# Modeling periodic autoregressive time series with multiple periodic effects

Harry Hurd and Vladas Pipiras

**Abstract** Two models of periodic autoregressive time series with multiple periodic effects are introduced and studied. In the first model, the autoregression coefficients vary periodically with several dominant components associated with two or more periods (for example, day and week for hourly data). In the second model, the autoregression coefficients consist of the additive periodic effects of several nominal variables (for example, the effect of hour in a given day and the effect of day in a given week for hourly data). Truncated Fourier representations of different periods are used to parametrize the autoregression coefficients in the two models. Model estimation and inference through ordinary and weighted least squares, and model selection based on diagnostics plots, in particular, are considered for the two approaches. An application to a real time series of hourly electricity volumes from the Nord Pool Spot Exchange is also presented, where the nature and use of the two models are contrasted.

## 1 Introduction

Many data collected in time exhibit cyclical variations, and call for time series models with cyclical features. One class of such models consists of time series with periodically varying dependence structures. The periodicity could be in the mean, the variance, but also in the model parameters such as with periodic autoregressive (PAR) models that play a central role in this class of models. See Ghysels and Osborne [13], Franses and Paap [12], Hurd and Miamee [16].

Harry Hurd
Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260, e-mail: hhurd1@nc.rr.com

Vladas Pipiras
Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260, e-mail: pipiras@email.unc.edu

In this work, we are interested in periodically correlated time series and, more specifically, PAR series where periodicity is driven by two or more periods. Having cyclical variations at multiple periods is expected in many data, especially when they are associated with natural cycles of 24 hours, 1 week (when modeling human related activity), 4 annual quarters or seasons, and so on. We shall introduce two classes of periodically non-stationary time series that will operate at two or more periods.

To motivate briefly the construction of the models and to explain the basic ideas, suppose the goal is to model just the deterministic mean function $\mu(t)$ of the series as a function with two periodic effects. As with the application considered below in this work, suppose time $t$ refers to hours and the two periodic effects are associated with the 24 hours (1 day) and 168 hours (1 week) periods. Two natural candidates for $\mu(t)$ operating at these two different periods come to mind, namely,

$$\mu(t) = \mu_{24}(t) + \mu_{168}(t), \tag{1}$$

where, for example, in the first case,

$$\mu_{24}(t) = 2 + 0.5\cos(\frac{2\pi t}{24}), \quad \mu_{168}(t) = -0.1\sin(\frac{2\pi t}{168}), \tag{2}$$

and, in the second case,

$$\begin{aligned} \mu_{24}(t) &= 1 - (0.2)1_{1\text{AM}}(t) + (0.3)1_{2\text{AM}}(t) + (0.7)1_{7\text{PM}}(t), \\ \mu_{168}(t) &= (0.3)1_{\text{Monday}}(t) - (0.1)1_{\text{Wednesday}}(t) + (4)1_{\text{Sunday}}(t), \end{aligned} \tag{3}$$

where $1_E(t)$ stands for the indicator function of "event" $E$, that is, it is equal to 1 if $t$ falls into $E$, and 0 otherwise. The mean function $\mu(t)$ in (1) and (2) consists of two dominant components, one with period 24 and the other with period 168. The mean function $\mu(t)$ in (1) and (3), on the other hand, expresses the idea that the mean effect can be due to the hour of a given day or the day of a given week.

Our models for PAR time series with multiple periodic effects will allow for such periodic behavior for all model parameters, not just the mean function. The model extending (2) will be referred to as the model of Type A, and that extending (3) as the model of Type B. As with (2), we shall use Fourier representations of periodic model coefficients that will often require estimating fewer coefficients.

A number of other authors also considered various models exhibiting cyclical variations at several periods. For example, Gould et al. [14], De Livera et al. [6] and others consider models involving multiple periods based on exponential smoothing. The use of double seasonal ARIMA models (that is, seasonal ARIMA models with two periods) goes back at least to Box et al. [4]. Basawa et al. [3] do not quite have multiple periods but consider a hybrid model exhibiting both seasonal and periodic dependence for the same period. Neural networks in the context of multiple periods were used by Dudek [10, 11] and others. Comparison of various available methods involving multiple periods can be found in Taylor et al. [20]. Applications

to electricity markets dominate many of these contributions; see also Weron [22], Dannecker [5].

Our data application is also related to electricity markets. But we do not seek to provide an exhaustive comparison of our approach to other methods. The goal is to explain how one could think of periodic autoregressive time series with multiple periods at a most basic level, and how the resulting models could be estimated and manipulated in other ways. Though we also note that the introduced models do seem relevant for the considered data set.

The structure of the paper is as follows. The models of Types A and B are defined in Section 2 below. Estimation issues are discussed in Section 3, and inference, model selection and other issues in Section 4. A data application is considered in Section 5. Conclusions can be found in Section 6.

## 2 PAR models with multiple periodic effects

For the sake of clarity, we focus on PAR models with two periodic effects and comment on the case of multiple periodic effects in Remarks 2 and 3 below. The two periodic effects will be associated with two periods that are denoted $s_1, s_2$. We shall suppose that $s_1 < s_2$ and $s_2/s_1$ is an integer. For example, in the application in Section 5 below, $s_1 = 24$ hours (1 day) and $s_2 = 24 \cdot 7 = 168$ hours (1 week).

### *2.1 Model A*

To introduce our first model with two periodic effects, we need several preliminary observations and definitions. A function $f(t)$ is $s$-periodic if $f(t+s) = f(t)$ for all $t \in \mathbb{Z}$. Note that an $s_1$-periodic function is also $s_2$-periodic (with the assumptions on $s_1, s_2$ stated above). An $s_2$-periodic function $f(t)$ can always be expressed through a Fourier representation as

$$f(t) = f_0 + \sum_{m=1}^{\lfloor s_2/2 \rfloor} \Big( f_{1,m} \cos(\frac{2\pi mt}{s_2}) + f_{2,m} \sin(\frac{2\pi mt}{s_2}) \Big), \tag{4}$$

where $f_0, f_{1,m}, f_{2,m} \in \mathbb{R}$. It can then also be expressed (uniquely) as

$$f(t) = f_0 + f_1(t) + f_2(t), \tag{5}$$

where

$$f_1(t) = \sum_{m_1=1}^{\lfloor s_1/2 \rfloor} \Big( f_{1,(s_2/s_1)m_1} \cos(\frac{2\pi(s_2/s_1)m_1 t}{s_2}) + f_{2,(s_2/s_1)m_1} \sin(\frac{2\pi(s_2/s_1)m_1 t}{s_2}) \Big)$$
$$= \sum_{m_1=1}^{\lfloor s_1/2 \rfloor} \Big( f_{1,(s_2/s_1)m_1} \cos(\frac{2\pi m_1 t}{s_1}) + f_{2,(s_2/s_1)m_1} \sin(\frac{2\pi m_1 t}{s_1}) \Big) \tag{6}$$

and

$$f_2(t) = \sum_{m=1,\ldots,\lfloor s_2/2 \rfloor; m/s_1 \notin \mathbb{Z}} \Big( f_{1,m} \cos(\frac{2\pi m t}{s_2}) + f_{2,m} \sin(\frac{2\pi m t}{s_2}) \Big). \tag{7}$$

We shall refer to $f_j(t)$ as the $s_j$-periodic component of $f(t)$, $j = 1,2$.

