Processes with Periodicity

by

Harry L. Hurd
University of North Carolina
Chapel Hill, NC

1. Periodic signal in additive noise
2. Periodically stationary, in 2nd order case, periodically correlated
3. PARMA
4. SARMA

Periodic signal in additive noise : an old problem but with many modern applications

- The beginning: meteorology, physical processes related to planetary motion
  - A. Shuster, "On lunar and solar periodicities of earthquakes," *Proc. Roy. Soc.*, v. 61, pp. 455-465, 1897.
  - A. Shuster, "On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena," *Terr. Magn.*, v. 3, pp. 13-41, 1898.
  - The periodogram. A method for illuminating the presence of periodic functions in noisy observations. As a quick reminder, given a finite sequence of real or complex numbers $X_N = \{x_0, x_1, \ldots, x_{N-1}\}$, the discrete Fourier transform of X_N is

$$F_N(\lambda) = \sum_{t=0}^{N-1} x_k \exp(i\lambda t) \tag{1}$$

The usual periodogram is defined simply as

$$I_N(\lambda) = \frac{1}{2\pi N} |F_N(\lambda)|^2. \tag{2}$$

Periodogram Illustration

$$Y_t = A\cos(\pi t/16) + \xi_t$$

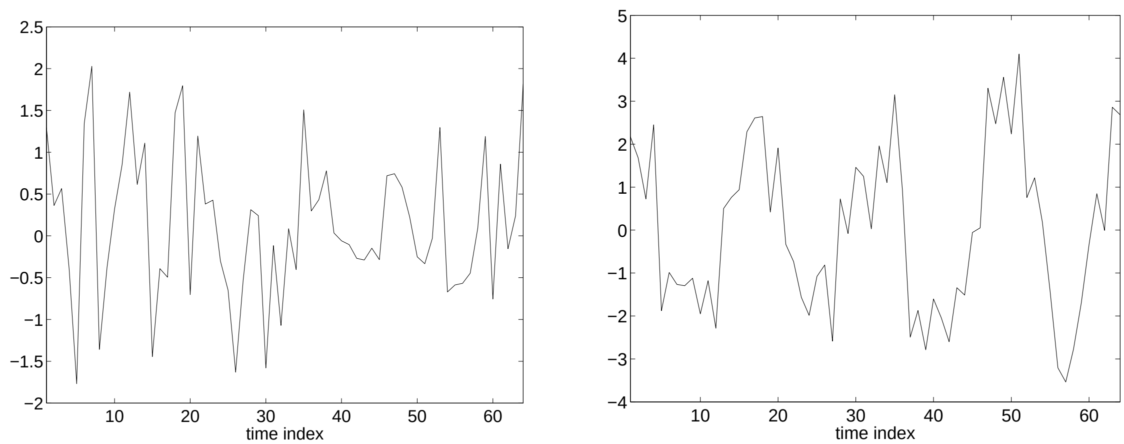


Figure 1: Time series $Y_t = A\cos(\pi t/16) + \xi_t, \sigma_\xi = 1$. Left: $A = .5/512$. Right: $A = 2.5/512$.

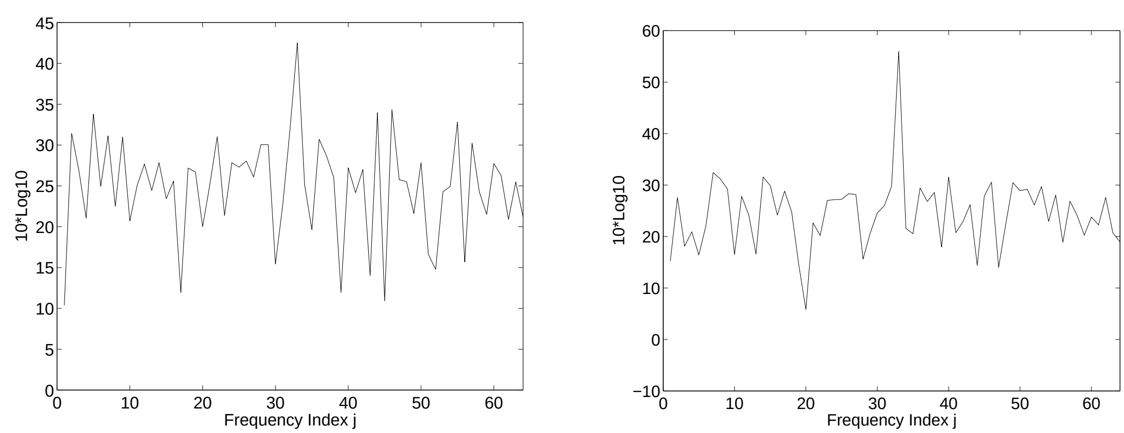


Figure 2: Periodograms of $Y_t = A\cos(\pi t/16) + \xi_t, \sigma_\xi = 1$ based on 1 FFT of length 512. Left: $A = .5/512$. Right: $A = 2.5/512$.

Periodogram and Periodic Signals

- The periodogram is the starting place for an important estimator of spectral density of a stationary process. When smoothed in the correct manner, this smoothed periodogram is an asymptotically consistent estimator of spectral density. See B and D. The modern topics of kernel smoothing seems to have originated here.

- Modern applications of periodic signals in additive noise.

