ | 100 | 1.5 | -0.180 | 0.253 | 0.022 | 0.120 | -0.013 | 0.054 |
| | | | 0.8 | -0.049 | 0.081 | 0.011 | 0.057 | -0.008 | 0.032 |
| | | | 1.2 | -0.082 | 0.126 | 0.016 | 0.083 | -0.009 | 0.036 |
| | | | 0.5 | -0.029 | 0.053 | -0.003 | 0.041 | -0.010 | 0.028 |
| | | 400 | 1.5 | -0.172 | 0.193 | 0.022 | 0.054 | -0.010 | 0.025 |
| | | | 0.8 | -0.041 | 0.050 | 0.015 | 0.028 | -0.005 | 0.015 |
| | | | 1.2 | -0.083 | 0.097 | 0.012 | 0.039 | -0.007 | 0.018 |
| | | | 0.5 | -0.024 | 0.033 | 0.003 | 0.020 | -0.005 | 0.014 |

Reisen et al. [2018], Reisen et al. [2017] and references therein. The data set considered in this paper is the pollutant Particulate Matter with diameter smaller than 10 µm ($PM_{10}$), measured hourly, in µg/m$^3$, collected at the station located in Enseada do Suá area.

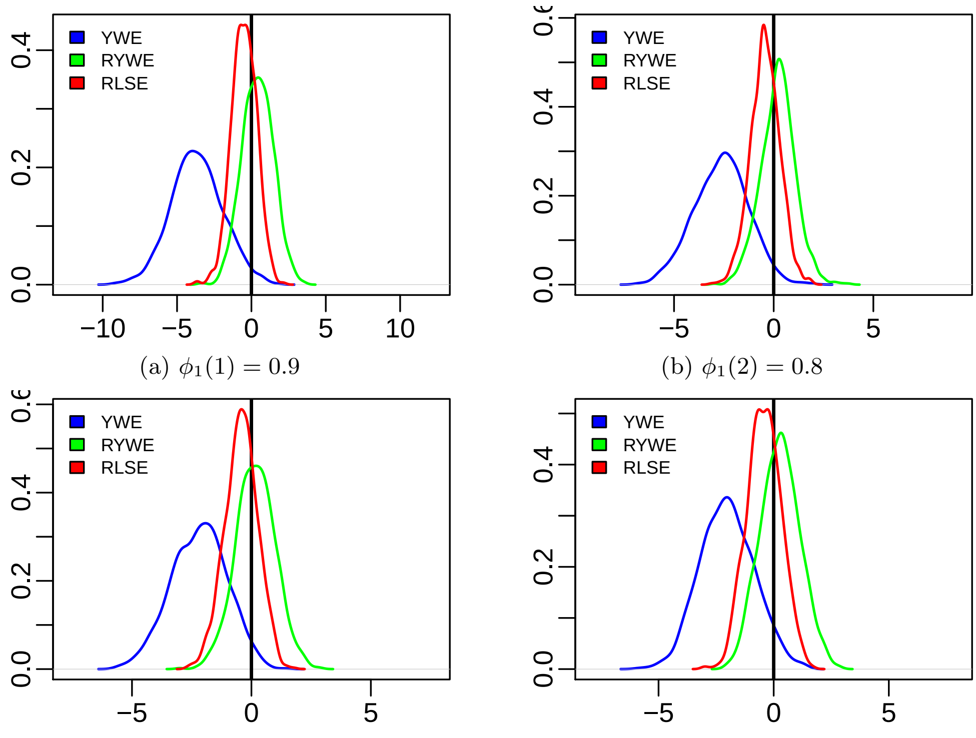
The $PM_{10}$ data set corresponds to daily average concentrations from January 1st, 2014 to December 29th, 2015 which kept the sample size multiple of the natural choice to the period length $\mathcal{S} = 7$. Due to skewness and some evidences of time varying variance, the natural logarithm transformation (log) was used and the plot of the log($PM_{10}$) is displayed in Figure 4. From this figure, one can see large peeks of $PM_{10}$ concentration which may be viewed here as outliers and, as mentioned previously, these high levels can provoke serious damage to some statistics, such as the mean and the standard deviation and, therefore, may affect the sample correlation structure of the series, causing misleading results. The existence of any outlier's effect will be discussed in the estimation parameter model (next subsection). It can also be seen the presence of sinusoidal deterministic trends. Analysis of the periodogram (Figure 5) corroborates to this result and indicates that the frequency $2/N$, corresponding approximately to a yearly cycle, has a large contribution to the overall variance of the data. The high frequency peaks of the periodogram