The following definition concerns our first model with two periodic effects.

**Definition 1.** A time series $\{X_t\}_{t\in\mathbb{Z}}$ is *type A periodic autoregressive of order p* (A–PAR($p$)) *with two periodic effects* if

$$X_t = \mu(t) + Y_t, \tag{8}$$
$$Y_t = \phi_1(t) Y_{t-1} + \ldots + \phi_p(t) Y_{t-p} + \sigma(t)\varepsilon_t \tag{9}$$

with $\{\varepsilon_t\}_{t\in\mathbb{Z}} \sim \mathrm{WN}(0,1)$ (that is, a white noise series with $\mathbb{E}\varepsilon_t = 0$ and $\mathbb{E}\varepsilon_t^2 = 1$) and $s_2$-periodic $\mu(t)$, $\sigma(t)^2$ and $\phi_r(t)$, $r = 1,\ldots,p$, with the decompositions

$$\begin{array}{rl} \mu(t) &= \mu_0 + \mu_1(t) + \mu_2(t), \\ \sigma(t)^2 &= \sigma_0^2 + \sigma_1^{(2)}(t) + \sigma_2^{(2)}(t), \\ \phi_r(t) &= \phi_{r,0} + \phi_{r,1}(t) + \phi_{r,2}(t), \quad r = 1,\ldots,p, \end{array} \tag{10}$$

as in (5), where at least one of the $s_1$-periodic components $\mu_1(t), \sigma_1^{(2)}(t), \phi_{r,1}(t)$, $r = 1,\ldots,p$, is non-zero.

In practice, motivated by the representations (5)–(7), we shall model the coefficients $\phi_r(t)$ and their components $\phi_{r,1}(t)$ and $\phi_{r,2}(t)$ as

$$\phi_{r,j}(t) = \sum_{m_j=1}^{H_j} \Big( a_{r,m_j}^{(j)} \cos(\frac{2\pi m_j t}{s_j}) + b_{r,m_j}^{(j)} \sin(\frac{2\pi m_j t}{s_j}) \Big), \quad j = 1,2, \tag{11}$$

assuming $H_2 < s_2/s_1$ (which ensures that indices $m_2$ in (11) are not multiples of $s_2/s_1$). The indices $j = 1$ and $j = 2$ in (11) correspond to $s_1$-periodic and $s_2$-periodic components, respectively. Modeling periodic time series through the (reduced) Fourier representations of their coefficients goes back at least to Jones and Brelsford [17]. See also Dudek et al. [8] and references therein.

The parameters $\mu_0$, $\mu_1(t)$, $\mu_2(t)$, $\sigma_0^2$, $\sigma_1^{(2)}(t)$, $\sigma_2^{(2)}(t)$, on the other hand, will be estimated in a nonparametric fashion, though a parametric route analogous to (11) is also a possibility. Note also that $\sigma_1^{(2)}(t)$, $\sigma_2^{(2)}(t)$ are not necessarily positive.

*Remark 1.* By the discussion above, the series $\{X_t\}_{t\in\mathbb{Z}}$ in Definition 1 is also PAR($p$) with the larger period $s_2$. We also note that our main interest here is in

such series $\{X_t\}_{t\in\mathbb{Z}}$ which are stable, that is, for which the multivariate VAR representation of the $s_2$-vector series $\{(X_{s_2(\tilde{t}-1)+1}, X_{s_2(\tilde{t}-1)+2}, \ldots, X_{s_2\tilde{t}})'\}_{\tilde{t}\in\mathbb{Z}}$ is stable. Here and throughout, a prime indicates a vector or matrix transpose. Conditions for the latter are well-known in the literature; see, for example, Lütkepohl [19].

*Remark 2.* The framework described above can be extended straightforwardly to the case of multiple periods $s_1, s_2, \ldots, s_K$, assuming that $s_1 < s_2 < \ldots < s_K$ and $s_K/s_j$ are integers. Though some caution would need to be exercised in how many terms in the Fourier representations are included when some multiples of two periods $s_{j_1}$ and $s_{j_2}$ are the same (and smaller than $s_K$).

## 2.2 Model B

We now turn to a different PAR model that builds on the idea behind the model (1) and (3) for the mean discussed in Section 1. We adopt the following quite general framework concerning two periodic effects.

We think of each time $t$ and observation $X_t$ as associated with two nominal variables, that vary periodically in time, and are interested to model their effects. We assume that the two variables have $k_1$ and $k_2$ levels, respectively. We shall represent the two nominal variables by two functions $g_1(t)$ and $g_2(t)$, assuming that they are $s_1$-periodic and $s_2$-periodic, respectively, and take values $\{1,\ldots,k_1\}$ and $\{1,\ldots,k_2\}$, respectively, that are associated with respective levels. As above, we assume that $s_1 < s_2$ and $s_2/s_1$ is an integer. It is not necessarily the case that $s_j = k_j$, as the following examples illustrate.

*Example 1.* In the application to hourly data in Section 5 below, the two periodic effects will be the effect of the hour in a day and the effect of the day in a week. For hourly data, these effects are periodic with periods $s_1 = 24$ hours (1 day) and $s_2 = 24\cdot 7 = 168$ hours (1 week), respectively. The corresponding nominal variables have $k_1 = 24$ (hours 1 through 24) and $k_2 = 7$ (Monday through Sunday) levels, respectively. The effects can be captured through the two corresponding functions $g_1(t)$ and $g_2(t)$ with the properties described above. They can also be represented as

$$g_1(t) = t, \quad t = 1,\ldots,24, \quad g_2(t) = \lceil \frac{t}{24} \rceil, \quad t = 1,\ldots,168, \tag{12}$$

where $\lceil x \rceil$ denotes the ceiling integer part of $x$, and then extended periodically with their respective periods.

*Example 2.* One could have the second variable (function) in Example 1 having only $k_2 = 2$ levels (values), for workdays and weekends. Similarly, the first variable (function) in Example 1 could have $k_2 = 4$ levels (values), for night hours (1-6AM), morning hours (6AM-12PM), afternoon hours (12-6PM) and evening hours (6PM-12AM).

