PARMA Models with Applications in R

Anna E. Dudek, Harry Hurd and Wioletta Wójtowicz

1 Introduction

Any second-order random process that is generated by the mixing (in the workings of a system) of randomness and periodicity will likely have the structure of periodic correlation. Precisely, a random sequence $\{X_t, t \in \mathscr{Z}\}$ with finite second moments is called periodically correlated with period T (PC-T) if it has periodic mean and covariance functions, e.g.

$$E\left(X_t\right) = E\left(X_{t+T}\right) \quad \text{and} \quad Cov\left(X_t, X_s\right) = Cov\left(X_{t+T}, X_{s+T}\right) \tag{1}$$

for each $t, s \in \mathscr{Z}$. To avoid ambiguity, the period T is taken as the smallest positive integer such that (1) holds.

The studies of PC sequences, which were initiated by [15], result in the appreciable theory and some practical approaches as well. Many real data have periodic structure, so they can be described by periodically correlated sequences. The cyclic nature of environmental and social phenomena impart a seasonal mean and correlation structure into many climatological [21, 26] as well as economical [11, 19, 20] time series. Other examples of PC data could be found in e.g. [9, 16, 22]. Gardner investigated perceptively the nature of cyclostationarity [12, 13]. Hurd and Miamee [16] provide substantial study of PC issues with many motivating and illustrative examples.

A.E. Dudek (✉)
AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland
e-mail: aedudek@agh.edu.pl

H. Hurd
Department of Statistics and Operations Research, University of North Carolina
at Chapel Hill, Chapel Hill 27599-3260, USA

W. Wójtowicz
Institute of Applied Computer Science, Cracow University of Technology,
Al. Jana Pawla II 37, 31-864 Cracow, Poland

In this paper we focus on identification and estimation of PC structure and on PARMA (periodic autoregressive–moving-average) modelling introduced in [27, 28]. PARMA sequences are formed by introducing time-varying coefficients into the ARMA set-up; under a mild restriction on the coefficients, the resulting PARMA sequence is periodically correlated (PC) [28]. The periodically correlated processes and PARMA modelling are studied in [4, 5, 9] and lately also in [1]. The methods proposed by those authors, although based on statistical principles, are not well known or available to a wide audience. Currently, difficulties connected with analysis of periodic sequences arise from deficient knowledge about mathematical achievement in this field. The intent of this work is to provide an accessible package for periodic time series analysis in R. In this paper, we present an R package called perARMA that is available from the Comprehensive R Archive Network (CRAN) (see online supplementary material and [10]). The package implements non-parametric analysis, characterization and identification, PARMA model fitting and prediction. Missing observations are allowed in some characterization procedures. The implemented procedures are loosely based on the Hurd's Matlab functions available from his Web page and introduced in [16]. As a result, the applied researcher obtains quite easy and very intuitive tool that can be easily used in many applications. To our knowledge there is no R package (under CRAN) for PARMA time series analysis although the pear [3] and partsm [18] packages provide for PAR analysis.

The paper is organized in the following way. In Sect. 2 we provide some theoretical background of PARMA time series. Section 3 presents analysis of real dataset from energy market. In Sect. 4 the estimation of full PARMA model for simulated dataset is performed. Finally, Sect. 5 concludes our study.

2 PARMA Time Series Analysis

PARMA modelling arises from the introduction of periodic correlation of PC sequences into a stationary ARMA model when the coefficients of the model are allowed to vary periodically in time. Precisely, the random sequence X_t is called PARMA(p,q) with period T if it satisfies

$$X_t = \sum_{j=1}^{p} \phi_j(t) X_{t-j} + \sum_{k=1}^{q} \theta_k(t) \xi_{t-k} + \sigma(t) \xi_t, \tag{2}$$

where $\phi_j(t) = \phi_j(t+T), \theta_k(t) = \theta_k(t+T), \sigma(t) = \sigma(t+T)$ for all $j = 1, \ldots, p$, $k = 1, \ldots, q$ are periodic coefficients, and ξ_t is mean zero white noise with variance equal to one.

The PARMA systems are widely applied in modelling climatology [4, 5], meteorology [24], hydrology [2, 21, 26, 28] and economics data [7, 11].

PARMA time series analysis is performed in three main processing steps: (1) identification, (2) parameter estimation and (3) diagnostic checking. The same general

process is used in the stationary case, where the tools are simpler. In the following sections, we describe the tools that are essential to achieve each of these three steps. But first we focus on the ARMA(p,q) fitting process to remind the reader of the general idea.

Identification in the stationary case refers to the determination of the model order parameters p and q, which provide an adequate fit to the data. Initial guesses of p and q are usually suggested from the identification tools, which are the sample covariance, sample autocorrelation (ACF) and sample partial autocorrelation (PACF). Parameter estimation refers to the process of estimating the values of the parameters in the chosen representation. For AR models we can use the Yule–Walker equations. For general ARMA we use maximum likelihood. Diagnostic checking in the stationary case consists of determining if the residuals (based on some parameter estimates) are consistent with white noise. If not, then modifications to p and q are made based essentially on the application of the identification step to the residuals (determine what structure is not yet explained) and estimation is rerun.

Below we present how those ideas can be transformed to the periodic nonstationary case. In the subsequent we assume that $\{X_1, X_2, \ldots, X_N\}$ is a mean zero PC time series with period T. Moreover, we assume without loss of generality that the data record contains d full cycles, e.g. N is an integer multiple of period T ($N = dT$).

2.1 Identification of PC-T Structure

There are two processes that we include under the heading of identification: (1) the determination of period T when it is unknown and some basic characterizing quantities such as sample periodic mean $\widehat{m}_{t,d}$, sample periodic deviation $\widehat{\sigma}_d(t)$ and sample periodic covariance $\widehat{R}_d(t + \tau, t)$; (2) the determination of p and q, the orders of the PAR and PMA parts of a PARMA model.

Preliminary Identification

An important preliminary step in the identification process is the determination of period T when it is not known. In this case, the *periodogram* and *squared coherence statistic* can be used. The usual periodogram can detect additive frequency components in the time series and this includes frequencies belonging to the additive periodic mean. So if a periodic mean is present, the periodogram can illuminate its frequencies and help in the determination of T. The value of T may also be inferred from spacing of the support lines in the (harmonizable) spectral measure of a PC process. These support lines are seen empirically in the images produced by the magnitude-squared coherence statistic. Using both simultaneously gives a complete picture.

To clarify the idea of the magnitude-squared coherence, we need to use some features of the spectral measure of a PC sequence. The spectral measure of periodically correlated sequence is determined on the two-dimensional set $[0, 2\pi) \times [0, 2\pi)$, so we always deal with the pairs of frequencies $(\lambda_p, \lambda_q) \in [0, 2\pi)^2$, and the support

of the measure is contained in the subset of parallel lines $\lambda_q = \lambda_p + 2j\pi/T$ for $j = -(T-1), \ldots, -1, 0, 1, \ldots, (T-1)$. For more details see [14, 16].

The concept of determining the period length using the squared coherence statistic directly corresponds to its features. We begin with Discrete Fourier Transform (DFT), $\tilde{X}_N(\lambda_j)$, of the series $\{X_1, X_2, \ldots, X_N\}$, defined for the Fourier frequencies $\lambda_j = 2\pi j/N$, $j = 0, 1, \ldots, N-1$. The squared coherence statistic,

$$\left|\hat{\gamma}(\lambda_p, \lambda_q, M)\right|^2 = \frac{\left|\sum_{m=1}^{M-1} \tilde{X}_N(\lambda_{p-M/2+m}) \overline{\tilde{X}_N(\lambda_{q-M/2+m})}\right|^2}{\sum_{m=1}^{M-1} \left|\tilde{X}_N(\lambda_{p-M/2+m})\right|^2 \sum_{m=1}^{M-1} \left|\tilde{X}_N(\lambda_{q-M/2+m})\right|^2} \qquad (3)$$

is computed for a specified collection of pairs $(\lambda_p, \lambda_q) \in [0, 2\pi)^2$; note it is the sample correlation between two M-dimensional complex-valued vectors and satisfies $0 \le \left|\hat{\gamma}(\lambda_p, \lambda_q, M)\right|^2 \le 1$. Having computed (3) for (λ_p, λ_q) in some subset of $[0, 2\pi)^2$, one may determine frequency pairs indexed by (p, q) for which the sample correlation is significant in the sense that threshold determined by the distribution of $\left|\hat{\gamma}(\lambda_p, \lambda_q, M)\right|^2$, under the null hypothesis of stationary white noise, is exceeded [16, p. 310]. Plotting those points on the square $[0, 2\pi)^2$, we can say something about the nature of the analyzed time series, according to some general rules:

- if in the square only the main diagonal appears, then X_t is a stationary time series;
- if there are some significant values of statistic and they seem to lie along the parallel equally spaced diagonal lines, then X_t is likely PC-T, where T is the "fundamental" line spacing. Algebraically, T would be the gcd of the line spacings from the diagonal; for a sequence to be PC-T, not all lines are required to be present;
- if there are some significant values of statistic but they occur in some non-regular places, then X_t is a nonstationary time series in other than periodic sense; but note there are many hypotheses being tested, so some threshold exceedances are to be expected.