(a) $\phi_1(1) = 0.9$ (b) $\phi_1(2) = 0.8$

(c) $\phi_1(3) = 0.7$ (d) $\phi_1(4) = 0.6$

Figure 1: Empirical Distributions of $\sqrt{n}(\hat{\phi}_i(\nu) - \phi_i(\nu))$ (blue lines), $\sqrt{n}(\tilde{\phi}_i(\nu) - \phi_i(\nu))$ (green lines) and the $\sqrt{n}(\breve{\phi}_i(\nu) - \phi_i(\nu))$ (red lines) for Model 1 with $\omega = 0$, $n = 400$ and normal errors.

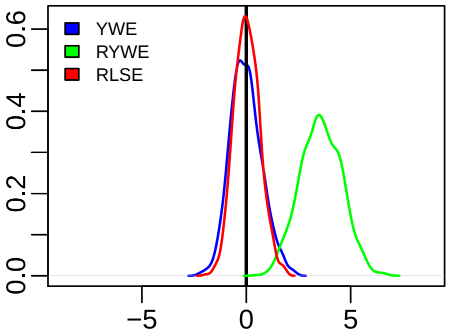



(a) $\phi_1(1) = 0.9$ (b) $\phi_1(2) = 0.8$

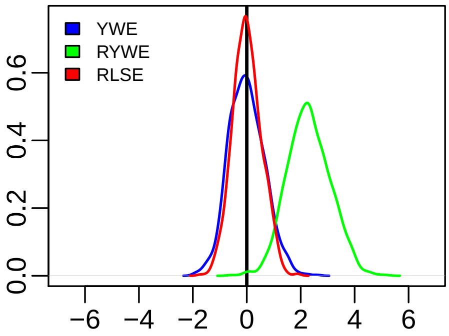
(c) $\phi_1(3) = 0.7$ (d) $\phi_1(4) = 0.6$

Figure 2: Empirical Distributions of $\sqrt{n}(\hat{\phi}_i(\nu) - \phi_i(\nu))$ (blue lines), $\sqrt{n}(\tilde{\phi}_i(\nu) - \phi_i(\nu))$ (green lines) and the $\sqrt{n}(\breve{\phi}_i(\nu) - \phi_i(\nu))$ (red lines) for Model 1 with $\omega = 7$, $n = 400$ and normal errors.

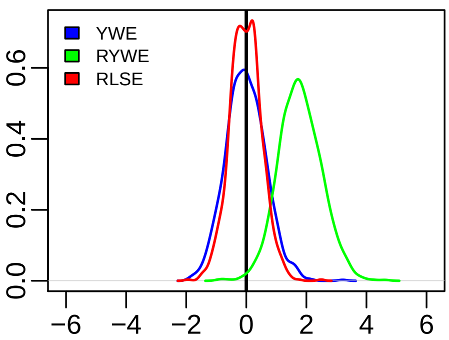
correspond to weekly periodicity and, according to the daily periodic boxplots displayed in Figure 6, they can be explained by a level decrease in the weekends. This is an expected finding due to fact that the traffic and civil construction decrease in the region in the weekend days.