**Definition 2.** A time series $\{X_t\}_{t\in\mathbb{Z}}$ is *type B periodic autoregressive of order p* (B–PAR($p$)) *with two periodic effects* if

$$X_t = \mu(t) + Y_t, \tag{13}$$
$$Y_t = \phi_1(t)Y_{t-1} + \ldots + \phi_p(t)Y_{t-p} + \sigma(t)\varepsilon_t \tag{14}$$

with $\{\varepsilon_t\}_{t\in\mathbb{Z}} \sim \mathrm{WN}(0,1)$ and

$$\begin{aligned} \mu(t) &= \mu_0 + \mu_1(g_1(t)) + \mu_2(g_2(t)), \\ \sigma(t)^2 &= \sigma_0^2 + \sigma_1^{(2)}(g_1(t)) + \sigma_2^{(2)}(g_2(t)), \\ \phi_r(t) &= \phi_{r,0} + \phi_{r,1}(g_1(t)) + \phi_{r,2}(g_2(t)), \quad r = 1,\ldots,p, \end{aligned} \tag{15}$$

where the functions $g_1(t)$ and $g_2(t)$ are defined before Example 1, are associated with two nominal variables and are $s_1$-periodic and $s_2$-periodic, respectively.

Definition 2 requires further clarification. With $f(t)$ denoting $\mu(t)$, $\sigma(t)^2$ or $\phi_r(t)$, let

$$f(t) = f_0 + f_1(g_1(t)) + f_2(g_2(t)) \tag{16}$$

be the decomposition analogous to those in (15). Recall from above that $g_j(t)$ takes an integer value from 1 to $k_j$, which we shall denote by $u_j$. Thus, $f_j$ acts on a value $u_j$ as $f_j(u_j)$, where $u_j = g_j(t)$. For identifiability purposes, we assume that

$$\sum_{u_j=1}^{k_j} f_j(u_j) = 0, \quad j = 1,2. \tag{17}$$

We also note that the function $f_j(g_j(t))$ is $s_j$-periodic, $j = 1,2$, and hence, with our assumptions on $s_1, s_2$, the function $f(t)$ is $s_2$-periodic with the larger $s_2$.

The function $f_j(u_j)$, $j = 1,2$, $u_j = 1,\ldots,k_j$, can be expressed through a Fourier representation as

$$f_j(u_j) = \sum_{m_j=1}^{\lfloor k_j/2 \rfloor} \Big( f_{1,m_j}^{(j)} \cos(\frac{2\pi m_j u_j}{k_j}) + f_{2,m_j}^{(j)} \sin(\frac{2\pi m_j u_j}{k_j}) \Big). \tag{18}$$

In practice, to have fewer coefficients to estimate, we shall model the coefficients $\phi_r(t)$ and their components as

$$\phi_{r,j}(u_j) = \sum_{m_j=1}^{H_j} \Big( a_{r,m_j}^{(j)} \cos(\frac{2\pi m_j u_j}{k_j}) + b_{r,m_j}^{(j)} \sin(\frac{2\pi m_j u_j}{k_j}) \Big), \tag{19}$$

where $H_j \le \lfloor k_j/2 \rfloor$. The parameters $\mu_j(u_j)$, $\sigma_j^{(2)}(u_j)$, $j = 1,2$, on the other hand, will be estimated in a nonparametric fashion, though again a parametric route analogous to (19) is also a possibility.

*Example 3.* We continue with the setting of Example 1. In this example, by combining (12) and (19), the functions $\phi_{r,j}(g_j(t))$ are modeled as

$$\phi_{r,1}(g_1(t)) = \sum_{m_1=1}^{H_1} \left( a_{r,m_1}^{(1)} \cos(\frac{2\pi m_1 t}{24}) + b_{r,m_1}^{(1)} \sin(\frac{2\pi m_1 t}{24}) \right) \tag{20}$$

and

$$\phi_{r,2}(g_2(t)) = \sum_{m_2=1}^{H_2} \left( a_{r,m_2}^{(2)} \cos(\frac{2\pi m_2 \lceil t/24 \rceil}{7}) + b_{r,m_2}^{(2)} \sin(\frac{2\pi m_2 \lceil t/24 \rceil}{7}) \right). \tag{21}$$

We note again that the function $\phi_{r,1}(g_1(t))$ is 24-periodic, and that $\phi_{r,2}(g_2(t))$ is 168-periodic but also constant over successive intervals of length 24.

*Remark 3.* The framework described above can be extended straightforwardly to the case of multiple periodic effects, by introducing additional functions $g_j(t)$ associated with these effects.

*Remark 4.* As A–PAR($p$) models discussed in Remark 1, B–PAR($p$) models are also PAR($p$) models with the larger period $s_2$. It is instructive here to contrast the two introduced models from the perspective of these standard PAR models. A PAR($p$) model with period $s_2$ has its coefficients vary periodically with period $s_2$. These coefficients can always be expressed through a Fourier representation. In the applications of the A–PAR model, only a small number of these Fourier coefficients are assumed to be non-zero, more specifically, the first few in the Fourier representation and also the first few in the component of the representation that is $s_1$-periodic. The B–PAR model, on the other hand, assumes that the periodicity consist of two additive effects associated with two periodic nominal variables. The latter effects do not need to be components of the Fourier representation of the model coefficients (as, for example, the coefficients (21) above).

*Remark 5.* The preceding remark also suggests that A–PAR and B–PAR models might serve quite different purposes. By increasing the number of non-zero coefficients in the A–PAR model Fourier representation, one could effectively get any PAR model with period $s_2$. From this perspective, the A–PAR model is quite flexible. With the B–PAR model, on the other hand, one might be more interested in which effects and which of their levels are more pronounced in the dynamics of the PAR process. This is illustrated further in our application to a real data set in Section 5.

## 3 Estimation procedure

We discuss here estimation of the parameters $\mu(t)$, $\sigma(t)^2$ and $\phi_r(t)$ of the A–PAR and B–PAR models, using the Fourier representations (11) and (19) of the parameters. The way the A–PAR and B–PAR models were introduced allows us to present essentially a unified estimation framework. We suppose that the observed data consist of observations $X_1, \ldots, X_T$, where the sample size $T$ is a multiple of both $s_1$ and $s_2$ for simplicity.

### 3.1 Estimation of mean

For an A–PAR model, we estimate the means as $\widehat{\mu}_0 = \overline{X}$ (the overall mean),

$$\widehat{\mu}_1(t) = \frac{1}{(T/s_1)} \sum_{n=1}^{T/s_1} (X_{t+s_1(n-1)} - \overline{X}), \quad t = 1, \ldots, s_1, \tag{22}$$

and extended periodically with period $s_1$ for other $t$'s, and

$$\widehat{\mu}_2(t) = \frac{1}{(T/s_2)} \sum_{n=1}^{T/s_2} (X_{t+s_2(n-1)} - \widehat{\mu}_1(t)), \quad t = 1, \ldots, s_2, \tag{23}$$

and extended periodically with period $s_2$ for other $t$'s. Once can check that $\widehat{\mu}(t) = \widehat{\mu}_0 + \widehat{\mu}_1(t) + \widehat{\mu}_2(t)$ is just the periodic mean at period $s_2$. For a B–PAR model, the mean effects are estimated through a least squares regression of $X_t$ on the two nominal variables described in the beginning of Section 2.2. Again, let $\widehat{\mu}(t)$ be the overall estimated mean which is generally different from that for the A–PAR model (see Figure 1 in Section 5 for an illustration).