We need to comment also the choice of the parameter M that controls the length of the smoothing window appearing in (3). Too small or too large values of M can affect the results [16, p. 311]. Hurd and Miamee suggest to observe the results for a collection of smoothing parameters; a suggested beginning choice is $M = 8, 16, 32$. Under the null case the sample-squared coherence statistics has probability density

$$p\left(|\gamma|^2\right) = (M-1)\left(1 - |\gamma|^2\right)^{M-2}, \quad 0 \le |\gamma|^2 \le 1, \qquad (4)$$

because, for X_t a Gaussian white noise, $\tilde{X}_N(\lambda_j)$ are complex Gaussian with uncorrelated real and imaginary parts for each j. As a result, to determine the squared coherence α-threshold we use

$$|\gamma|_\alpha^2 = 1 - \exp\left(\log(\alpha)/(M-1)\right). \qquad (5)$$

For more details we refer the reader to [14, 16].

When the period length is known, the estimation of basic sequence statistics: periodic mean, periodic deviation and periodic covariance is possible. Detailed description of all estimators presented below can be found in [16].

The sample periodic mean function is given by

$$\widehat{m}_{t,d} = \frac{1}{d} \sum_{k=0}^{d-1} X_{t+kT}. \tag{6}$$

The sample periodic covariance function is defined for any lag τ as

$$\widehat{R}_d\,(t+\tau, t) = \frac{1}{d} \sum_{k=0}^{d-1} \left(X_{t+kT+\tau} - \widehat{m}_{t+\tau,d}\right)\left(X_{t+kT} - \widehat{m}_{t,d}\right) \tag{7}$$

for $t = 0, 1, \ldots, T-1$. The number of τ for which (7) is evaluated can be determined (or estimated) from the usual (stationary) ACF. Then it is quite straightforward to get the periodic sample deviation by putting

$$\widehat{\sigma}_d\,(t) = \widehat{R}_d\,(t, t)\,. \tag{8}$$

The Fourier representation of the covariance function [15] for PC-T processes is based on the periodicity $R_d(t + \tau, t) = R_d(t + \tau + T, t + T)$ and the Fourier series representation

$$R_d(t + \tau, t) = \sum_{k=0}^{T-1} B_k(\tau) \exp(i2kt\pi/T), \tag{9}$$

where $B_k(\tau)$ ($\tau \in \mathscr{Z}$ and $k = 0, 1, \ldots, T-1$) are the Fourier coefficients given by

$$B_k(\tau) = \frac{1}{T} \sum_{t=0}^{T-1} \exp(-i2kt\pi/T) R_d(t + \tau, t). \tag{10}$$

The problem of computing the Fourier representation is thus reduced to the problem of determining the coefficients $B_k(\tau)$. Moreover, $B_0(\tau)$ is always non-negative definite and hence is the covariance of stationary sequence. The Fourier coefficients of the sample covariance, for $k = 0, 1, \ldots, T-1$ and $\tau \in$ set of lags, can be estimated

$$\widehat{B}_{k,dT}(\tau) = \frac{1}{dT} \sum_{t \in I_{dT,\tau}} \left(X_{t+\tau} - \widehat{m}_{t+\tau,d}\right)\left(X_t - \widehat{m}_{t,d}\right) e^{-i2\pi kt/T}, \tag{11}$$

where

$$I_{dT,\tau} = \begin{cases} [0, dT - \tau - 1] & \text{for } \tau \geq 0 \\ [-\tau, dT - 1] & \text{for } \tau < 0. \end{cases}$$

Hurd and Miamee proposed a test with the null hypothesis $B_k(\tau) = 0$ for fixed τ and k, which is based on the properties of the sample Fourier transform of $Y_{t,\tau} = (X_{t+\tau} - \widehat{m}_{t+\tau,d})(X_t - \widehat{m}_{t,d})$. The p values for the test of $|B_k(\tau)|^2 = 0$ are based on the ratio of magnitude squares of amplitudes of a high resolution Fourier decomposition. Magnitudes for the frequency corresponding to index k are compared to the magnitudes of neighbouring frequencies (via the F distribution) [16, pp. 272–282, 288–292].

Determination of p and q

Before fitting a periodic model (PARMA(p,q)) to data, the orders of maximum lag p and/or q are required. Similarly to the stationary ARMA, we need to look at the sample autocovariance (ACF) and the sample partial autocovariance (PACF). Of course, their periodic versions are essential in that case. Below we present formulas for both of them.

Sample periodic ACF for $t = 0, 1, \ldots, T-1$ is defined as

$$\widehat{\rho}_d(t+\tau, t) = \frac{\widehat{R}_d(t+\tau, t)}{\widehat{\sigma}_d(t+\tau)\widehat{\sigma}_d(t)}, \tag{12}$$

where τ is a lag. To calculate confidence bands for $\widehat{\rho}_d(t+\tau, t)$ the Fisher transformation

$$Z = \frac{1}{2}\log\frac{1+\widehat{\rho}}{1-\widehat{\rho}} \tag{13}$$

is used, producing Z that are approximately normal $N(\mu_Z, \sigma_Z^2)$, where $\mu_Z \approx \zeta + (1/2d)\rho$, ζ is the Fisher transformation of ρ and $\sigma_Z^2 = 1/(d-3)$. The confidence limits for ρ can be calculated using those for Z with the assumption that the term $(1/2d)\rho$ can be ignored.

Two useful tests may be additionally constructed. First, a test for null hypothesis $\rho(t+\tau, t) \equiv \rho(\tau)$, where $\rho(\tau)$ is some unknown constant, (no time variation at lag τ) is based on statistic

$$S_\rho^2(\tau) = \frac{(Z_t - \widehat{Z}_t)^2}{1/(d-3)}, \tag{14}$$

which under the null hypothesis is approximately $\chi^2(T-1)$. The second, a test for null hypothesis $\rho(t+\tau, t) \equiv 0$ for some specific τ, (no correlation at this lag τ) is equivalent to the test for $\mu_Z = 0$. For more details about those tests we refer the reader to [8, p. 399] and [16, p. 285].

The sample periodic PACF, for $t = 0, 1, \ldots, T-1$ and $n \in$ set of orders is defined as

$$\hat{\pi}_{n+1}(t) = \frac{\hat{R}_d(t+1, t-n) - \hat{\mathbf{r}}'_{t-n,t:t-n+1}\hat{\boldsymbol{\alpha}}_n^{(t+1)}}{\hat{\sigma}_n(t+1)\hat{\sigma}_n(t-n)}, \tag{15}$$

where

$$\hat{\sigma}_n^2(t+1) = \hat{R}_d(t+1, t+1) - \hat{\mathbf{r}}'_{t+1,t:t-n+1}\hat{\boldsymbol{\alpha}}_n^{(t+1)},$$

$$\hat{\sigma}_n^2(t-n) = \hat{R}_d(t-n,t-n) - \hat{\mathbf{r}}'_{t-n,t:t-n+1}\hat{\boldsymbol{\beta}}_n^{(t-n)}$$

and

$$\hat{\boldsymbol{\alpha}}_n^{(t+1)} = \hat{\mathbf{r}}^{-1}(t,n)\hat{\mathbf{r}}_{t+1,t:t-n+1},$$
$$\hat{\boldsymbol{\beta}}_n^{(t-n)} = \hat{\mathbf{r}}^{-1}(t,n)\hat{\mathbf{r}}_{t-n,t:t-n+1}.$$

The confidence bands are based on the asymptotic normality assumption [6, Chap. 3].

2.2 Estimation of PARMA Parameters

Parameter estimation refers to the process of estimating the values of the parameters in the chosen representation. The full PARMA model has T parameters for each autoregressive order, for each moving average order and for $\sigma(t) = \theta_0(t), t = 0, 1, \ldots, T-1$, and thus a total of $T(p+q+1)$ parameters, so the cost of computing all of them could be quite high. Therefore, it is recommended first trying to fit simple models, which explain the data with the fewest possible parameters.