  1. Signal processing of sonar and radar signals.
  2. Coding and decoding of communication signals.
  3. Meteorlogical, climatological, economic processes;

- Spectral meaning of periodic signals: Correlation functions $r(\tau)$ of stationary sequences are Fourier transforms (Herglotz)

$$r(\tau) = \int_0^{2\pi} \exp(i\lambda\tau) dF(\lambda) \tag{3}$$

  where $F(\lambda)$ is a bounded, non-decreasing function on $[0, 2\pi)$. (equivalently a non negative measure). If λ is a jump point $F(\lambda+) - F(\lambda) > 0$, then the process will have a periodic component with frequency λ.

Solar radiation I

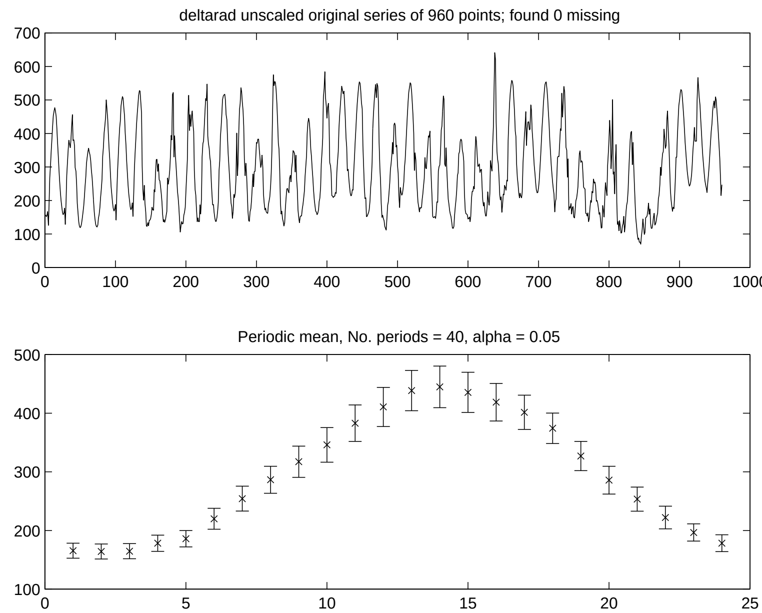


Figure 3: (Top) Solar radiation time series. (Bottom) Sample periodic mean of radiation series using $T = 24$. The error bars are the $\alpha = .05$ confidence limits determined by the student's t with sample size $N_p = 40$. From *Taconite Inlet Project* (`http:/climate4/geo.umass.edu/TILPHTMLhomepage.html`.)

The *sample periodic mean* for period T is computed by

$$\hat{m}_N(t) = \frac{1}{N} \sum_{p=0}^{N-1} X_{t+pT}, \quad t = 1, 2, ..., T \tag{4}$$

Solar radiation II

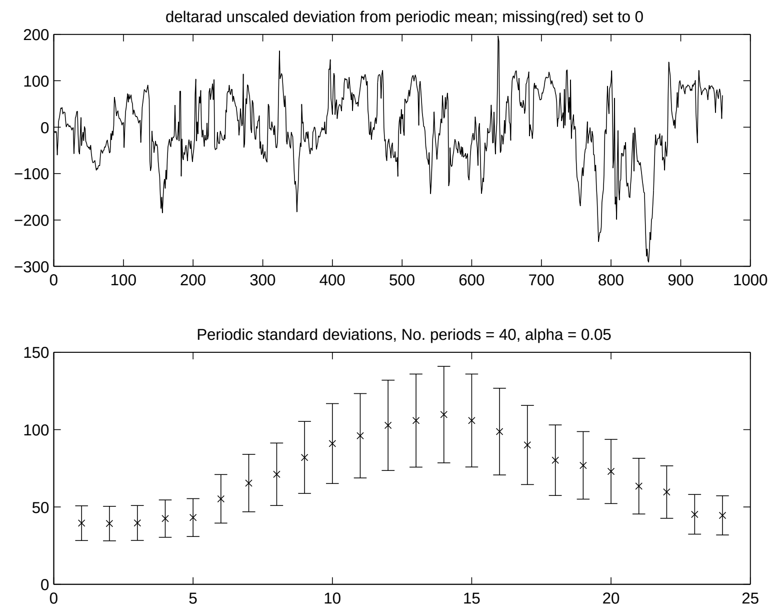


Figure 4: (Top) Deviation around sample periodic mean. (Bottom) Sample periodic standard deviation of radiation series using $T = 24$. The error bars are the $\alpha = .05$ confidence limits determined by the chi-squared distribution with $N_p - 1 = 39$ degrees of freedom.

The *sample periodic variance* is given by

$$S_N^2(t) = \frac{1}{N-1} \sum_{p=0}^{N-1} Y_{t+pT}^2, \quad t = 1, 2, ..., T \tag{5}$$

where $Y_t = X_t - \hat{m}_N(t)$ is the deviation of X_t from the sample periodic mean $\hat{m}_N(t)$.
So this series has a visually clear (suggesting very significant) periodicity in it's mean and variance. This motivates the examination of processes with some sort of structural time periodicity.

Processes invariant under T-shifts

Strict sense A process $X_t(\omega) : \Omega \longrightarrow \mathbf{C}$ or $\mathbf{R}$ is called periodically stationary with period T if for every n, collection of times $t_1, t_2, ..., t_n$ in $\mathbf{Z}$ or $\mathbf{R}$, collection of Borel sets $A_1, A_2, ..., A_n$ of $\mathbf{C}$ or $\mathbf{R}$,

$$\begin{aligned} & Pr[X_{t_1+T} \in A_1, X_{t_2+T} \in A_2, ..., X_{t_n+T} \in A_n] \\ = \; & Pr[X_{t_1} \in A_1, X_{t_2} \in A_2, ..., X_{t_n} \in A_n]. \end{aligned} \tag{6}$$

and there are no smaller values of $T > 0$ for which (6) holds.
Synonyms for *periodically stationary* include *periodically non-stationary*, *cyclostationary* (think of cyclically stationary), *processes with periodic structure*, and a few others. If $T = 1$, the process is strictly stationary.