### 3.2 OLS estimation

Let $\widehat{Y}_t = X_t - \widehat{\mu}(t)$. In applying the ordinary least squares (OLS), the model parameters are estimated as

$$\left\{ \widetilde{\phi}_{r,0}, \widetilde{a}^{(j)}_{r,m_j}, \widetilde{b}^{(j)}_{r,m_j} \right\}_{r=1,\ldots,p, m_j=1,\ldots,H_j, j=1,2}$$

$$= \underset{\phi_{r,0},, a^{(j)}_{r,m_j}, b^{(j)}_{r,m_j}}{\operatorname{argmin}} \sum_t (\widehat{Y}_t - \phi_1(t)\widehat{Y}_{t-1} - \ldots - \phi_p(t)\widehat{Y}_{t-p})^2, \tag{24}$$

where $\phi_r(t) = \phi_{r,0} + \phi_{r,1}(t) + \phi_{r,2}(t)$ and $\phi_{r,j}(t)$ are given in (11) or (19), depending on the type of the model. Let $\widetilde{\phi}_r(t)$ be the resulting OLS parameter estimators. Consider also the errors

$$\widetilde{\eta}_t = \widehat{Y}_t - \widetilde{\phi}_1(t)\widehat{Y}_{t-1} - \ldots - \widetilde{\phi}_p(t)\widehat{Y}_{t-p}. \tag{25}$$

The model parameter $\sigma(t)^2$ and its components $\sigma_0^2$, $\sigma_1^{(2)}(t)$, $\sigma_2^{(2)}(t)$ could then be estimated analogously to the mean $\mu(t)$ and its three components as in Section 3.1 but replacing $X_t$ with $\widetilde{\eta}_t^2$. We shall refer to $\widetilde{\eta}_t/\widetilde{\sigma}(t)$ as the residuals from the OLS estimation.

*Remark 6.* There are several potential issues with the suggested estimation of $\sigma(t)^2$ that, in particular, are encountered in the application in Section 5. When $T/s_2$ is small (e.g. $T/s_2 = 6$ in the application considered below) and $\widetilde{\sigma}(t)^2$ is computed

as the $s_2$-periodic sample mean, note that the estimation of each $\sigma(t)^2$ involves just $T/s_2$ error terms (e.g. 6 in the application below). The quality of estimation of $\sigma(t)^2$ is then dubious, and we try to rectify this by slightly smoothing the estimates over time. This procedure does have some minor effect on the estimates and their standard errors, and might call for further investigation in the future. (We do not perform smoothing when estimating the mean $\mu(t)$ since we expect these estimates to be already quite smooth.) On the other hand, for Model B, we also note that the suggested procedure is not guaranteed to yield nonnegative estimates of $\sigma(t)^2$, which also happens in our application. In this case, we use the estimates of $\sigma(t)^2$ obtained for Model A.

### *3.3 WLS estimation*

Having the OLS estimate $\widetilde{\sigma}(t)^2$ of the variance of the error terms, the model parameters could be reestimated by using the weighted least squares (WLS) as

$$\left\{\widehat{\phi}_{r,0}, \widehat{a}^{(j)}_{r,m_j}, \widehat{b}^{(j)}_{r,m_j}\right\}_{r=1,\ldots,p, m_j=1,\ldots,H_j, j=1,2}$$

$$= \underset{\phi_{r,0}, a^{(j)}_{r,m_j}, b^{(j)}_{r,m_j}}{\operatorname{argmin}} \sum_t (\widehat{Y}_t - \phi_1(t)\widehat{Y}_{t-1} - \ldots - \phi_p(t)\widehat{Y}_{t-p})^2 / \widetilde{\sigma}(t)^2. \tag{26}$$

Likewise, the variance $\sigma(t)^2$ could be reestimated as $\widehat{\sigma}(t)^2$ by using the model errors based on the WLS estimates (and this process could be iterated till convergence occurs), with possible modifications discussed in Remark 6 above. Letting $\widehat{\eta}_t$ be the error terms from the WLS estimation, defined similarly to (25), the WLS residuals are defined as $\widehat{\eta}_t/\widetilde{\sigma}(t)$.

## 4 Inference and other tasks

In the implementation of the OLS and WLS estimation, a PAR$(p)$ model is expressed in the form of a linear regression as

$$Y = R\alpha + Z. \tag{27}$$

For example, for an A–PAR$(p)$ model, $Y = (\widehat{Y}_{p+1}, \ldots, \widehat{Y}_T)'$ is a $(T-p)$–vector of periodically demeaned observations $\widehat{Y}_t$, $\alpha = (\alpha_1' \ldots \alpha_p')'$ is a $((1+2H_1+2H_2)p)$–vector of parameters with

$$\alpha_r = (\phi_{r,0}, a^{(1)}_{r,1}, \ldots, a^{(1)}_{r,H_1}, b^{(1)}_{r,1}, \ldots, b^{(1)}_{r,H_1}, a^{(2)}_{r,1}, \ldots, a^{(2)}_{r,H_2}, b^{(2)}_{r,1}, \ldots, b^{(2)}_{r,H_2})',$$

the regressors $R$ can be expressed as a $(T-p)\times((1+2H_1+2H_2)p)$ matrix $(R_{p+1}\dots R_T)'$ with $R_t=(I_p\otimes B_t)Y_{t,lags}$, $Y_{t,lags}=(Y_{t-1},\dots,Y_{t-p})'$,

$$B_t=\Big(1,\cos(\frac{2\pi t}{s_1}),\dots,\cos(\frac{2\pi H_1 t}{s_1}),\sin(\frac{2\pi t}{s_1}),\dots,\sin(\frac{2\pi H_1 t}{s_1}),\\ \cos(\frac{2\pi t}{s_2}),\dots,\cos(\frac{2\pi H_2 t}{s_2}),\sin(\frac{2\pi t}{s_2}),\dots,\sin(\frac{2\pi H_2 t}{s_2})\Big)'$$

and $Z$ refers to the error terms. Within the linear formulation (27), the OLS and WLS parameter estimators and their standard errors have well-known expressions in terms of $R$ and $Y$, which we use here as well but omit for the shortness sake.