Two useful methods for estimation of periodic models coefficients are implemented in perARMA package: the Yule–Walker moment estimators for the general PAR model and approximate maximum likelihood estimation for the general PARMA. It is easy to show that for PAR models both techniques are equivalent, e.g. for PAR(1) model provide following estimators of coefficients:

$$\hat{\phi}(t) = \frac{\hat{\gamma}_t(1)}{\hat{\gamma}_{t-1}(0)}, \qquad \hat{\sigma}^2(t) = \hat{\gamma}_t(0) - \hat{\phi}(t)\hat{\gamma}_t(1), \qquad t = 0, 1, \ldots, T-1, \quad (16)$$

where $\hat{\gamma}_t(j) = \frac{1}{d}\sum_{n=0}^{d-1} X_{nT+t-j}X_{nT+t}$ and $d = N/T$ is number of cycles.

The Yule–Walker estimation is a simple method, but whenever PMA terms are present, it will be inadequate. In these cases maximum likelihood estimation is used. But whenever there are a large number of parameters, the optimization has difficulty due to presence of local maxima and due to slowness caused by the approximation of derivatives.

Rather than searching over the entire Φ, Θ parameter space, the dimension of the search space can often be substantially reduced by transforming Φ, Θ to the Fourier coefficients $\mathbf{A}$, $\mathbf{B}$ appearing in the DFT representation introduced in [17]:

$$\phi_j(t) = a_{j,1} + \sum_{n=1}^{\lfloor T/2 \rfloor} a_{j,2n}\cos(2\pi nt/T) + \sum_{n=1}^{\lfloor T/2 \rfloor} a_{j,2n+1}\sin(2\pi nt/T), \quad (17)$$

$$\theta_k(t) = b_{k,1} + \sum_{n=1}^{\lfloor T/2 \rfloor} b_{k,2n}\cos(2\pi nt/T) + \sum_{n=1}^{\lfloor T/2 \rfloor} b_{k,2n+1}\sin(2\pi nt/T) \quad (18)$$

for $j = 1, \ldots, p$, $k = 0, \ldots, q$ and $t = 0, 1, \ldots, T - 1$. Reduction of the parameter search to subspaces of **A**, **B** can be easily accomplished by constraining some (or even most) frequencies to have zero amplitude. Often this can be justified by physical considerations that correspond to smoothness in the time dependence. Since the mapping (DFT) from Φ, Θ to **A**, **B** is linear and invertible, a linear subspace of **A**, **B** transforms back to a linear subspace of Φ, Θ, although our methods make no explicit use of this fact. In our experience, reduction of the search space in this manner not only reduces computation time, but tends to produce unique solutions, whereas the likelihood often has many local extremes in the entire space Φ, Θ.

This parameterization was first used in [28] in the context of maximum likelihood parameter estimation and in [23] where are developed expressions for the harmonizable spectral density in terms of the parameters **A**, **B**.

Computation speed of PARMA maximum likelihood estimation can be further improved by the use of Ansley's transformation, first applied to PARMA(p,q) sequences in [27, 28]. This method focuses on the conditional version which ignores the first $m = \max(p; q)$ samples in order to avoid the cumbersome calculation of the full covariance.

2.3 Goodness of Fit and Model Selection

To confirm that fitted model is appropriate the residuals should be examined to ensure independence (whiteness) and normality, but this does not differ from stationary case. However, in the periodic case we do permit the residuals to be periodic white noise, this can be easily converted to white noise by scaling out the empirical periodic variance.

Sometimes there are several sets of model parameters that give reasonable fits. Then to choose the simplest model that explains data the best the penalized likelihoods could be computed. The AIC and BIC methods compute penalties for the number of parameters used and thus encourage the simplicity (or parsimony) of the selected fit. The parameter set that minimizes the penalized likelihoods is considered to the best fit. To calculate AIC and BIC values we use following (k is the total number of parameters in the parameter set Φ, Θ or **A**, **B** in the Fourier parametrization and N is the number of linearly independent samples):

$$AIC^{(k)} = -2 \ln L\left(\widehat{\Phi}, \widehat{\Theta}\right) + 2k,$$
$$BIC^{(k)} = -2 \ln L\left(\widehat{\Phi}, \widehat{\Theta}\right) + 2k \log N.$$

3 Real Data Example

In this section identification of PC-T structure, PAR modelling and forecasting procedure is developed for a non-volatile segment of real time series contained hourly observations from volumes of energy traded on the Nord Pool Spot exchange. The Nord Pool Spot runs the largest market for electrical energy in the world, measured in volume traded (TWh) and in market share. The data were found on the Nord Pool Spot Exchange Web page http://www.npspot.com.

3.1 Data

Analyzed data aggregates volume in MWh bought and sold by participants in each bidding area of this market. Thus, this data reflects the demand on the energy on the daily basis. In Fig. 1 there is presented `volumes_2010` time series that contains all hourly observations within 2010 year with non-volatile segment identified. This segment (after removing weekends records) as `volumes` time series is analyzed in our study. It contains hourly observations of the volumes of energy traded from 6 July to 31 August of 2010 ($N = 984$ records). We decided to omit the weekends, as including them would increase the complexity of the model with unknown benefit. So for simplicity and clarity, we focus on understanding the dynamics within the trading part of the week and leave the weekend effects for a future study.
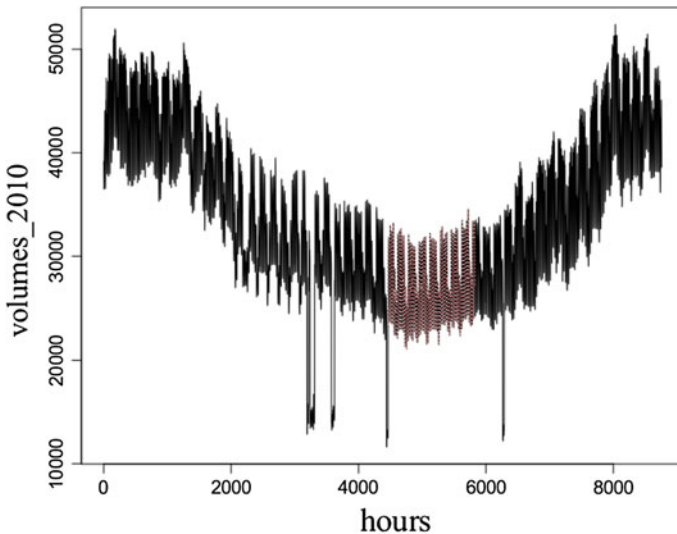

Fig. 1 Volumes of energy traded hourly on the Nord Pool Spot Exchange in 2010 with the non-volatile segment of volumes data (from 6 July 2010 to 31 August 2010)

3.2 Non-parametric Spectral Analysis

As one may expect for `volumes` time series presented in Fig. 1, the daily periodic structure is observed. Even though the period length $T = 24$ is rather obvious in this case, in some cases it may need to be estimated, as the dominant frequencies of the process may not be known at the beginning of analysis. Therefore in our example, we confirm that the period length was properly chosen.

In the `perARMA` package two procedures, based on the spectral analysis, are useful for estimating the period: `pgram()` and `scoh()`. Function `pgram()` plots periodogram of the series and provides test to find candidate for the period length assuming that the period of the second-order structure is the same as the period of the first order structure (i.e. in the instantaneous mean $E\{X(t)\}$ of the series itself). Recall that the FFT index j (where a big peak occurs) corresponds to the number of cycles in the FFT window, so the period corresponding to the index $j = 41$ where the first highly significant peak occurs can be easily computed as $T = 984/41 = 24$ (see Fig. 2). The function `scoh()`, which computes and plots the magnitude-squared coherence, can be used to confirm the presence of the PC structure, to estimate the period of the PC structure and to determine the strength of the PC structure on the different support lines. Specifically, the magnitude-squared coherence $|\hat{\gamma}(\lambda_p, \lambda_q, M)|^2$ is computed in a specified square set of (λ_p, λ_q) and using a specified smoothing window, M. The perception of this empirical spectral coherence is aided by plotting the coherence values only at points where threshold is exceeded. To ensure that periodic structure seen in the spectral coherence image is not a consequence of an additive periodic mean, it is recommended that the periodic mean should be first removed from the series. In Fig. 2 there are presented periodogram plot and magnitude-squared coherence values. In the right, the first significant
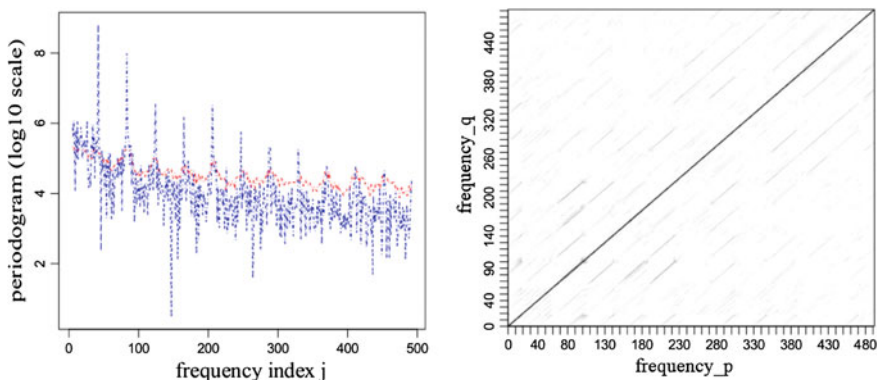