Weak sense A second order process $X_t \in L_2(\Omega, \mathcal{F}, P)$ with $t \in \mathbf{Z}$ is called *periodically correlated* (or wide-sense cyclostationary) with period T if

$$m(t) = E\{X_t\} = m(t+T) \forall t, \text{ and} \tag{7}$$

$$R(s,t) = E\{X_s \overline{X_t}\} = R(s+T, t+T) \forall s, t \in \mathbf{Z} \tag{8}$$

and there are no smaller values of $T > 0$ for which (7) and (8) hold. If $T = 1$, the process is weakly (or wide-sense) stationary.

Periodically correlated processes have a spectral theory that is an extension of the spectral theory for stationary processes.
Just as ARMA models provide parametric models of *stationary processes*, the PARMA models provide parametric models for *periodically correlated* processes.

Mean and covariance for PC processes

The periodicity leads to Fourier series representations

$$m(t) = E\{X_t\} = \sum_{k=0}^{T-1} m_k e^{i2\pi kt/T}$$

and

$$R(t+\tau, t) = E\{X_{t+\tau}\overline{X_t}\} = \sum_{k=0}^{T-1} e^{i2\pi kt/T} B_k(\tau)$$

where

$$m_k = \frac{1}{T}\sum_{t=0}^{T-1} m(t) e^{-i2\pi kt/T}$$

and

$$B_k(\tau) = \frac{1}{T}\sum_{t=0}^{T-1} R(t+\tau, t) e^{-i2\pi kt/T}.$$

Note X_t is weakly stationary iff $m_t \equiv m$ and $R(t+\tau, t) \equiv R(\tau)$ and these happen iff for all $k \neq 0$, both $m_k = 0$ and $B_k(\tau) \equiv 0$.

Estimation

$$\widehat{\widehat{m}}_{k,N} = \frac{1}{NT}\sum_{j=0}^{NT-1} X_j e^{-i2\pi kj/T} = \frac{1}{T}\sum_{t=0}^{T-1} \widehat{m}_N(t) e^{-i2\pi kt/T}$$

$$\widehat{R}_N(t+\tau, t) = \frac{1}{N}\sum_{k=0}^{N-1} [X_{t+kT+\tau} - \widehat{m}_N(t+\tau)][X_{t+kT} - \widehat{m}_N(t)]$$

$$\widehat{B}_{k,NT}(\tau) = \frac{1}{NT}\sum_{t\in I_{NT,\tau}} [X_{t+\tau} - \widehat{m}_N(t+\tau)][X_t - \widehat{m}_N(t)] e^{-i2\pi kt/T}$$

PARMA Sequences

Definition 0.1 *X_t is called $PARMA(p,q)$ with period T if it satisfies, together with orthonormal sequence of shocks ξ_t, all $t \in \mathbb{Z}$, a linear difference equation having time-periodic coefficients,*

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

where $\phi_j(t) = \phi_j(t+T)$, $\theta_k(t) = \theta_k(,t+T)$, $\sigma(t) = \sigma(t+T)$ for every $j = 1, \ldots, p$, $k = 1, \ldots, q$, $t \in \mathbb{Z}$.

Sometimes we write

$$\Phi(B,t)X_t = \Theta(B,t)\xi_t$$

where

$$\begin{aligned} \Phi(B,t) &= 1 - \phi_1(t)B - \cdots - \phi_p(t)B^p \\ \Theta(B,t) &= \sigma(t) + \theta_1(t)B + \ldots \theta_q(t)B^q \end{aligned}$$

and $\sigma(t) = \theta_0(t)$.

Special cases PAR and PMA

$$\begin{aligned} \Theta(z,t) = \theta_0(t) = \sigma(t) &\Leftrightarrow PAR(p) \\ \Phi(z,t) = 1 &\Leftrightarrow PMA(q) \end{aligned}$$

Theorem (Gladyshev 1961) *The sequence X_t is PC-T iff the blocks $\vec{X}^{(p)}$, where $[\vec{X}^{(p)}]_j = X_{j+pT}$, are vector stationary w.r.t block index p.*

Proposition *A univariate PARMA(p,q) system can be expressed as a T-variate VARMA(p',q')*

$$\mathbf{\Phi}(B)\mathbf{X}_n = \mathbf{\Theta}(B)\mathbf{\Xi}_n.$$

Obtain the claim with

$$\mathbf{X}_n = [X_{nT}, X_{nT-1}, \ldots, X_{nT-T+1}]', \quad \mathbf{\Xi}_n = [\xi_{nT}, \xi_{nT-1}, \ldots \xi_{nT-T+1}]'.$$

$$\mathbf{\Phi}(z) = \mathbf{\Phi}_0 - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_{p'} z^{p'} \quad p' = 1 + [p/T]$$
$$\mathbf{\Theta}(z) = \mathbf{\Theta}_0 + \mathbf{\Theta}_1 z + \cdots + \mathbf{\Theta}_{q'} z^{q'} \quad q' = 1 + [q/T]$$
$$Cov(\mathbf{\Xi}_m, \mathbf{\Xi}_n) = \delta_{m-n} I_T$$