In addition to the OLS and WLS estimation as outlined above, we also use their counterparts when some of the coefficients are set to 0. We shall refer to the corresponding models as *restricted* PAR models. Estimation and computing standard errors for restricted PAR modelds are carried out in a standard way by expressing zero constraints through

$$\alpha=C\gamma, \tag{28}$$

where $\gamma$ is a $k$–vector of non-zero coefficients and $C$ is a $((1+2H_1+2H_2)p)\times k$ restriction matrix, with rows of zeros corresponding to the zero elements of $\alpha$, and rows with a single entry of 1 corresponding to non-zero elements of $\alpha$. The OLS and WLS estimation and inference are then performed essentially by replacing $R$ by $RC$.

If needed, model selection can be guided by some information criterion, such as BIC and AIC defined in the usual way as $(-2)$ multiplied by the log-likelihood, with an appropriate penalty. In the data application below, we shall be guided by looking at parameter "significance" and suitable diagnostics plots of model residuals. Similarly, the introduced PAR models can be used in forecasting in a straightforward way as with standard AR models and their PAR extensions. Out-of-sample forecasting performance could also be employed as another tool for selecting a model.

*Remark 7.* Under mild assumptions on the residuals $\{\varepsilon_t\}$ in the A–PAR and B–PAR models (with typical assumptions being the i.i.d. property and finiteness of the 4th moment), the parameter estimators $\{\widetilde{\phi}_{r,0},\widetilde{a}^{(j)}_{r,m_j},\widetilde{b}^{(j)}_{r,m_j}\}$ in (24) and $\{\widehat{\phi}_{r,0},\widehat{a}^{(j)}_{r,m_j},\widehat{b}^{(j)}_{r,m_j}\}$ in (26) (assuming the true variance $\sigma^2(t)$ is used in estimation) are expected to be asymptotically normal. Indeed, these estimators are linear transformations of the analogous PAR model parameter estimators $\{\widetilde{\phi}_r(t)\}$ and $\{\widehat{\phi}_r(t)\}$. The asymptotic normality of the latter under mild assumptions is proved in Basawa and Lund [2], Anderson and Meerschaert [1]. The analogous linear transformation argument to establish the asymptotic normality of the coefficient estimators in the Fourier representation of the parameters is also used in Tesfaye et al. [21].

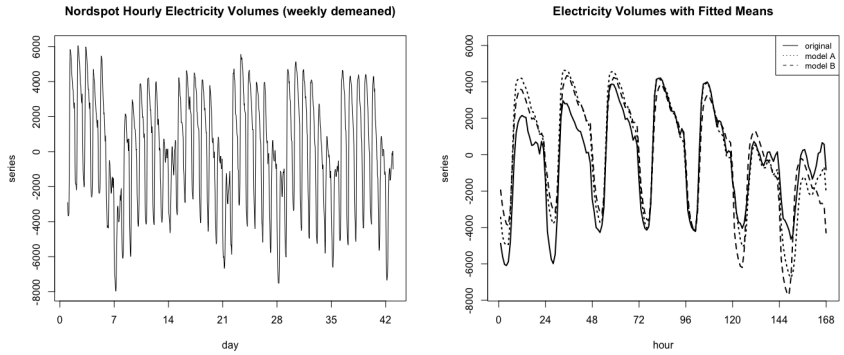


**Fig. 1** Left: Weekly demeaned energy volume series for 6 weeks. Right: The volume series for Week 2 with estimated means according to Models A and B.

## 5 Data application

To illustrate our proposed models, we shall work with a time series of hourly electricity volumes from Nord Pool Spot Exchange.[1] This data was considered in a number of other works related to periodically correlated series, for example, Dudek et al. [7]. We consider the series for 6 weeks in 2008, and remove the weekly mean from the data. The length of the series is thus $T = 1,008$. Note that 6 weeks (1 week being the period of the underlying PAR model) are sufficient for our modeling purposes since the number of parameters is reduced considerably through the Fourier representations. For example, a small number of non-zero coefficients in the Fourier representation could be estimated, in principle, even from the data covering just one period. The resulting series is presented in Figure 1, left plot. The right plot of the figure presents one week of the series with the mean effects estimated according to Models A and B. In the rest of the section, we shall fit Models A and B to the periodically demeaned series, that is, the difference between the observed and fitted values in Figure 1, right plot.

### *5.1 Fitting Model A*

Figure 2 depicts the periodically demeaned series according to Model A, and its sample PACF. The sample PACF suggests including lags 1, 2 and 24 into an autoregressive model. Figure 3 presents two commonly used plots to detect periodic correlations: the spectral coherence plot according to Hurd and Gerr [15] (left plot of the figure), and a related test statistic with a critical value line from Lund [18] (right plot; with a tuning parameter $M = 10$ in Lund [18]). See also Hurd and Mi-

[1] http://www.npspot.com

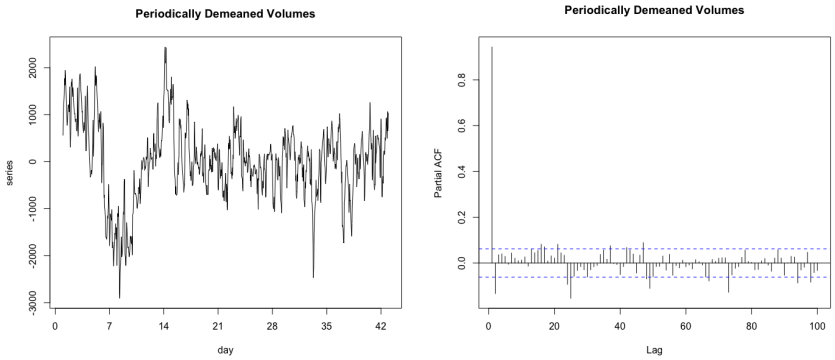


**Fig. 2** Left: Periodically demeaned volume series for 6 weeks (Model A). Right: The corresponding sample PACF.

amee [16], Sections 10.4 and 10.5 The spectral coherence is plotted using the R package perARMA [9].

If a series exhibits periodic correlations at period $s$, the spectral coherence plot should have diagonal lines emerging at multiples of the index $T/s$. Here, $T/s = 1,008/s$. The plot in Figure 3 suggests the first major diagonal line around the index 40. In fact, it corresponds to the period $s_1 = 24$ with $T/s_1 = 42$. There are also traces of diagonal lines at indices smaller than 42 but it is difficult to say for sure what these indices are. The latter could be determined easier from the Lund test statistic plot, which essentially averages the spectral coherence statistic at different indices along the corresponding diagonals, and also provides a critical value (the horizontal dashed line in the plot). As expected, the Lund test statistic has a large value at index 42. But note also that the values are larger, some above the critical values, at multiples of the index 6. This index corresponds to the period $s_2 = 168$ (1 week) since $T/s_2 = 6$. We thus conclude from these plots that periodic correlations are present in the periodically demeaned series at both periods $s_1 = 24$ and $s_2 = 168$.