Fig. 2 Periodogram of 'volumes' in the logarithmic scale (*solid line*) with threshold line (*dashed*) for $\alpha = 0.05$ computed from $F(2, 2 * M)$ statistic with $M = 4$. Peaks at $j = 41, 82, \ldots$ (*left-hand side*). Squared coherence statistic, values exceeding an $\alpha = 0.05$ threshold, applied to the 'volumes' series after removing periodic mean with $M = 16$ (*right-hand side*)

off-diagonal is at $|p - q| = 41$ which verifies the first significant peak at $j = 41$ in the periodogram plot on the left. This shows that there is a second-order PC structure with period $T = 494/41 = 24$ in the data.

3.3 Preliminary Identification and Conditioning

Knowing the period length enables one to compute the basic periodic characteristics of the series: periodic mean and periodic standard deviations, which are computed and plotted by procedures `permest()` and `persigest()`, respectively. Both are plotted as a function of season with $1 - \alpha$ confidence intervals based on all non-missing values present for each particular season. For periodic mean confidence intervals are based on the t distribution, whereas for periodic standard deviations the chi-square distribution was used. Figure 3 presents plots of the periodic sample mean and periodic sample deviation, along with 95 % confidence intervals, for the `volumes` data with $T = 24$; it is clearly indicated that the sample mean and periodic sample deviation are not constant and thus the data are non-stationary. The p value for a one-way ANOVA test for equality of seasonal means and p value for Barttlet's test for homogeneity of variance $\sigma(t) \equiv \sigma$ are also computed; by $\sigma(t) \equiv \sigma$ we mean $\sigma(t) = \sigma, t = 0, 1, \ldots, T - 1$. These tests resulted in p value $= 9.74 * 10^{-60}$ for the ANOVA test and p value $= 0$ for Barttlet's test. Rejection of homogeneity based on the p value indicates a properly periodic variance, but leaves open whether or not series is simply the result of a stationary process subjected to amplitude-scale modulation. To resolve this, $R(t + \tau, t)$ for some set of τ values with $\tau \neq 0$ needs to be calculated. If we cannot reject '$R(t, t + \tau)$ is properly periodic' for some $\tau \neq 0$ then the series is an amplitude-modulated stationary sequence.
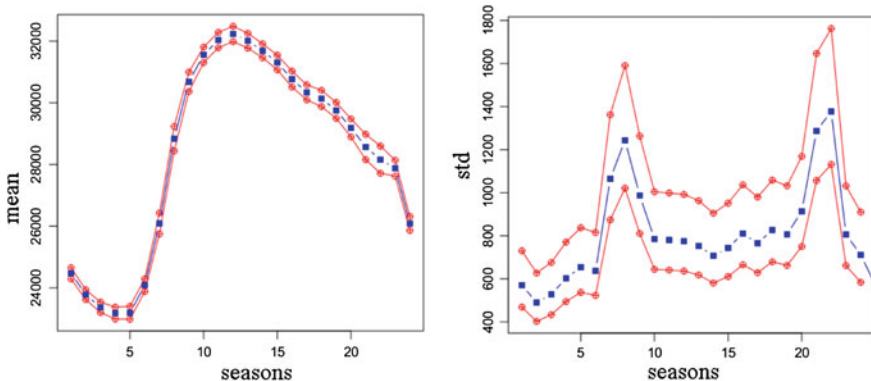
Fig. 3 Estimated periodic mean (*left-hand side*) and periodic standard deviations (*right-hand side*), with $1 - \alpha$ confidence intervals for $\alpha = 0.05$, for 'volumes' series. Number of samples per season = number of periods = 41

3.4 Identification

In this section the determination of p and/or q orders will be carried out. Therefore, periodic version of ACF and PACF functions for `volumes` data will be computed.

3.4.1 Autocovariance Function of PC Processes

At the identification step the covariance function $R(t + \tau, t)$ (see (7)) and/or correlation coefficients $\rho(t + \tau, t)$ for each particular season $t = 0, 1, \ldots, T - 1$ and lags τ should be computed. The essential calculation could be performed in two different ways: direct and indirect.

- Direct method: tests directly on $\rho(t + \tau, t)$
  We first compute correlation coefficients $\rho(t + \tau, t)$ (formula (12)), where $t = 0, 1, \ldots, T - 1$ are seasons and τ is lag (procedure `peracf()`). For each possible pair of t and τ confidence limits for these coefficients are also computed using Fisher transformation, see Fig. 4. Two important tests are also performed:

  (a) $\rho(t + \tau, t) \equiv 0$ meaning $\rho(t + \tau, t) = 0, t = 0, 1, \ldots, T - 1$ the rejection for any $\tau \neq 0$ indicates that the sequence is not PC white noise.
  (b) $\rho(t + \tau, t) \equiv \rho(\tau)$ meaning $\rho(t + \tau, t) = \rho(\tau), t = 0, 1, \ldots, T - 1$, rejection for any $\tau \neq 0$ indicates that series is properly PC and is not just an amplitude-modulated stationary sequence. That is, there exist lags τ for which $\rho(t, \tau)$ is properly periodic in variable t.

An issue is the number of lags that need to be tested. A quick estimate can be obtained from examining the usual ACF applied to the data and using the largest lag producing significant non-zero correlation.
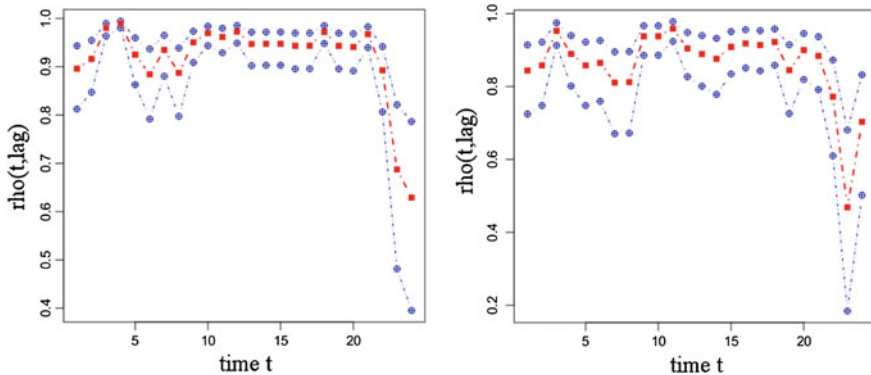
Fig. 4 Correlation coefficients of 'volumes' series: $\hat{\rho}(t, 1)$ versus t and $\hat{\rho}(t, 2)$ versus t with $1 - \alpha$ confidence intervals for $\alpha = 0.05$

Table 1 Constancy of correlation coefficients for 'volumes' data: for each fixed τ we give the p values for tests of constancy in correlations for $\tau = 1, 2, \ldots, 12$

τ	$(a) : \rho_t(\tau) \equiv 0$	$(b) : \rho_t(\tau) = \text{const}$
1	0.00	0.00
2	0.00	4.86623×10^{-9}
3	0.00	2.30767×10^{-4}
4	0.00	0.02262
5	0.00	0.15584
6	0.00	0.41075
7	0.00	0.48365
8	0.00	0.25407
9	0.00	0.038110
10	0.00	0.00770
11	0.00	0.00846
12	0.00	0.00202

Tests are (a) if the coefficients are zero for all t: $\rho_t(\tau) \equiv 0$ and (b) if they are constant: $\rho_t(\tau) \equiv \rho(\tau)$

We note that if $\rho(t + \tau, t) \equiv 0$ is rejected for some $\tau \neq 0$ then also $\rho(t + \tau, t) \equiv \rho(\tau)$ is rejected for that τ. This follows also because if a process is PC white noise, then it is also an amplitude-modulated stationary sequence. Table 1 presents results for both tests when lags $\tau = 1, 2, \ldots, T$ were considered.