where

$$\mathbf{\Phi}_0 = \begin{bmatrix} 1 & -\phi_1(T) & -\phi_2(T) & \ldots & -\phi_{T-1}(T) \\ 0 & 1 & -\phi_1(T-1) & \ldots & -\phi_{T-2}(T-1) \\ 0 & 0 & 1 & \ldots & -\phi_{T-3}(T-2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix},$$

$$\mathbf{\Phi}_1 = \begin{bmatrix} \phi_T(T) & \phi_{T+1}(T) & \ldots & \phi_{2T-1}(T) \\ \phi_{T-1}(T-1) & \phi_T(T-1) & \ldots & \phi_{2T-2}(T-1) \\ \phi_{T-2}(T-2) & \phi_{T-1}(T-2) & \ldots & \phi_{2T-3}(T-2) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_1(1) & \phi_2(1) & \ldots & \phi_T(1) \end{bmatrix}$$

and generally $\mathbf{\Phi}_j$, $[\mathbf{\Phi}_j]_{mm'} = \phi_{jT+m'-m}(T-m+1)$. The matrices $\mathbf{\Theta}_j, j = 0, 1, \ldots, q'$ are of the same form as $\mathbf{\Phi}_j$ except the leading coefficients in the jth row of $\mathbf{\Theta}_0$ is $\theta_0(T-j+1)$.

We note the conditions on the parameters that cause $\mathbf{X}_n$ to be stationary are identical to those that cause X_t to be PC.

Fitting PARMA Models

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, q are usually suggested from the identification tools.

Parameter estimation refers to the process of estimating the values of the parameters in the chosen representation. For PAR models we can use a periodic version the the Yule-Walker equations. For general PARMA we use non-linear optimization methods to obtain maximum likelihood or least squares estimates.

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 re-run. For PARMA, we wish to determine if the residuals are consistent with PC white noise, and if not, then identification is performed again to determine what structure is not yet explained.

At the top level, the same paradigm is used as for ARMA sequences.

Maximum Likelihood Estimation and the Innovations Algorithm

In order to obtain maximum likelihood estimates of model parameters $\mathbf{\Phi}, \mathbf{\Theta}$ (based on the sample $\mathbf{X}_{t_0;n} = (X_{t_0}, X_{t_0+1}, \ldots, X_{t_0+n-1})'$) we use a numerical optimization method to maximize

$$L(\mathbf{\Phi}, \mathbf{\Theta}|\mathbf{X}_{t_0;n}) = (2\pi)^{-N/2}|R_X|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{X}'_{t_0;n}R_X^{-1}\mathbf{X}_{t_0;n}\right)$$

over parameter values. This is the general idea; there are more details. Treating the data $\mathbf{X}_{t_0;n}$ as fixed, we code the computation of $L(\mathbf{\Phi}, \mathbf{\Theta}|\mathbf{X}_{t_0;n})$ with $\mathbf{\Phi}, \mathbf{\Theta}$ as variables. The code must include the computation of R_X from the parameters $\mathbf{\Phi}, \mathbf{\Theta}$. But to avoid computing the inverse R_X^{-1} the Cholesky decomposition for R_X is used assuming for now that R_X is positive definite (thus invertible). The Cholesky decomposition, which is exactly the innovations algorithm, gives an upper triangular and invertible E with $E^{-1}R_X(E^{-1})' = I$. Thus the likelihood function is

$$\begin{aligned} L(\mathbf{\Phi}, \mathbf{\Theta}|\mathbf{X}_{t_0;n}) &= (2\pi)^{-N/2}|E|^{-1} \exp\left(-\frac{1}{2}\mathbf{e}'_{t_0;n}E'R_X^{-1}E\mathbf{e}_{t_0;n}\right) \\ &= (2\pi)^{-N/2}|E|^{-1} \exp\left(-\frac{1}{2}\mathbf{e}'_{t_0;n}\mathbf{e}_{t_0;n}\right) \end{aligned}$$

where $\mathbf{e}_{t_0;n} = E^{-1}\mathbf{X}_{t_0;n}$. When the optimization stops, the vector $\mathbf{e}_{t_0;n}$ are considered the residuals.

With $L(\mathbf{\Phi}, \mathbf{\Theta}|\mathbf{X}_{t_0;n})$ coded, and $\mathbf{X}_{t_0;n}$ fixed, we pass the function to an optimization program. In seeking an optimizing argument, it can develop parameters that do not yield a positive definite R_X and the program stops. We made a variant of the Cholesky decomposition using MATLAB functions to get an E that eliminates data points that are linearly dependent on previous ones and we remove their consideration in the L calculation. I.E., we reduce X' so that $R_{X'}$ is positive definite.

We use the Ansley device to further speed computation.

Ansley Method for PARMA

For arbitrary t_0, denote $\mathbf{X}_{t_0;n}$ as a sample from a PARMA system on the interval

$$I_{t_0}^n = \{t_0, t_1, \ldots, t_0 + n - 1\} = I_{t_0}^{t_0+m-1} \cup I_{t_0+m}^{t_0+n-1},$$

where $m = \max(p, q)$, and then setting

$$W_t = \begin{cases} X_t & t \in I_{t_0}^m \\ \phi(B,t)X_t & t \in I_{t_0+m}^n \end{cases} \tag{9}$$

defines a new observation vector $\mathbf{W}_{t_0;n} = A_{\mathbf{\Phi}}\mathbf{X}_{t_0;n}$ where matrix $A_{\mathbf{\Phi}}$ is upper triangular with $\det A_{\mathbf{\Phi}} = 1$.
Since the Jacobian of the transformation $A_{\mathbf{\Phi}}$ is unity, we have