We also see the presence of periodic correlations at the two periods $s_1 = 24$ and $s_2 = 168$ when fitting Model A. We shall report here on our fitting attempts for A–PAR($p$) models of orders $p = 2$ and $p = 26$, to accomodate the partial autocorrelations seen at these lags in Figure 2. Experimenting with various restricted A–PAR(2) models, we settled on the model with the following non-zero WLS estimated coefficients, with the standard errors indicated in the parentheses: at lag 1,

$$\begin{aligned}
&\widehat{\phi}_{1,0} = 1.104\,(0.025),\\
&\widehat{a}^{(1)}_{1,5} = -0.291\,(0.038), \quad \widehat{a}^{(1)}_{1,10} = -0.102\,(0.037),\\
&\widehat{b}^{(1)}_{1,7} = 0.202\,(0.036), \quad \widehat{b}^{(1)}_{1,9} = 0.081\,(0.041),\\
&\widehat{a}^{(2)}_{1,1} = 0.023\,(0.012)
\end{aligned}$$

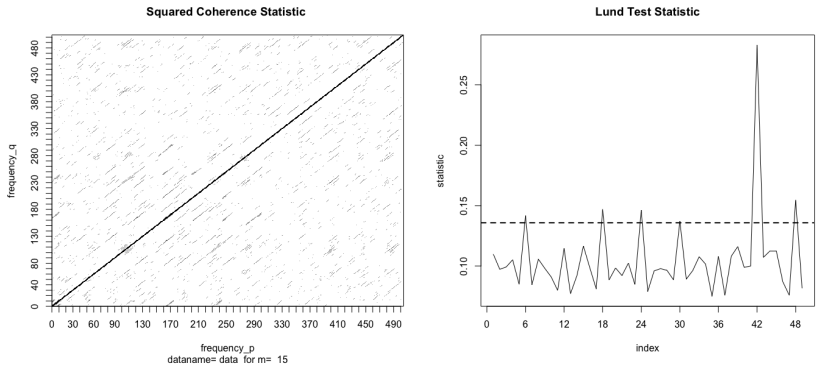


**Fig. 3** Left: The spectral coherence plot for periodically demeaned volume series for 6 weeks (Model A). Right: The Lund test statistic for the same series with a horizontal dashed line indicating the critical value.

and at lag 2,

$$\begin{aligned}
&\widehat{\phi}_{2,0} = -0.178\,(0.025),\\
&\widehat{a}^{(1)}_{2,5} = 0.245\,(0.038), \quad \widehat{a}^{(1)}_{2,10} = 0.084\,(0.037),\\
&\widehat{b}^{(1)}_{2,7} = -0.195\,(0.036), \quad \widehat{b}^{(1)}_{2,9} = -0.082\,(0.040).
\end{aligned}$$

Note that only one non-zero coefficient, namely $\widehat{a}^{(2)}_{1,1}$, is included in the component for period $s_2 = 168$. The resulting WLS estimated parameter functions $\widehat{\phi}_1(t)$ and $\widehat{\phi}_2(t)$ are plotted in Figure 4. The component of the mean with the non-zero coefficient $\widehat{a}^{(2)}_{1,1}$ at period $s_2 = 168$ produces a "global" trend in the coefficients $\widehat{\phi}_1(t)$ over the 168 hours, which is clearly visible in the left plot. Without this global trend, the coefficients can be checked to be close to what one would get from fitting a standard PAR$(2)$ model with period $s_1 = 24$.

Figure 5 depicts the sample ACF and the Lund test statistic for the WLS residuals of the fitted A–PAR$(2)$ model. Note some remaining autocorrelations around lag 24, which should not be surprising since we fitted a PAR model of order $p = 2$. The plot with the Lund test statistic is depicted using the same vertical scale as in Figure 3: the peaks at dominant indices have become smaller in general but are not completely negligible.

To remove the remaining autocorrelations in the residuals, one could fit an A–PAR$(p)$ model of higher order $p$. (Another possibility would be to use a seasonal PAR model as in Basawa et al. [3].) In analogy to non-periodic seasonal models, we have experimented with fitting restricted A–PAR$(26)$, by allowing some of the coefficients at lags 24, 25 and 26 to be non-zero. We shall not report here the fitted models but rather indicate several key observations. We found significant periodicity in the coefficients $\phi_{24}(t)$, $\phi_{25}(t)$ and $\phi_{26}(t)$, but also only in the component with

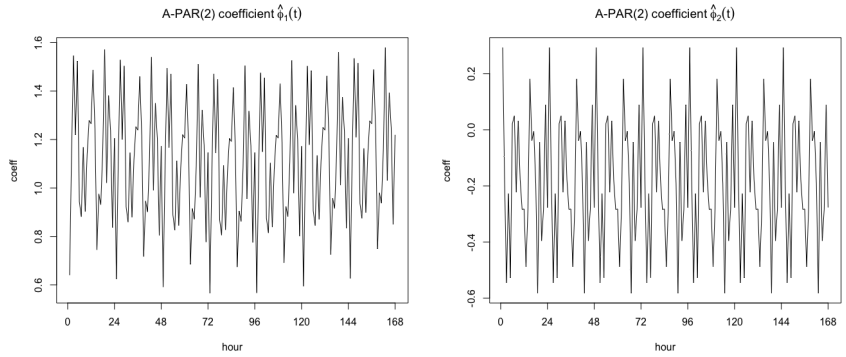


**Fig. 4** The WLS estimated parameter functions $\widehat{\phi_1}(t)$ and $\widehat{\phi_2}(t)$ of the fitted A–PAR$(2)$ model.

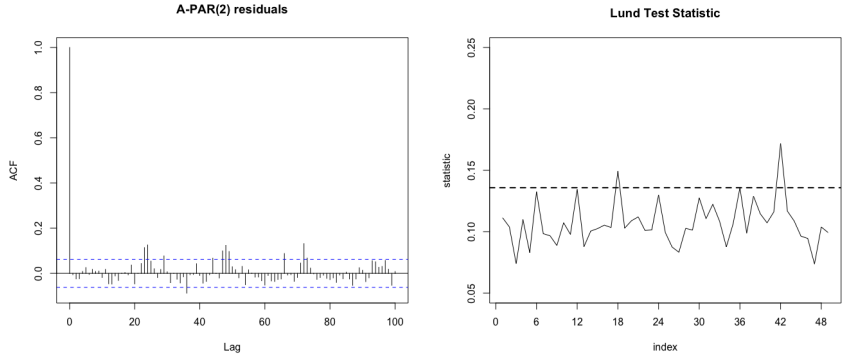


**Fig. 5** The sample ACF and the Lund test statistic for the WLS residuals of the fitted A–PAR$(2)$ model.

period $s_1 = 24$. Typical sample ACF and Lund statistic plots for the WLS residuals of a fitted restricted A–PAR$(26)$ are presented in Figure 6. Note the smaller autocorrelations around multiples of lag 24 compared to those in Figure 5. The Lund statistic plot continues having several peaks above the critical value line but their locations are no longer multiples of 6. (For example, the largest peak is no longer at 42.) It remains to clarify what might cause this shift in indices where peaks are present.