Conclusions from direct tests:

1. test (a): hypothesis $\rho_t(\tau) \equiv 0$ for all analyzed lags is rejected. It means that the process exhibits non-zero correlation at non-zero lags, meaning it is not PC neither stationary white noise;
2. test (b): hypothesis $\rho_t(\tau) \equiv \rho(\tau)$ is rejected for lags $\tau = 1, 2, 3$, weakly rejected for $\tau = 10, 11, 12$ and not rejected for other lags. Significant values at lag $\tau = 1, 2, 3$ show that there is periodicity in correlations even when periodic variance is scaled out.

- Indirect method: tests on Fourier coefficients $B_k(\tau)$

  In the second approach computation of the complex estimators $\hat{B}_k(\tau)$ (see (10)) for particular lags τ and $k = 0, 1, \ldots, T - 1$ is performed (procedure `Bcoeff()`). Moreover, p values for the test of $B_k(\tau) = 0$ for all τ and k are returned. Additionally, note that testing if $R_d(t, t) = \sigma_t^2 \equiv \sigma^2$ is equivalent to testing if $B_k(0) = 0$ for $k = 1, \ldots, T - 1$.

  Computations are made for each specified τ, the values of $\hat{B}_k(\tau)$ only for $k = 0, 1, \ldots, \lfloor (T - 1)/2 \rfloor$ because $B_k(\tau) = \overline{B_{T-k}(\tau)}$. In addition, the p values for the test if $B_k(\tau) = 0$ are presented in Table 2. These p values should be treated with caution because the requisite assumptions may not be met (see [16]). In the first two columns we have results for $\tau = 0, 1$ for the original series `volumes`, whereas in the next columns there are results for $\tau = 1, 2$ for the normalized

Table 2 The results of testing the nullity of Fourier coefficients $B_k(\tau)$ and $\rho_k(\tau)$

k	$B_k(0)$	$B_k(1)$	$\rho_k(1)$	$\rho_k(2)$
0	0.00	0.00	0.00	0.00
1	1.48×10^{-1}	8.68×10^{-2}	7.09×10^{-1}	5.36×10^{-1}
2	1.11×10^{-16}	2.30×10^{-12}	8.25×10^{-1}	8.04×10^{-1}
3	8.02×10^{-5}	4.35×10^{-3}	5.51×10^{-1}	4.13×10^{-1}
4	1.43×10^{-3}	3.76×10^{-3}	7.62×10^{-1}	7.43×10^{-1}
5	4.98×10^{-5}	1.16×10^{-3}	8.17×10^{-1}	7.54×10^{-1}
6	5.06×10^{-1}	6.46×10^{-1}	9.43×10^{-1}	8.39×10^{-1}
7	3.46×10^{-3}	4.25×10^{-2}	9.14×10^{-1}	9.23×10^{-1}
8	4.66×10^{-1}	6.54×10^{-1}	8.99×10^{-1}	8.54×10^{-1}
9	3.05×10^{-1}	6.12×10^{-1}	9.57×10^{-1}	9.37×10^{-1}
10	5.75×10^{-1}	7.44×10^{-1}	1.00	9.63×10^{-1}
11	7.54×10^{-1}	9.65×10^{-1}	9.76×10^{-1}	9.49×10^{-1}
12	7.55×10^{-1}	7.49×10^{-1}	9.84×10^{-1}	9.99×10^{-1}

Significant values of $B_k(\tau)$ occurs for $k = 0, 1$ and of $\rho_k(\tau)$ for $k = 0$

series (`volumes` normalized by the sample variance $\hat{\sigma}^2(t)$). If the series is the result of an amplitude-scale modulation of a stationary series, then we expect that $\rho_k(\tau) = 0$ will be rejected only for $k = 0$ and $\tau = 0$ and possibly some other τ and it will never be rejected for any other $k \geq 1$ and lag τ.

Conclusions from indirect tests:

1. Rejection for $\tau = 0$ of hypothesis $B_0(0) = 0$ was expected because $B_0(0)$ is the average variance of the sequence and therefore is non-zero for nontrivial sequences;
2. The strong rejection of that $B_2(\tau) = 0$ for $\tau = 0$ indicates the periodicity in the variance (result is consistent with the `persigest()` function output);
3. Hypothesis that $B_0(1) = 0$ and $B_2(1) = 0$ are strongly rejected, showing very significant periodic variation in correlation. It also indicates that covariance function $R(t + \tau, t)$ is periodic for lags τ with frequency $\lambda = 4\pi/T$;
4. Hypothesis $\rho_k(\tau) = 0$ is strongly rejected for $k = 0$ and $\tau = 1, 2$, which means that there exist high correlation coefficients in the normalized `volumes` series, so it is not white noise;
5. Hypothesis that $\rho_k(\tau) = 0$ is never rejected for $k > 0$ and $\tau = 1, 2$, so using this method we cannot reject the hypothesis that the correlation coefficients for lag τ are constant (again consistent with normalized series being stationary, not necessarily white noise).

Note that indirect method of covariance function estimation provides the opposite conclusion in comparison to the direct method. In this case we cannot reject the possibility that analyzed series is a result of amplitude modulation of stationary series (the hypothesis about correlation coefficients equal to zero was not rejected). Surprisingly, it can happen because direct method examines the sample time-dependent

correlations in the period, whereas indirect method returns estimators for Fourier coefficients. It is argued that usually we can reject $\rho(t+1, t) \equiv \rho_1$ better in time domain than in the frequency domain [16, pp. 228–292].

3.4.2 Partial Autocovariance of PC-T Processes

Computation of periodic sample correlation coefficients $\hat{\pi}_{n+1}(t) = \hat{\pi}(n+1, t)$ (see (15)) is provided by the perpacf() function, see Fig. 5. Also ppfcoeffab() procedure is applied to represent $\pi(n, t)$ by its Fourier coefficients. If the variation in time of $\pi(n, t)$ is sufficiently smooth over the period T, then looking at these Fourier coefficients may be a more sensitive detector of linear dependence of x_{t+1} on the preceding n samples (n is fixed here) than looking at $\pi(n, t)$ for individual times. Two additional hypothesis tests are also provided:
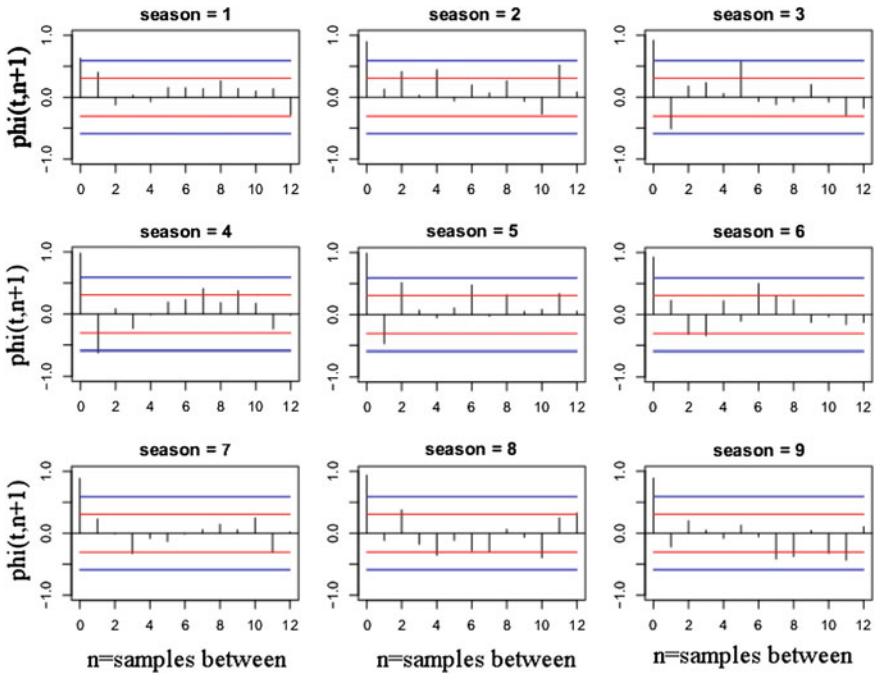
Fig. 5 Partial correlation coefficients $\hat{\pi}(n+1, t)$ of the 'volumes' series for specified number of samples $n = 1, \ldots, 12$ and seasons $t = 1, \ldots, 9$ with $1 - \alpha$ confidence intervals, $\alpha = 0.05$; inner lines are confidence intervals for the null hypothesis of $\pi(n+1; t) = 0$ and are based on the asymptotic null distribution; the outer lines are based on the same test but with Bonferroni correction for the number of n; coefficients values are presented on separate plots for particular $t = 1, \ldots, T$

(a) if $\pi(n_0 + 1, t) = 0$ for $t = 0, 1, \ldots, T - 1$ and each fixed n_0 for $n_0 = 1, 2, \ldots, nmax$,

(b) if $\pi(n + 1, t) = 0$ for $t = 0, 1, \ldots, T - 1$ and $n_0 \leq n \leq nmax$.