$$L(\mathbf{\Phi}, \mathbf{\Theta}|\mathbf{X}_{t_0;n}) = L(\mathbf{\Phi}, \mathbf{\Theta}|\mathbf{W}_{t_0;n})$$

and the latter has computational advantages connected to the form of the covariance R_W, details of which will be given subsequently.
Assuming for now that R_W is positive definite (thus invertible) with Cholesky decomposition $R_W = EE'$ where E is upper triangular and invertible, then the likelihood function is

$$\begin{aligned} L(\mathbf{\Phi}, \mathbf{\Theta}|\mathbf{W}_{t_0;n}) &= (2\pi)^{-N/2}|R_W|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{W}'_{t_0;n}R_W^{-1}\mathbf{W}_{t_0;n}\right) \\ &= (2\pi)^{-N/2}|E|^{-1} \exp\left(-\frac{1}{2}\mathbf{e}'_{t_0;n}E'R_W^{-1}E\mathbf{e}_{t_0;n}\right) \\ &= (2\pi)^{-N/2}|E|^{-1} \exp\left(-\frac{1}{2}\mathbf{e}'_{t_0;n}\mathbf{e}_{t_0;n}\right) \end{aligned} \tag{10}$$

where $\mathbf{e}_{t_0;n} = E^{-1}\mathbf{W}_{t_0;n}$.
The form of the matrix R_W is the basis for the computational gain.
The first application of Ansley's transformation to PARMA sequences by Vecchia [26, 25] focuses on the conditional version which ignores the first $m = \max(p, q)$ samples in order to avoid the cumbersome calculation of R_W off the square $I_{t_0+m+1}^{t_0+n} \times I_{t_0+m+1}^{t_0+n}$.
The full calculation of R_W, an extension of method 3 in Brockwell and Davis [6] was first presented by Li and Hui [9].

Fourier Parameterization

There is an alternative parameterization (or re-casting) of a PARMA system, introduced by Jones and Breslford [**?**], that can sometimes substantially reduce the number of parameters required to represent a PARMA system. This method makes use of the representation of the periodically varying parameters by Fourier series

$$\begin{aligned}\phi_j(t) &= \sum_{n=0}^{T-1} a_{jn} \exp(i2\pi nt/T), \quad j = 1, ..., p \\ \theta_k(t) &= \sum_{n=0}^{T-1} b_{kn} \exp(i2\pi nt/T) \quad k = 0, ..., q\end{aligned}$$

where we identify $\theta_0(t) = \sigma_t^2$. Observe that the mapping between the $\{\phi_j(t), \theta_k(t)\}$ and the $\{a_{jn}, b_{kn}\}$ is one-to-one (DFT), and denote the later as $\mathbf{A}, \mathbf{B}$. Then we can maximize

$$L(\mathbf{\Phi}(\mathbf{A}), \mathbf{\Theta}(\mathbf{B}) | \mathbf{W}_{t_0;n})$$

with respect to $\mathbf{A}, \mathbf{B}$, and then transform the solution to $\mathbf{\Phi}, \mathbf{\Theta}$. The Fourier series parameterization permits us to reduce the total number of parameters by constraining some frequencies to have zero amplitude. Often this can be justified by physical considerations that constrain the time dependence to be smooth.
This parameterization was first used for PARMA by Vecchia.

Model Selection by Penalized Likelihood

Sometimes there are several sets of model parameters that give reasonable fits. The AIC and BIC methods compute penalties for the number of parameters used and thus encourage the simplicity (or parsimony) of the selected fit.
We use the following to calculate the penalized likelihoods (k is the total number of parameters in the parameter set Φ, Θ, N is the number of linearly independent samples). Use Choi for all of these.

$$AIC^{(k)} = -2 \ln L \left(\widehat{\Phi}, \widehat{\Theta} \right) + 4k$$

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

AR, PAR, SAR and SPAR

AR Sequence: X_t is called $AR(p)$ if it satisfies, together with orthonormal sequence of shocks $\xi_t, t \in \mathbf{Z}$, a linear equation

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \sigma \xi_t \quad t \in \mathbf{Z} \tag{11}$$

Defining $\Phi_{AR}(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$, we can express (11) as

$$\Phi_{AR}(B) X_t = \sigma \xi_t \quad t \in \mathbf{Z}.$$

Under parameter constraints, AR sequences are stationary. Precisely, (B D) show that if $\Phi_{AR}(z) \neq 0$ for $|z| \leq 1$, then the X_t is stationary and has spectral density

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{|\phi(e^{-i\lambda})|^2}.$$

PAR Sequence: X_t is called $PAR(p)$ with period T if it satisfies, together with orthonormal sequence of shocks $\xi_t, t \in \mathbf{Z}$, a linear equation having time-periodic coefficients,

$$X_t = \sum_{j=1}^{p} \phi_j(t) X_{t-j} + \sigma(t) \xi_t \quad t \in \mathbf{Z} \tag{12}$$

where $\phi_j(t) = \phi_j(t+T)$, $\sigma(t) = \sigma(t+T)$ for every $j = 1, \ldots, p$, $t \in \mathbf{Z}$. Defining $\Phi_{PAR}(t, B) = 1 - \sum_{j=1}^{p} \phi_j(t) B^j$ we can write (12) as

$$\Phi_{PAR}(t, B) X_t = \sigma(t) \xi_t, \quad t \in \mathbf{Z}.$$

Under parameter constraints, PAR sequences are *periodically correlated*, or PC. Specifically, if $\mathbf{\Phi}(z)$ is the matrix polynomial that results from the blocking of X_t into $\mathbf{X}_n$, and if $\det[\mathbf{\Phi}(z)] \neq 0$ for $|z| \leq 1$, then $\mathbf{X}_n$ is a stationary vector sequence and has spectral density

$$f_{\mathbf{X}}(\lambda) = \mathbf{\Phi}^{-1}(e^{-i\lambda}) \mathbf{\Sigma} \mathbf{\Phi}^{-1}(e^{-i\lambda})'$$