## *5.2 Fitting Model B*

We now turn to fitting Model B and follow a similar presentation structure as for Model A in the previous section. Figure 7 presents similarly the periodically demeaned volume series according to Model B and its sample PACF. Figure 8 depicts

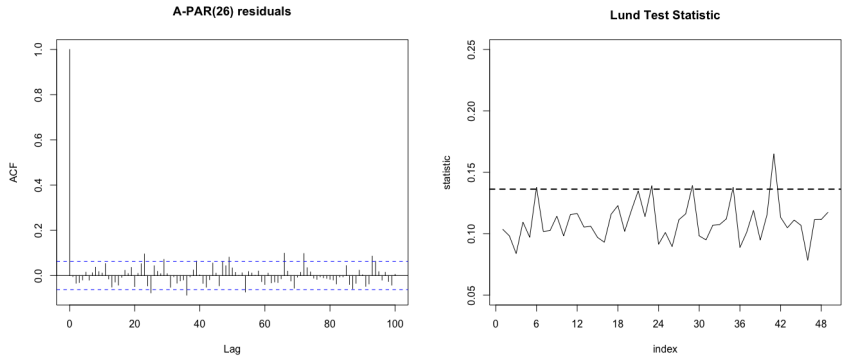


**Fig. 6** The sample ACF and the Lund test statistic for the residuals of the fitted restricted A–PAR(26) model.

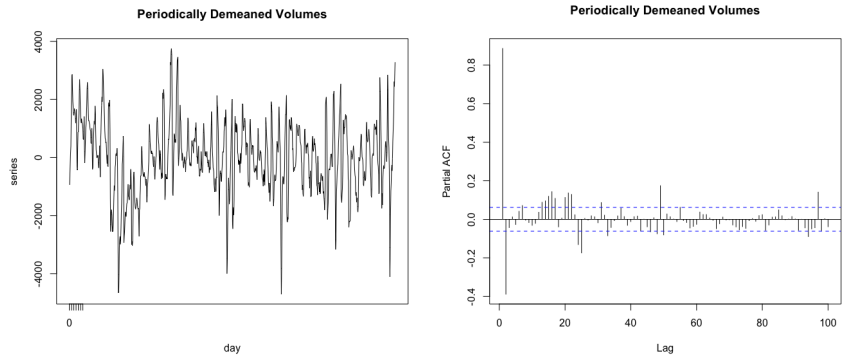


**Fig. 7** Left: Periodically demeaned volume series for 6 weeks (Model B). Right: The corresponding sample PACF.

the spectral coherence and Lund statistic plots. Note that the diagonal lines at the multiples of the indices 6 and 42 in the coherence plot, as well as the peaks at these indices in the Lund statistic plot, are much more pronounced compared to those in Figure 3. This interesting difference is due to the way the mean effect is computed in Model B.

When fitting a B–PAR(2) model with $H_1 = 10$ and $H_2 = 3$ in the representations (20) and (21), and then reestimating it through a restricted B–PAR(2) model when including only the significant coefficients from the non-restricted model, leads to the following significant non-zero coefficients: at lag 1, $\phi_{1,0}$,

$$a^{(1)}_{1,m_1} : m_1 = 2,3,6,7,8,9, \quad b^{(1)}_{1,m_1} : m_1 = 3,4,6,9,10,$$
$$a^{(2)}_{1,m_2} : m_2 = 1,3, \quad b^{(2)}_{1,m_2} : m_2 = 2,$$

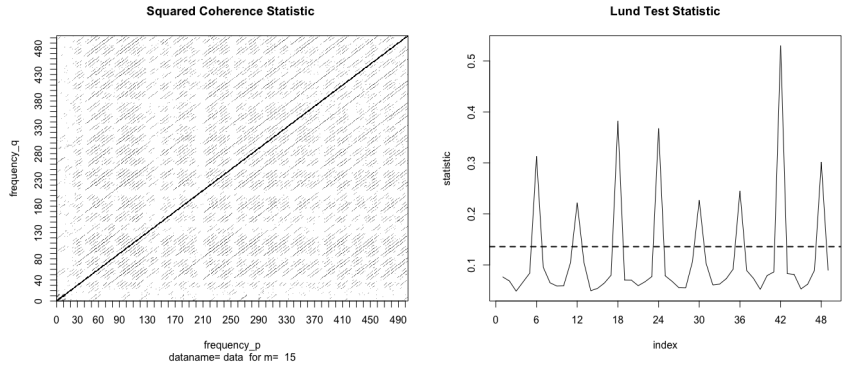


**Fig. 8** Left: The spectral coherence plot for periodically demeaned volume series for 6 weeks (Model B). Right: The Lund test statistic for the same series with a horizontal dashed line indicating the critical value.

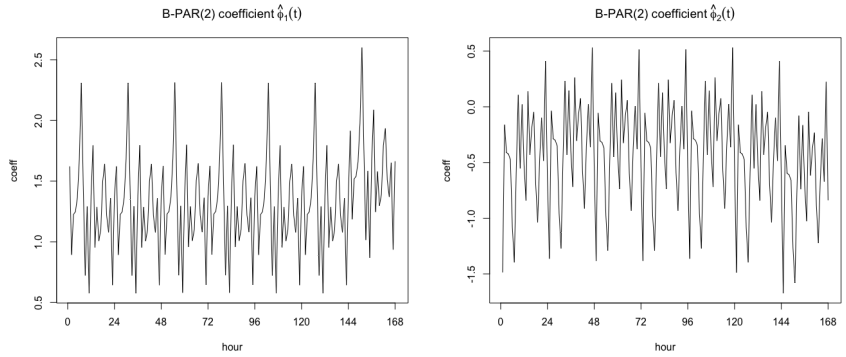


**Fig. 9** The estimated parameter functions $\widehat{\phi}_1(t)$ and $\widehat{\phi}_2(t)$ of the fitted B–PAR$(2)$ model.

and at lag 2, $\phi_{2,0}$,

$$a^{(1)}_{2,m_1} : m_1 = 1,2,3,7,8,10, \quad b^{(1)}_{2,m_1} : m_1 = 4,6,10, \quad a^{(2)}_{2,m_2} : m_2 = 1,3.$$

We shall not indicate here the values and standard errors of the corresponding WLS estimates but rather present a few revealing plots of the coefficient functions. More specifically, Figure 10 shows the WLS estimated parameter functions $\widehat{\phi}_1(t)$ and $\widehat{\phi}_2(t)$ of the fitted B–PAR$(2)$ model. Note that the effect of the day of a week, especially that of Sunday, is more apparent in the figure when compared to Figure 4. This can also be seen clearer through the two components $\widehat{\phi}_{r,1}(g_1(t))$ and $\widehat{\phi}_{r,2}(g_2(t))$ depicted in Figure 9, where the effects of the day (solid line) is more pronounced towards Sunday for lag 1 and Saturday through Monday for lag 2 coefficients.