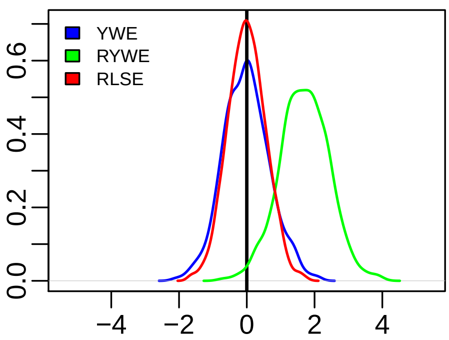

(a) $\phi_1(1) = 0.9$




(b) $\phi_1(2) = 0.8$




(c) $\phi_1(3) = 0.7$




(d) $\phi_1(4) = 0.6$

Figure 3: Empirical Distributions of $\sqrt{n}(\hat{\phi}_i(\nu) - \phi_i(\nu))$ (blue lines), $\sqrt{n}(\tilde{\phi}_i(\nu) - \phi_i(\nu))$ (green lines) and the $\sqrt{n}(\breve{\phi}_i(\nu) - \phi_i(\nu))$ (red lines) for Model 1 with $\omega = 0$, $n = 400$ and asymmetric errors.

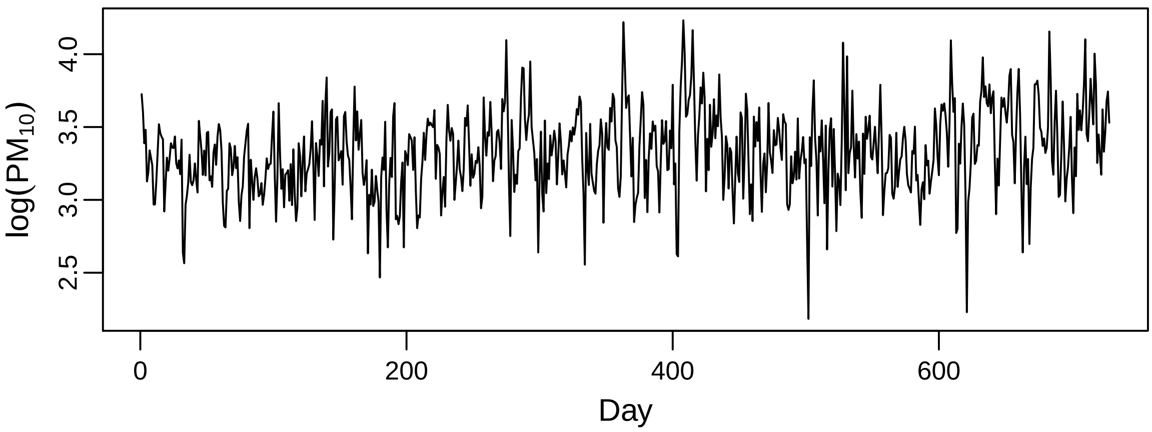



Figure 4: Plot of the $\log(\mathrm{PM}_{10})$ time series.

The above preliminary analysis of the series suggests that a deterministic trend must be firstly removed from $\log(\mathrm{PM}_{10})$ before further analysis and this is discussed in the next subsection in which a linear model with errors following a PAR model is fitted to the series.

### 4.1 Estimated Model

According to the previous statistical analysis of the $\log(\mathrm{PM}_{10})$ series, the following model is suggested here to fit the data

$$\log(\mathrm{PM}_{10,t}) = \mu + \alpha_1 \mathrm{sat}_t + \alpha_2 \mathrm{sun}_t + \beta_1 \cos_t + \beta_2 t + Y_t; \tag{14}$$