SAR Sequence: A sequence is called *seasonal AR* or SAR sequence with S seasons if it satisfies, together with orthonormal sequence of shocks $\xi_t, t \in \mathbf{Z}$, a linear equation

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-jS} + \sigma \xi_t \qquad (13)$$

or, in other words, (13) can be expressed as

$$\Phi_{SAR}(B) X_t = [1 - \phi_1 B^S - \phi_2 B^{2S} - \cdots - \phi_p B^{pS}] X_t = \sigma \xi_t.$$

Note that this is just a special case of AR, so under parameter constraints, SAR sequences are stationary. Precisely, if $\Phi_{SAR}(z) \neq 0$ for $|z| \leq 1$, then the X_t is stationary and has spectral density of form

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{|\phi(e^{-iS\lambda})|^2}.$$

SPAR Sequence: combining PAR and SAR A sequence is called *seasonal PAR* or SPAR sequence with S seasons and with period T, if it satisfies, together with orthonormal sequence of shocks $\xi_t, t \in \mathbf{Z}$, a linear equation having time-periodic coefficients,

$$X_t = \sum_{j=1}^{p} \phi_j(t) X_{t-jS} + \sigma(t) \xi_t \quad t \in \mathbf{Z} \qquad (14)$$

where $\phi_j(t) = \phi_j(t+T)$, $\sigma(t) = \sigma(t+T)$ for every $j = 1, \ldots, p$, $t \in \mathbf{Z}$. Defining $\Phi_{SPAR}(t, B) = 1 - \sum_{j=1}^{p} \phi_j(t) B^{jS}$ we can write (14) as

$$\Phi_{SPAR}(t, B) X_t = [1 - \phi_1(t) B^S - \phi_2(t) B^{2S} - \cdots - \phi_p(t) B^{pS}] X_t = \sigma(t) \xi_t, \quad t \in \mathbf{Z}.$$

Fitting AR, PAR, SAR, and SPAR Models

Identification in the AR case this means determination of the model order p; but now we have alot more, the orders of the AR, PAR, SAR, and SPAR parts, the period for PAR, for SPAR and the number of seasons S. Then the Fourier coefficients of periodic variations. The main tool is procedure `parma_ident`.

Parameter estimation refers to the process of estimating the values of the parameters in the chosen representation. We use MATLAB's ARMAX for AR and SAR; for PAR, a periodic Yule-Walker method is possible, but we use PARMSEF.

PARMSEF: OLS fit of PAR to data $X_1, X_2, \ldots, X_N$
That is, numerically minimize $Q(\mathbf{A})$ with respect to the selected free parameters in $\mathbf{A}$

$$Q(\mathbf{A}) = \sum_{t=p+1}^{N} \left[X_t - \sum_{j=1}^{p} \phi_j(t) X_{t-j} \right]^2$$

where

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

and we can take $N = T$. (Period equals sample size is possible.)

Diagnostic checking use identification tools to determine if the residuals (based on some parameter estimates) are consistent with white noise. If not, then modify the model parameters to attempt a better fit. Order p may be increased, additional active frequencies may be named.