For `volumes` data only for $n = 0, 2$ null hypothesis are rejected. Thus, only $\pi(1, t) \neq 0$ and $\pi(3, t) \neq 0$ for some values of t. For $n = 1$ and $n > 2$ and all t we can assume that $\pi(n + 1, t) = 0$. The obtained results suggest fitting a PAR(3) model to the data.

3.4.3 Usual ACF and PACF

Finally, to check the strength of dependence between the variables usual (meaning not periodic) ACF and PACF functions are plotted. It happens that for PC time series the usual ACF and PACF are still useful in the identification of model orders and in some cases can be more sensitive than the periodic versions. When subjected to a truly PC sequence, the usual ACF and PACF produce an average (of sorts; the exact relationship is an open question) of the instantaneous (time indexed) values produced by periodic ACF and periodic PACF. Depending therefore on correlations between these averaged quantities, the usual ACF and PACF may be more or less sensitive than the instantaneous ones. Function `acfpacf()` plots values of usual ACF and PACF functions with confidence intervals. It is possible to run this procedure on original data which include the periodic mean as a deterministic component. But typically the periodic mean can distort our understanding (or view) of the random fluctuations, thus using data after removing periodic mean is recommended as well (see Fig. 6). As a result of identification stage the orders of lags p and/or q for model of PARMA type: PARMA(p,q), PAR(p) or PMA(q), should be determined. For `volumes` data, periodic ACF values point to significant periodic correlations to many lags, which is consistent with PAR model with large coefficients. One can observe the significant periodic PACF values at $n = 0$ and maybe $n = 1$, which suggests starting with

Fig. 6 Usual ACF and PACF values of 'volumes' series together with 95 % confidence intervals (*top* and *bottom* plot, respectively); think inner dotted line is 95 % CI for null hypothesis of $\hat{\pi}(n + 1) = 0$ and outer is same but Bonferonni corrected for the number of n plotted (i.e. 24)
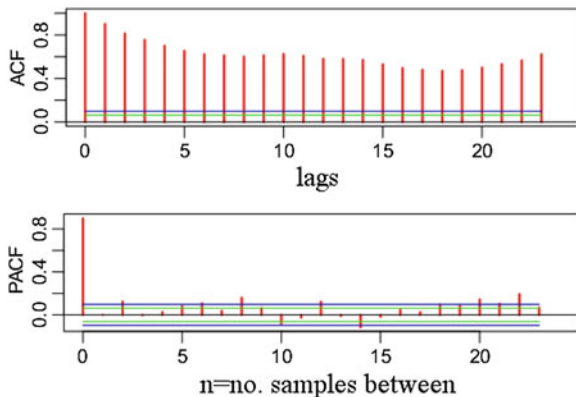
PAR(1) model and increasing the order of lags if it will be necessary. Significant values of the usual ACF and PACF are consistent with the periodic ones (strong autocorrelation values on ACF plot for all specified lags and significant values of PACF for some n, i.e. $n = 0, 2$, mainly for $n = 0$, p value equal to 0). This indicates a strong average lag 1 correlation and a weaker lag 3 ($lag = n + 1$) correlation suggesting a PAR(3) with a high average $\phi_1(t)$ and a low average $\phi_3(t)$.

3.5 Model Fitting

The PARMA(p,q) model has in total $T(p + q + 1)$ coefficients, thus, especially for long period length, the problem of estimation is computationally burdensome. Therefore, at least at the beginning, to reduce the number of parameters, it is recommended to fit models which explain that data with the fewest possible parameters. Then, only if proposed model is not sufficient, the order of lags should be increased. In the `perARMA` package, the Yule–Walker moment estimates for the general PAR model (procedure `perYW()`) and the approximate maximum likelihood parameter estimates for the general PARMA model (procedure `parmaf()`) are implemented. To illustrate the functionality of these procedures we apply them to the `volumes` data, after removing periodic mean, elaborating identification clues. First, we fit the simplest possible model, i.e. PAR(1)

$$X_t - \phi(t)X_{t-1} = \sigma(t)\xi_t,$$

where ξ_t is white noise with zero mean and variance equal to 1. The total number of parameters in this model is equal to $2T$. As it is shown in Figs. 7 and 8 the
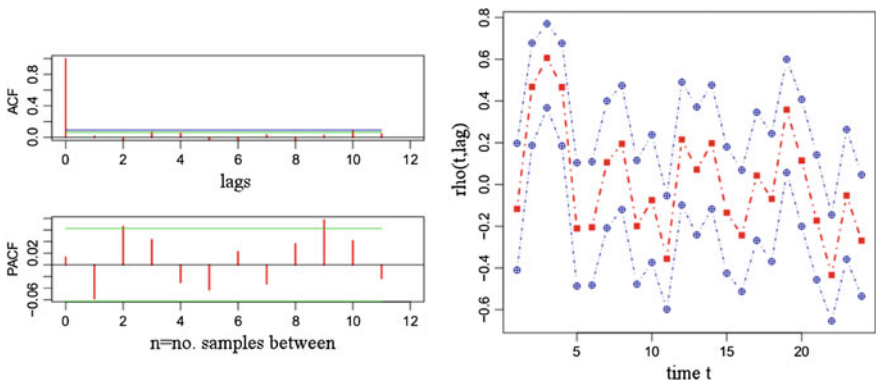
Fig. 7 PAR(1) model residuals evaluation: Usual ACF and PACF values together with 95 % confidence intervals (*top* and *bottom* plot, respectively); think inner dotted line is 95 % CI for null hypothesis of $\hat{\pi}(n + 1) = 0$ and outer is same but Bonferonni corrected for the number of n plotted (i.e. 12) (*left-hand side*). Correlation coefficients $\hat{\rho}(t, 1)$ versus t with confidence intervals for $\alpha = 0.05$ (*right-hand side*)
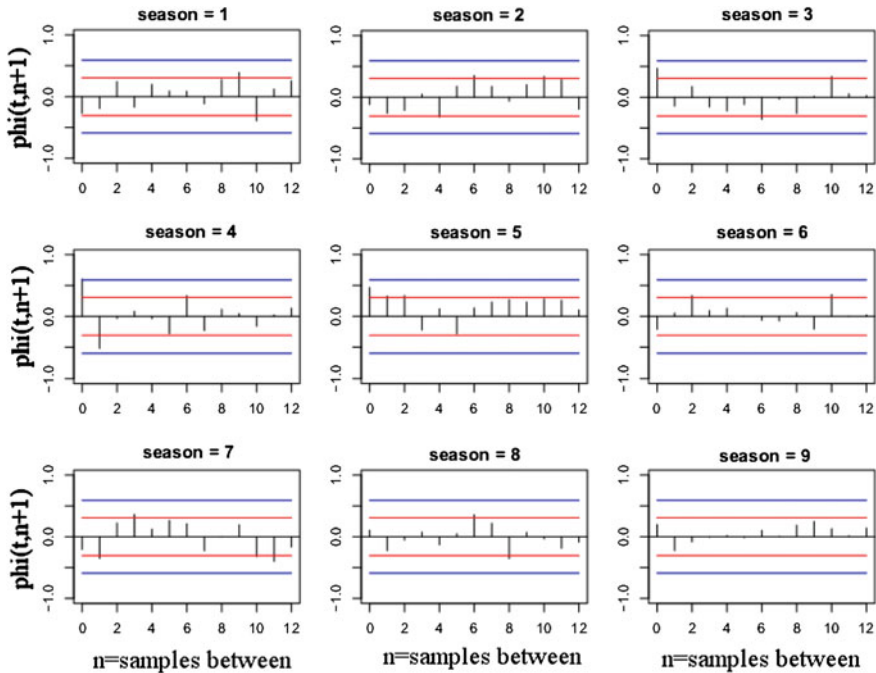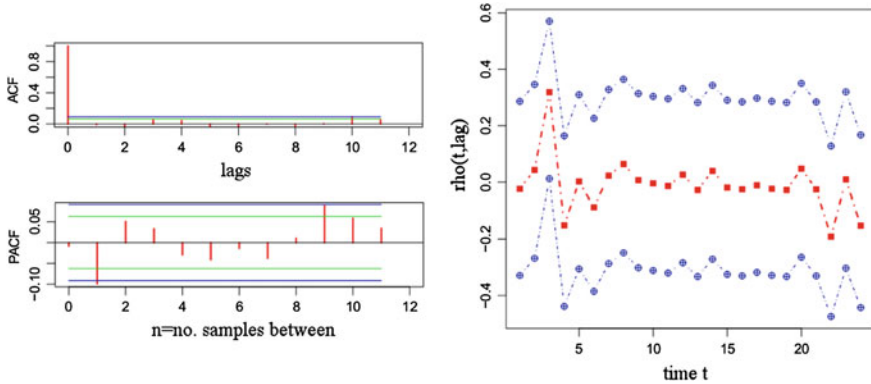
Fig. 8 Partial correlation coefficients $\hat{\pi}(n+1, t)$ of PAR(1) model residuals for specified number of samples $n = 1, \ldots, 12$ and seasons $t = 1, \ldots, 9$ with $1 - \alpha$ confidence intervals, $\alpha = 0.05$; inner lines are confidence intervals for the null hypothesis of $\pi(n+1, t) = 0$ and are based on the asymptotic null distribution; the outer lines are based on the same test but with Bonferroni correction for the number of n; coefficients values are presented on separate plots for particular $t = 1, \ldots, 12$