$$Y_t = \sum_{i=1}^{p_t} \phi_i(t) Y_{t-i} + \sigma_t \epsilon_t, \tag{15}$$

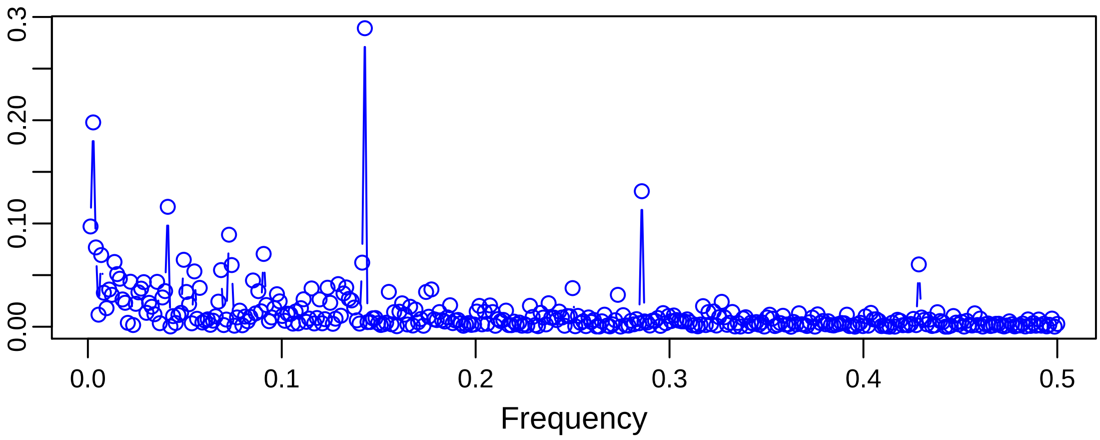



Figure 5: Periodogram of the $\log(PM_{10})$ time series.

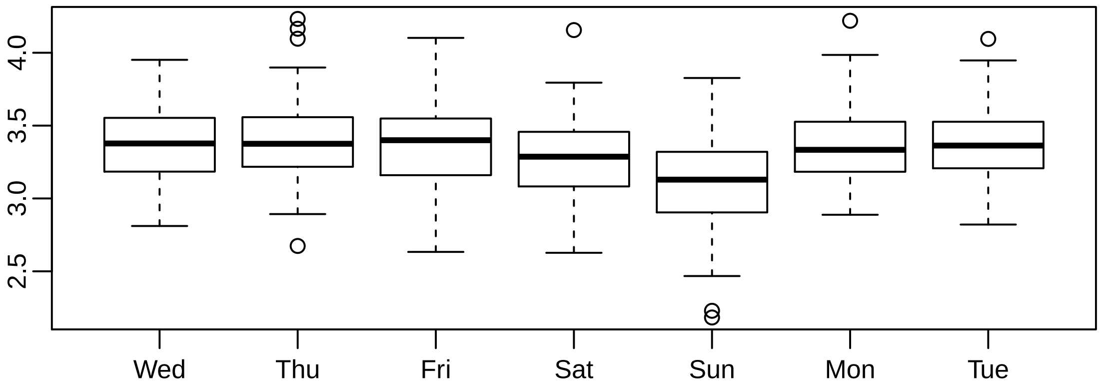



Figure 6: Daily box-plots of the $\log(PM_{10})$ time series.

with the sinusoidal covariate: $\cos_t = \cos(\frac{2\pi t}{365.25})$, $t = 1, \ldots, 728$; the linear term $t$; and a "day of the week" factor with the levels: Week (the reference level); Saturday (represented by the dummy variable $\text{sat}_t$ which takes value 1 for Saturdays); and Sunday ($\text{sun}_t$ which takes value 1 for Sundays) and $\mathcal{S} = 7$. The above model means that in the business days the regular level of $\log(PM_{10})$ is $\mu$, on Saturdays it suffers an increase of $\alpha_1$ and on Sundays it is increased by $\alpha_2$ and it has a long-run cyclic trend, represented by $\cos_t$ and a linear term $t$. The terms $\sin_t$ and $\cos_t$ turned out to be statistically insignificant and they were removed from the model.

The model in Equations 14 and 15 will be fitted based on following two steps procedure: (1) the linear model in (14) will be estimated through the ordinary least squares procedure; and (2) the PAR model in (15) will be fitted to the residuals of the linear model in step (1), where the AR orders $p_1, \ldots, p_{\mathcal{S}}$ will be identified through the Schwartz Information Criterion (BIC) proposed by Schwarz [1978] and adapted to the periodic scenario by McLeod [1994].

At the first step, the linear model in Equation 14 was fitted and the estimated coefficients are displayed in Table 4. As expected, there were negative effects of Saturday and Sunday, which led to a decrease of $\log(PM_{10})$ levels during the weekends.