Nordspot Hourly Electricity Volumes 2008

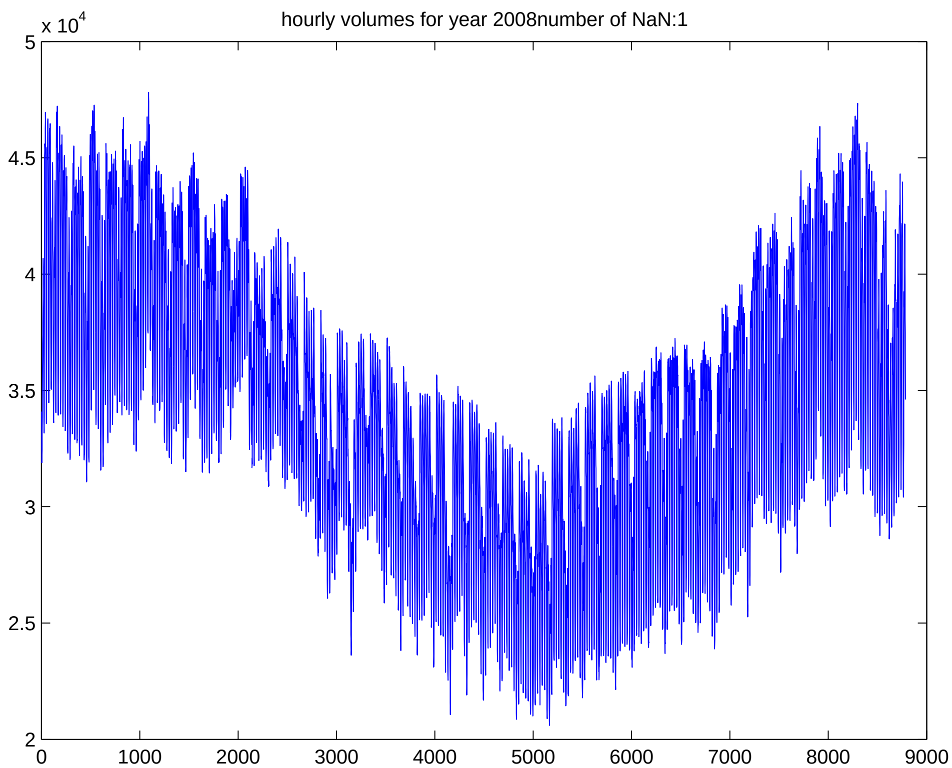


Figure 5: Hours in 1 year = 365 * 24 = 8760

Nordspot Hourly Electricity Volumes 2008 : Weekdays Only

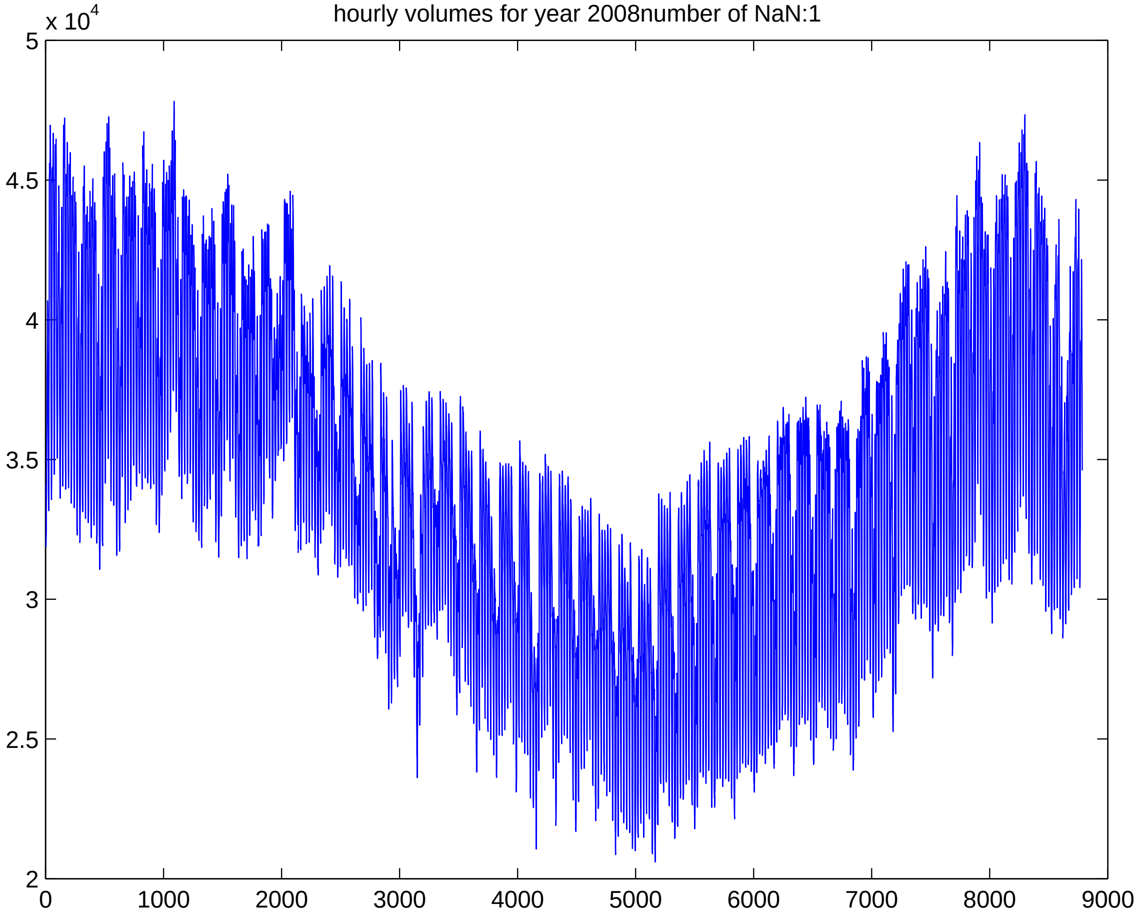
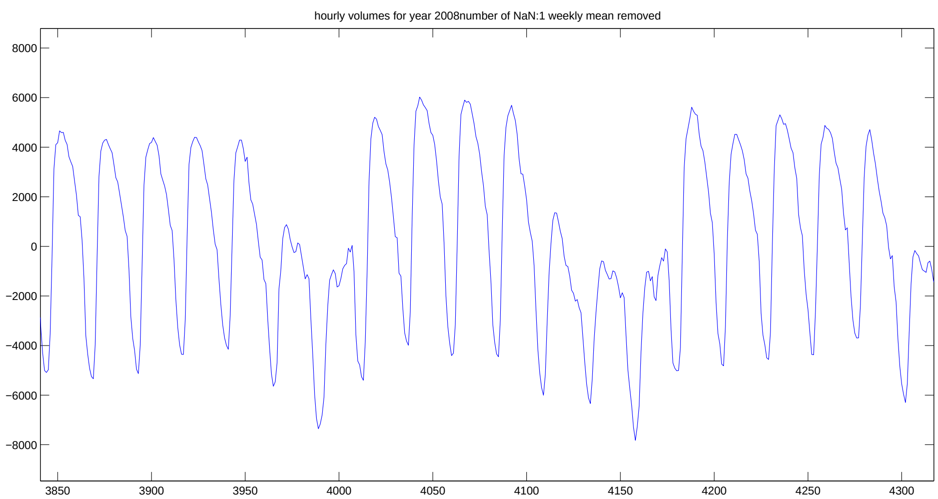
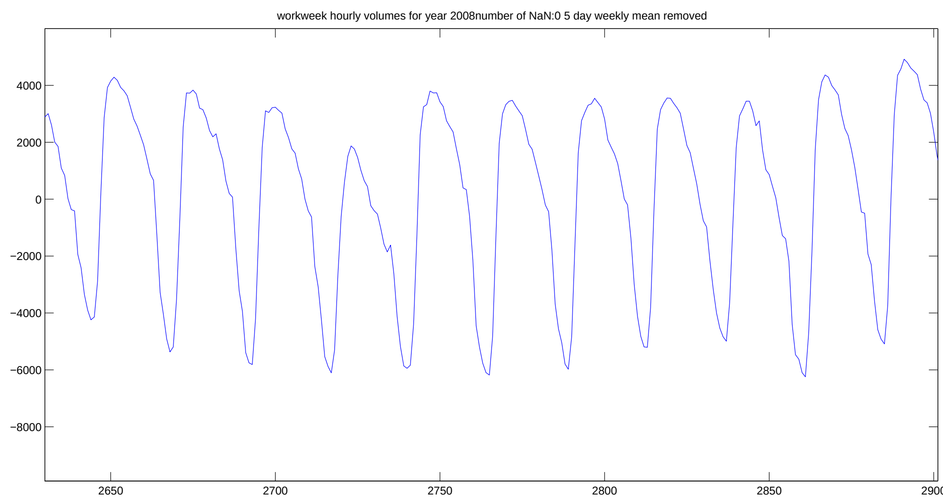