Figure 11 depicts the sample ACF and the Lund test statistic for the WLS residuals of the fitted B–PAR$(2)$ model. The conclusions are not very different from those

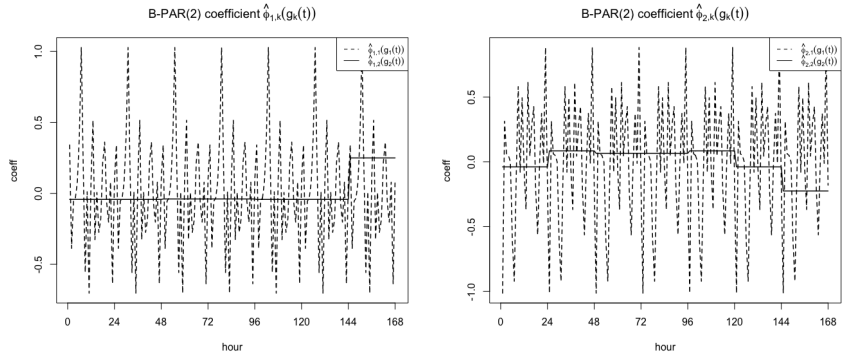


**Fig. 10** The estimated parameter functions $\widehat{\phi}_{1,k}(g_k(t))$ and $\widehat{\phi}_{2,k}(g_k(t))$ of the fitted B–PAR$(2)$ model.

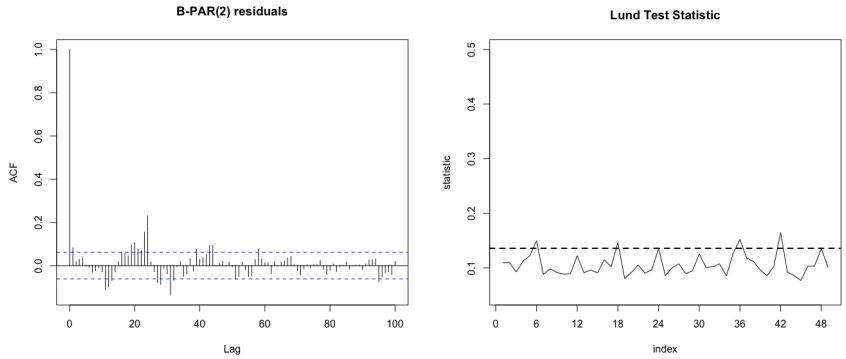


**Fig. 11** The sample ACF and the Lund test statistic for the WLS residuals of the fitted B–PAR$(2)$ model.

for the A–PAR$(2)$ model from Figure 5. In particular, as with Model A above, one could fit a B–PAR$(p)$ model with higher order $p$ to remove the remaining autocorrelations around lag 24 in the WLS nresiduals.

## 6 Conclusions

In this work, we introduced two periodic autoregression models with two or more periodic effects, discussed their inference and presented an application, showing their relevance for real data. Some of the issues that can be explored in the future include: incorporating moving average components into our models, comparing out-of-sample forecasting performance between the introduced and among competing

models, applications to other data sets, clarifying the role of the used estimation methods for error variances, and others.

**Acknowledgements** The second author was supported in part by NSF Grant DMS-17-12966 at the University of North Carolina.

## References

1. P. L. Anderson and M. M. Meerschaert. Parameter estimation for periodically stationary time series. *Journal of Time Series Analysis*, 26(4):489–518, 2005.
2. I. V. Basawa and R. Lund. Large sample properties of parameter estimates for periodic ARMA models. *Journal of Time Series Analysis*, 22(6):651–663, 2001.
3. I. V. Basawa, R. Lund, and Q. Shao. First-order seasonal autoregressive processes with periodically varying parameters. *Statistics & Probability Letters*, 67(4):299–306, 2004.
4. G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis*. Prentice Hall, Inc., Englewood Cliffs, NJ, third edition, 1994. Forecasting and control.
5. L. Dannecker. *Energy Time Series Forecasting: Efficient and Accurate Forecasting of Evolving Time Series from the Energy Domain*. Springer, 2015.
6. A. M. De Livera, R. J. Hyndman, and R. D. Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496):1513–1527, 2011.
7. A. E. Dudek, H. Hurd, and W. Wójtowicz. PARMA models with applications in R. In *Cyclostationarity: Theory and Methods-II*, pages 131–153. Springer, 2015.
8. A. E. Dudek, H. Hurd, and W. Wójtowicz. Periodic autoregressive moving average methods based on Fourier representation of periodic coefficients. *Wiley Interdisciplinary Reviews: Computational Statistics*, 8(3):130–149, 2016.
9. A. E. Dudek, H. L. Hurd, and W. Wójtowicz. *perARMA: Periodic Time Series Analysis*, 2016. R package version 1.6.
10. G. Dudek. Forecasting time series with multiple seasonal cycles using neural networks with local learning. In *International Conference on Artificial Intelligence and Soft Computing*, pages 52–63. Springer, 2013.
11. G. Dudek. Generalized regression neural network for forecasting time series with multiple seasonal cycles. In *Intelligent Systems' 2014*, pages 839–846. Springer, 2015.
12. P. H. Franses and R. Paap. *Periodic Time Series Models*. Advanced Texts in Econometrics. Oxford University Press, Oxford, 2004.
13. E. Ghysels and D. R. Osborn. *The Econometric Analysis of Seasonal Time Series*. Themes in Modern Econometrics. Cambridge University Press, Cambridge, 2001. With a foreword by Thomas J. Sargent.
14. P. G. Gould, A. B. Koehler, J. K. Ord, R. D. Snyder, R.J. Hyndman, and F. Vahid-Araghi. Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research*, 191(1):207–222, 2008.
15. 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.
16. H. L. Hurd and A. Miamee. *Periodically Correlated Random Sequences*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2007. Spectral theory and practice.
17. R. H. Jones and W. M. Brelsford. Time series with periodic structure. *Biometrika*, 54:403–408, 1967.
18. R. B. Lund. Choosing seasonal autocovariance structures: PARMA or SARMA. In W. R. Bell, S. H. Holan, and T. S. McElroy, editors, *Economic Time Series: Modelling and Seasonality*, pages 63–80. Chapman and Hall, 2011.

19. H. Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin, 2005.
20. J. W. Taylor, L. M. De Menezes, and P. E. McSharry. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, 22(1):1–16, 2006.
21. Y. G. Tesfaye, P. L. Anderson, and M. M. Meerschaert. Asymptotic results for Fourier-PARMA time series. *Journal of Time Series Analysis*, 32(2):157–174, 2011.
22. R. Weron. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. John Wiley & Sons, 2007.