Table 4: Estimated coefficients of the linear model.

| Parameter | $\mu$ | $\alpha_1$ | $\alpha_2$ | $\beta_1$ | $\beta_2$ |
|---|---|---|---|---|---|
| Estimate | 3.2650 | -0.0921 | -0.2579 | 0.0640 | 0.0003 |

The BIC criterion was used to identify the order of the model (see Sarnaglia et al. [2010] for more details) and the results are displayed in Table 5. In order to keep consistency with the simulation study, $c = 3.06$ was fixed in the Huber function (Equation 12). Note that the PAR model with better (smaller) BIC was obtained by the RYWE, which indicates that this estimator provides a good compromise between adjustment and parsimony.

Table 5: Selected AR orders using the BIC.

| Estimator | BIC | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $p_6$ | $p_7$ |
|---|---|---|---|---|---|---|---|---|
| YWE | $-2156.25$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| RYWE | $-2205.94$ | 1 | 1 | 4 | 2 | 2 | 1 | 1 |
| RLSE | $-2200.15$ | 1 | 1 | 4 | 1 | 2 | 1 | 1 |

The estimates of the PAR coefficients provided by YWE, RYWE and RLSE methods are given in Table 6. Based on these results, it is clear the presence of Periodic Correlation in the data, since the AR coefficients and orders are not constant over the seasons. In general, the methods selected different orders and presented quite different coefficient estimates. This indicate that the high levels of the pollutant were stronger enough to provoke changes in the parameter estimates, that is, this reveals that the high levels of the pollutant $PM_{10}$ presented the effects of additive outliers according to the discussion presented in the Simulation Section.

Table 6: Estimates of the AR coefficients for YWE, RYWE and RLSE.

| Estimator | $i$ | $\nu$ | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $\hat{\phi}_i(\nu)$ | 1 | 0.626 | 0.532 | 0.485 | 0.374 | 0.595 | 0.312 | 0.498 |
| $\tilde{\phi}_i(\nu)$ | 1 | 0.614 | 0.482 | 0.630 | 0.529 | 0.595 | 0.361 | 0.474 |
| | 2 | 0.000 | 0.000 | -0.107 | -0.143 | -0.258 | 0.000 | 0.000 |
| | 3 | 0.000 | 0.000 | 0.451 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 4 | 0.000 | 0.000 | -0.374 | 0.000 | 0.000 | 0.000 | 0.000 |
| $\breve{\phi}_i(\nu)$ | 1 | 0.661 | 0.513 | 0.479 | 0.376 | 0.638 | 0.293 | 0.522 |
| | 2 | 0.000 | 0.000 | -0.018 | 0.000 | -0.167 | 0.000 | 0.000 |
| | 3 | 0.000 | 0.000 | 0.271 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 4 | 0.000 | 0.000 | -0.240 | 0.000 | 0.000 | 0.000 | 0.000 |

The fitting performance will be accessed through the in-sample Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), symmetric MAPE (sMAPE) and Median of Absolute Deviation (MAD). The RMSE and the MAD are well-known and, for a discussion of MAPE and sMAPE quantities see Flores [1986]. The results are presented in Table 7. As can be seen, the RLSE are RYWE very competitive by presenting very similar results and they are slightly smaller than YWE method. This may corroborate the previous discussion related the effect of high level concentrations on the model estimation.

Table 7: Fitting performance of the estimated models.

| Statistic | Estimator | | |
|---|---|---|---|
| | YWE | RYWE | RLSE |
| RMSE | 0.2246 | 0.2245 | 0.2235 |
| MAPE | 5.3420 | 5.2545 | 5.2707 |
| sMAPE | 2.6454 | 2.6024 | 2.6104 |
| MAD | 0.2120 | 0.1988 | 0.2078 |

Figures 7, 8 and 9 present the classic ACF of the residuals of each model. It can be seen that all the models were able to fully explain the correlation structure of the data, despite the eventual outliers effect. Based on the ACF of the residuals, the three estimation methods are comparable since all the estimated residuals look like a white noise process.

Finally, for all models, the residuals have not passed the Jarque-Bera normality test [Jarque and Bera, 1980], presenting $p$-values $< 0.05$ which is an expected result due to the skewness revealed in the data.