residuals are not consistent with periodic white noise ones. Even if correlations significantly decreased, for $\tau = 1$ test for equality still rejects the null hypothesis with p value equal to $1.54E - 7$. Then we fit another model with the increased order of autoregression p, i.e. PAR(2)

$$X_t - \hat{\phi}_1(t)X_{t-1} - \hat{\phi}_2(t)X_{t-2}(t) = \hat{\sigma}(t)\xi_t,$$

where ξ_t is white noise with zero mean and variance equal to 1. The total number of parameters is equal to 3T. We repeat the whole procedure of examining residuals and checking if residuals are consistent with periodic white noise. In Fig. 9 one can observe that significant periodic correlations in the residuals are completely absent in residuals of PAR(2) model. For lag $\tau = 1$ both null hypothesis (provided by `peracf()` procedure) were not rejected with p values equal to 0.99 and 0.81. Also usual ACF and PACF functions for residuals are consistent with periodic white noise ones. In Fig. 10 there are presented coefficients $\hat{\pi}(n+1, t)$ for PAR(2) residuals ($n = 1, \ldots, 8$ and $t = 1, \ldots, 12$) together with confidence intervals for $\alpha = 0.05$: inner for $\pi(n+1, t) = 0$ and outer for Bonferroni correction. There are finally

Fig. 9 PAR(2) model residuals evaluation: Usual ACF and PACF values together with 95 % confidence intervals (*top* and *bottom* plot, respectively); think inner dotted line is 95 % CI for null hypothesis of $\hat{\pi}(n+1) = 0$ and outer is same but Bonferonni corrected for the number of n plotted (i.e. 12) (*left-hand side*). Correlation coefficients $\hat{\rho}(t, 1)$ versus t with confidence intervals for $\alpha = 0.05$ (*right-hand side*)
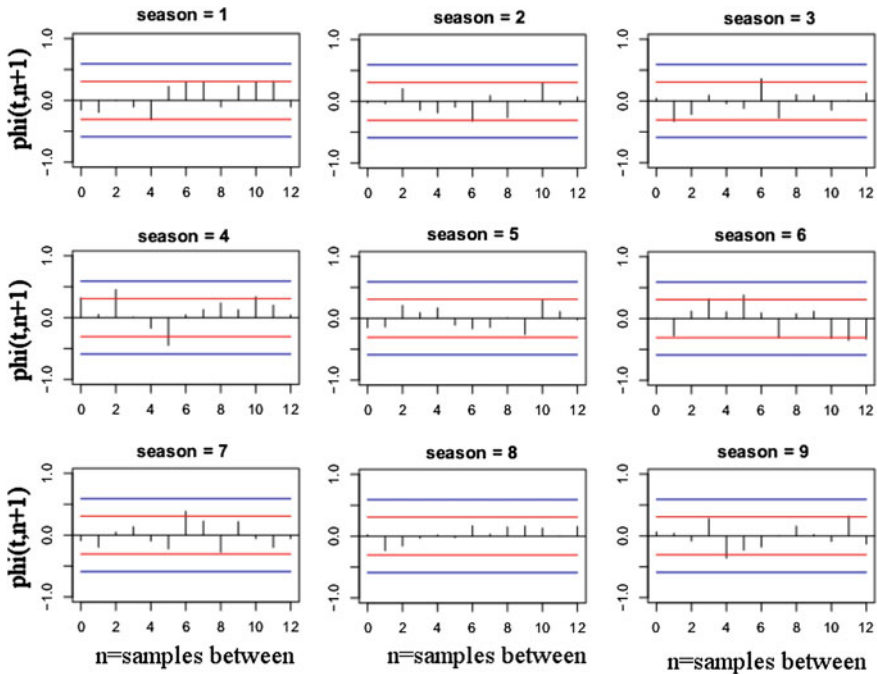

Fig. 10 Partial correlation coefficients $\hat{\pi}(n+1, t)$ of PAR(2) model residuals for specified number of samples $n = 1, \ldots, 12$ and seasons $t = 1, \ldots, 9$ with $1 - \alpha$ confidence intervals, $\alpha = 0.05$; inner lines are confidence intervals for the null hypothesis of $\pi(n+1, t) = 0$ and are based on the asymptotic null distribution; the outer lines are based on the same test but with Bonferroni correction for the number of n; coefficients values are presented on separate plots for particular $t = 1, \ldots, 12$)

no values which exceed the outer threshold. For periodic white noise we expect $\pi(n + 1, t) = 0$, for all t and n and in this case this hypothesis is definitely not rejected (p value $= 0.27$).

Additionally, the penalized likelihoods for both PAR models were computed using `loglikec()` procedure. AIC values were equal to 6873.231 and 6821.9 for PAR(1) and PAR(2) models, respectively. BIC values were equal to 6883.012 and 6836.569 for PAR(1) and PAR(2) models, respectively. Both criteria prefers PAR(2) model and analysis of residuals confirms PAR(2) model as a better fit. Furthermore, as residuals of PAR(2) model are consistent with periodic white noise, it seems that `volumes` data did not require a full PARMA model.

3.6 Short Time Prediction for PAR Models

In this section a diagnostic checking is carried out by comparing predicted values of `volumes` series (based on PAR models coefficients) with the corresponding real values of this series on 1 and 2 September 2010. First procedure `predictperYW()`,
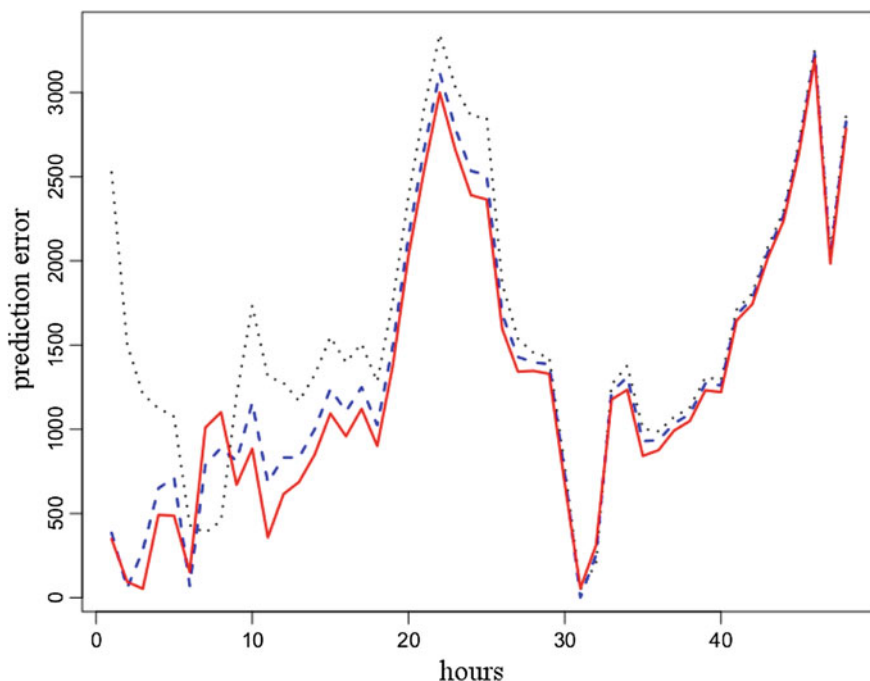
Fig. 11 Prediction errors (i.e. difference between predicted and true values of the series) on 1 and 2 September 2010 for various ways of predicting: PAR(1) (*dashed line*), PAR(2) (*solid line*) and sample periodic mean (*dotted line*)

based on the Yule–Walker estimators, is applied to demeaned `volumes` series to forecast values for the next 2 days (48 observations). Then sample periodic mean is added to obtained predictors and results are compared with the real (observed) values of the series. Instead of analyzing particular observations values the prediction error, i.e. observed values minus predicted values is considered. In Fig. 11 there are presented prediction errors for both PAR(1) and PAR(2) models and just sample periodic mean used as a predictor as well. Taking these results into account, it seems that the forecast based on PAR(2) model approximates the real values the best.