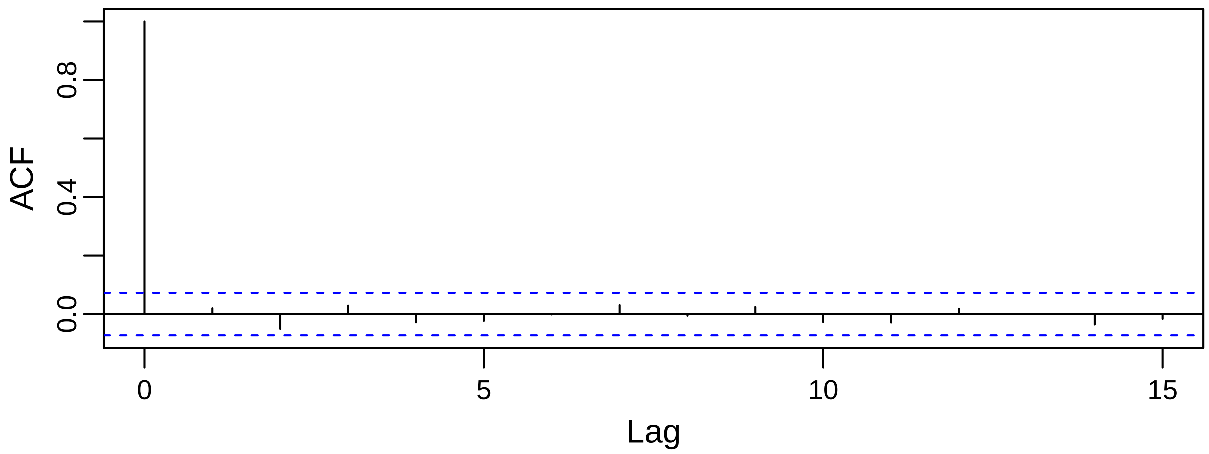



Figure 7: ACF of the residuals of the YWE fit.

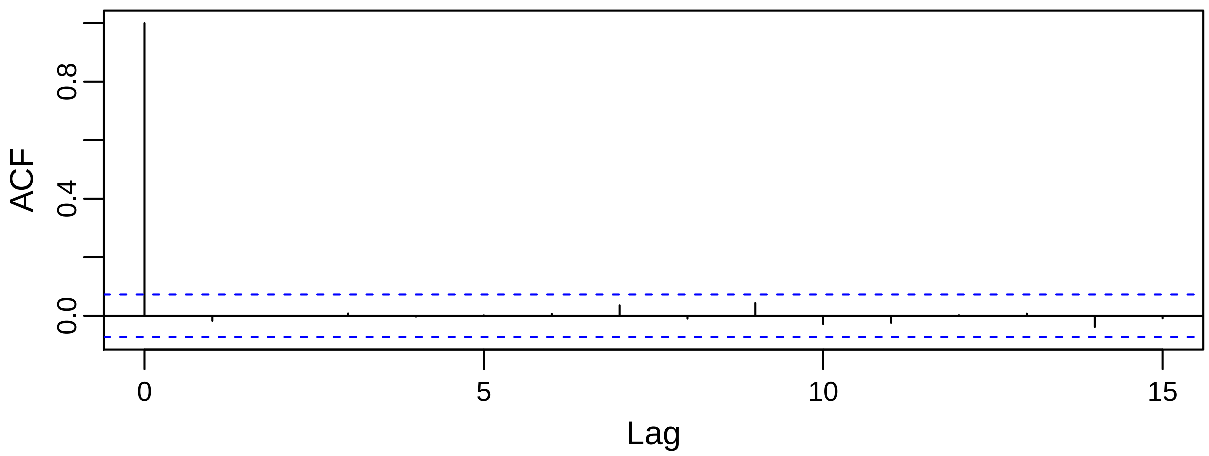



Figure 8: ACF of the residuals of the RYWE fit.