4 Simulated Data Example

In the previous section PAR(2) model was the optimal choice and Yule–Walker estimators were sufficient but not always this needs to be the case. Therefore, in this section, to complete our considerations about PARMA modelling, we illustrate full PARMA model estimation method, based on maximization of log-likelihood function. We use simulated dataset to test the performance of the estimation procedures for the PARMA(p,q) model.

We decided to consider the simulated datasets to check the performance of our procedures for full PARMA model estimation. In the example below, a PARMA(2,1) sequence is generated with period length $T = 12$ and series length equal to $N = 480$. Knowing the orders of the original PARMA sequence we can compare them with obtained output from estimation procedures presented in the next section.

For general PARMA we use non-linear optimization methods to obtain maximum logarithm of likelihood function. In `perARMA` package procedure `parmaf()` enables to estimate the values of the PARMA(p,q) model. This method of computation of log-likelihood function is based on the representation of the periodically varying parameters by Fourier series (see (17)). Inside the procedure the negative logarithm of likelihood function from the PARMA sequence for matrices of coefficients in their Fourier representation is computed. This alternative parametrization of PARMA system, introduced by [17], can sometimes substantially reduce the number of parameters required to represent a PARMA system. This permits estimation of the values of the chosen representation of PARMA(p,q) model using non-linear optimization methods. Initial values of parameters are computed using Yule–Walker equations.

To illustrate functionality of these procedure we try to fit PARMA-type model to the series generated as PARMA(2,1). We consider three PARMA-type models: PARMA(0,1), PARMA(2,0) and PARMA(2,1) and compare the result with reality. According to log-likelihood, AIC and BIC values presented in Table 3, PARMA(2,1) is considered to be the best fitted. It is consistent with our expectations as analyzed series y was generated using orders $p = 2$, $q = 1$.

Table 3 The information criteria for validation of PARMA models fits when true model was PARMA(2,1)

model	negloglik	AIC	BIC
PARMA(0,1)	969.05	977.05	993.74
PARMA(2,0)	748.98	762.98	792.17
PARMA(2,1)	653.38	673.38	715.07

5 Conclusions

In the present study we showed that PARMA model approach works for real data with periodic structure. We follow all steps of standard model fitting procedure in regard to PARMA models through non-parametric spectral analysis, model identification, parameter estimation, diagnostic checking (model verification). Additionally, we perform test for period length detection, prediction for PAR models and estimation of full PARMA for simulated data. The results presented here illustrate how PARMA approach can be applied to model periodic structure of energy market data and confirm the considerable value of this method in forecasting short-term energy demand.

To our knowledge `perARMA` package is the only existing tool in statistical software programs that provide PARMA time series analysis and allow to follow all steps of standard model fitting procedure. Additionally, it deals with missing data, provides period length detection tools and prediction for PAR models. Moreover, `perARMA` package permits simulation of PARMA time series.

References

1. Anderson, P. L., Meerschaert, M. M., & Zhang, K. (2013). Forecasting with prediction intervals for periodic autoregressive moving average models. *Journal of Time Series Analysis, 34*(2), 187–193.
2. Anderson, P. L., Tesfaye, Y. G., & Meerschaert, M. M. (2007). Fourier-PARMA models and their application to river flows. *Journal of Hydrologic Engineering, 12*(5), 462–472.
3. Balcilar, M., & McLeod, A. I. (2011). pear: Package for periodic autoregression analysis. http://cran.r-project.org/web/packages/pear/.
4. Bloomfield, P., Hurd, H. L., & Lund, R. (1994). Periodic correlation in stratospheric ozone data. *Journal of Time Series Analysis, 15*, 127–150.
5. Bloomfield, P., Hurd, H. L., Lund, R. B., & Smith, R. (1995). Climatological time series with periodic correlation. *Journal of Climate, 8*, 2787–2809.
6. Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting*. New York: Springer.
7. Broszkiewicz-Suwaj, E., Makagon, A., Weron, R., & Wyłomańska, A. (2004). On detecting and modeling periodic correlation in financial data. *Physica A, 336*, 196–205.
8. Cramér, H. (1961). *Methods of mathematical statistics*. New York: Princeton University Press.

9. Dehay, D., & Hurd, H. L. (1994). Representation and estimation for periodically and almost periodically correlated random processes. In W. A. Gardner (Ed.), *Cyclostationarity in communications and signal processing*. New York: IEEE Press.

10. Dudek, A. E., Hurd, H., & Wójtowicz, W. perARMA: Package for periodic time series analysis. R package version 1.5, http://cran.r-project.org/web/packages/perARMA/.

11. Franses, P. H. (1996). *Periodicity and stochastic trends in economic time series*. Oxford: Oxford Press.

12. Gardner, W. A. (1994). *Cyclostationarity in communications and and signal processing*. New York: IEEE Press.

13. Gardner, W. A., Napolitano, A., & Paura, L. (2006). Cyclostationarity: Half a century of research. *Signal Processing, 86*, 639–697.

14. Gerr, N. L., & Hurd, H. L. (1991). Graphical methods for determining the presence of periodic correlation in time series. *Journal of Time Series Analysis, 12*, 337–350.

15. Gladyshev, E. G. (1961). Periodically correlated random sequences. *Soviet Mathematics, 2*, 385–388.

16. Hurd, H. L., & Miamee, A. G. (2007). *Periodically correlated random sequences: Spectral theory and practice*. Hoboken: Wiley.

17. Jones, R., & Brelsford, W. (1967). Time series with periodic structure. *Biometrika, 54*, 403–408.

18. López-de-Lacalle, J. (2012). partsm: Periodic autoregressive time series models. R package version 1.1, http://cran.r-project.org/web/packages/partsm/.

19. Pagano, M. (1978). On periodic and multiple autoregressions. *Annals of Statistics, 6*(6), 1310–1317.

20. Parzen, E., & Pagano, M. (1979). An approach to modelling seasonally stationary time series. *Journal of Econometrics, 9*(1–2), 137–153.

21. Rasmussen, P. F., Salas, J., Fagherazzi, L., Rassam, J. C., & Bobée, B. (1996). Estimation and validation of contemporaneous PARMA models for streamflow simulation. *Water Resources Research, 32*(10), 3151–3160.

22. Sabri, K., Badaoui, M., Guillet, F., Belli, A., Millet, G., & Morin, J. B. (2010). Cyclostationary modelling of ground reaction force signals. *Signal Processing*, *90*, 1146–1152.

23. Sakai, H. (1991). On the spectral density matrix of a periodic ARMA process. *Journal of Time Series Analysis, 12*(2), 73–82.

24. Smadi, A. A. (2009). Periodic auto-regression modeling of the temperature data of Jordan. *Pakistan Journal of Statistics, 25*(3), 323–332.

25. Swider, D. J., & Weber, C. (2007). Extended ARMA models for estimating price developments on day-ahead electricity markets. *Electric Power Research, 77*, 583–593.

26. Tesfaye, Y., Meerschaert, M. M., & Anderson, P. L. (2006). Identification of periodic autoregressive moving average models and their application to the modeling of river flows. *Water Resources Research, 42*, W01419.

27. Vecchia, A. V. (1985). Maximum likelihood estimation for periodic autoregressive moving average models. *Technometrics, 27*, 375–384.

28. Vecchia, A. V. (1985). Periodic autoregressive moving average (PARMA) modeling with applications to water resources. *Water Resources Bulletin, 21*, 721–730.