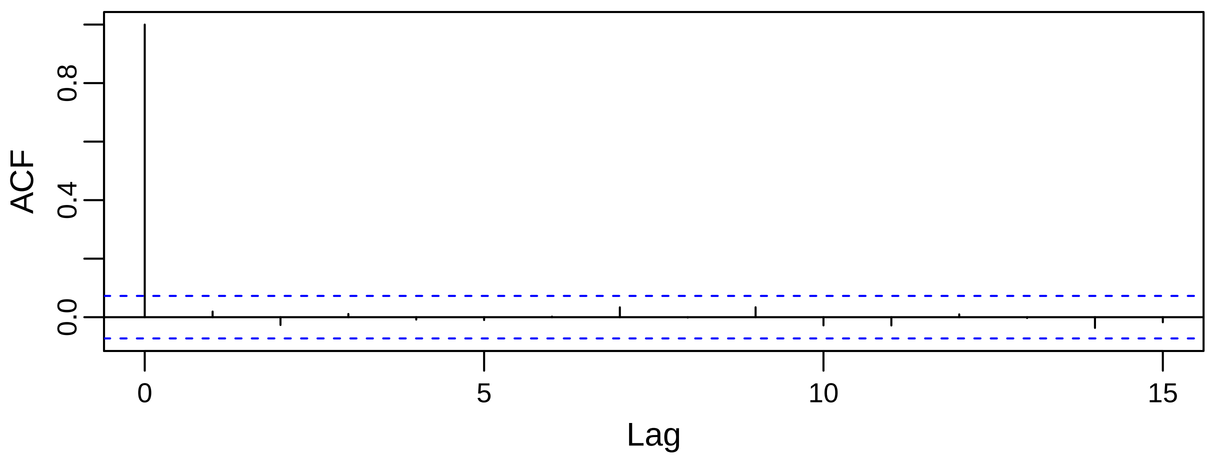



Figure 9: ACF of the residuals of the RLSE fit.

## 5 Conclusions

This paper reviews different estimation methodologies for PAR models. More specifically, the method considered are: the so-called YWE [McLeod, 1994], the RYWE [Sarnaglia et al., 2010] and the RLSE [Shao, 2008]. The finite sample performance of these methods was compared through a Monte Carlo experiment. The performance of RLSE is remarkably good under uncontaminated and contaminated scenarios, even under asymmetric errors, which violates Assumption 3. The RYWE is quite resistant to outliers, however it has a poor performance under asymmetric errors, mainly under weak correlation scenarios. As expected, YWE empirical distribution is resistant to departures from normality, however this estimator is completely affected by the presence of outliers. In order to illustrate the methodologies considered in this paper, the daily mean $PM_{10}$ concentrations collected at the air quality monitoring station, located at Enseada do Suá, ES, Brazil, was considered as an application. The estimation and modelling results revealed outlier effects on the estimates.

## 6 Acknowledgements

The results in this paper are part of the Master thesis of the first author in the PPGEA-UFES under supervision of the second and third authors. The authors gratefully acknowledge partial financial support from FAPES/ES, CAPES/Brazil and CNPq/Brazil. This paper was revised when Prof. Valderio Reisen was visiting CentraleSupélec (01 to 03/2018). This author is indebted to CentraleSupélec for its financial support. The authors are grateful to the Editor and the referee for the time and efforts in providing very constructive and helpful comments that have led to clarify and substantially improve the quality of the paper.

## References

P. L. Anderson and M. M. Meerschaert. Parameter estimation for periodically stationary time series. *Journal of Time Series Analysis*, 26(4):489–518, 2005.

I. Basawa and R. Lund. Large sample properties of parameter estimates for periodic ARMA models. *Journal of Time Series Analysis*, 22(6):651–663, 2001.

P. Bloomfield, H. L. Hurd, and R. B. Lund. Periodic correlation in stratospheric ozone data. *Journal of Time Series Analysis*, 15(2):127–150, 1994.

G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control.* Holden day, San Francisco, CA, 1970.

P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods.* Springer Verlag, New York, second edition, 1991.

F. A. Fajardo, V. A. Reisen, C. Lévy-Leduc, and M. S. Taqqu. M-periodogram for the analysis of long-range-dependent time series. *Statistics*, pages 1–19, 2018.

B. E. Flores. A pragmatic view of accuracy measurement in forecasting. *Omega*, 14(2):93–98, 1986.

W. A. Fuller. *Introduction to Statistical Time Series.* Wiley, New York, 1976.

W. A. Gardner and L. E. Franks. Characterization of cyclostationary random signal processes. *IEEE Transactions on Signal Processing*, 21:4–14, 1975.

E. G. Gladyshev. Periodically correlated random sequences. *Sov. Math.*, 2:385–388, 1961.

H. L. Hurd and N. L. Gerr. Graphical methods for determining the presence of periodic correlation. *Journal of Time Series Analysis*, 12(4):337–350, 1991.

C. M. Jarque and A. K. Bera. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259, 1980.

R. Lund and I. Basawa. *Asymptotics, Nonparametrics, and Time Series*, chapter Modeling and Inference for Periodically Correlated Time Series, pages 37–62. CRC Press, 1999.

R. B. Lund, H. L. Hurd, P. Bloomfield, and R. L. Smith. Climatological time series with periodic correlation. *Journal of Climate*, 8:2787–2809, 1995.

Y. Ma and M. G. Genton. Highly robust estimation of the autocovariance function. *Journal of time series analysis*, 21(6):663–684, 2000.

A. McLeod. Diagnostic checking of periodic autoregression models with application. *Journal of Time Series Analysis*, 15(2):221–223, 1994.

D. J. Noakes, A. I. McLeod, and K. W. Hipel. Forecasting monthly riverflow time series. *International Journal of Forecasting*, 1(2):179–190, 1985.

M. B. Priestley. *Spectral Analysis and Time Series*. Academic press, 1981.

V. A. Reisen, A. J. Q. Sarnaglia, N. C. Reis Jr, C. Lévy-Leduc, and J. M. Santos. Modeling and forecasting daily average $PM_{10}$ concentrations by a seasonal long-memory model with volatility. *Environmental Modelling & Software*, 51:286–95, 2014.

V. A. Reisen, C. Lévy-Leduc, and M. S. Taqqu. An M-estimator for the long-memory parameter. *Journal of Statistical Planning and Inference*, 187(2):44–55, 2017.

V. A. Reisen, E. Z. Monte, G. d. C. Franco, A. M. Sgrancio, F. A. F. Molinares, P. Bondon, F. A. Ziegelmann, and B. Abraham. Robust estimation of fractional seasonal processes: Modeling and forecasting daily average $SO_2$ concentrations. *Mathematics and Computers in Simulation*, 146: 27–43, 2018.

P. J. Rousseeuw and C. Croux. Alternatives to the median absolute deviation. *Journal of the American Statistical association*, 88(424):1273–1283, 1993.

H. Sakai. Circular lattice filtering using Pagano's method. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(2):279–287, 1982.

A. J. Q. Sarnaglia, V. A. Reisen, and C. Lévy-Leduc. Robust estimation of periodic autoregressive processes in the presence of additive outliers. *Journal of Multivariate Analysis*, 101(9):2168–2183, 2010.

A. J. Q. Sarnaglia, V. A. Reisen, and P. Bondon. Periodic ARMA models: Application to particulate matter concentrations. In *23rd European Signal Processing Conference*, pages 2181–2185, 2015.

A. J. Q. Sarnaglia, V. A. Reisen, P. Bondon, and C. Lévy-Leduc. A robust estimation approach for fitting a PARMA model to real data. In *IEEE Statistical Signal Processing Workshop*, pages 1–5, 2016.

G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.

Q. Shao. Robust estimation for periodic autoregressive time series. *Journal of Time Series Analysis*, 29(2):251–263, 2008.

Q. Shao and R. Lund. Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *Journal of Time Series Analysis*, 25(3):359–372, 2004.

R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications – With R Examples*. Springer Texts in Statistics. Springer, Cham, fourth edition, 2017.

J. B. Souza, V. A. Reisen, G. C. Franco, M. Ispány, P. Bondon, and J. M. Santos. Generalized additive model with principal component analysis: An application to time series of respiratory disease and air pollution data. *JRSS-Appl. Statist., Series C*, 6(2):453—480, 2018.

G. Tiao and M. Grupe. Hidden periodic autoregressive-moving average models in time series data. *Biometrika*, pages 365–373, 1980.