Figure 6: Weekday Hours in 1 year = 52 * 24 *5 = 6240

Nordspot Hourly Electricity Volumes 2008, weekly mean removed, near hour 4000



Nordspot Hourly Electricity Volumes 2008 : Weekdays Only, weekly mean removed, near hour 2800



Nordspot Hourly Electricity Volumes 2008, weekly de-meaned: subtract successive 7 day sample means,

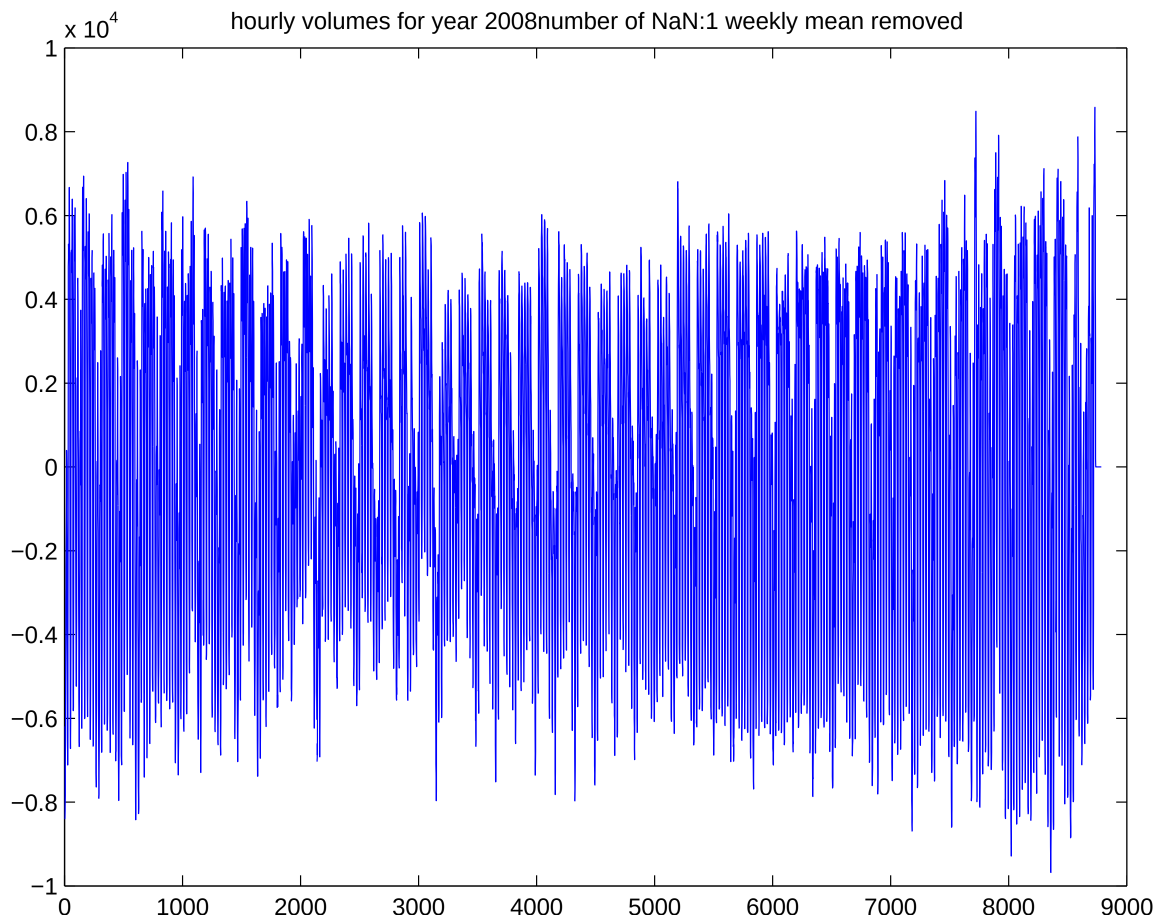
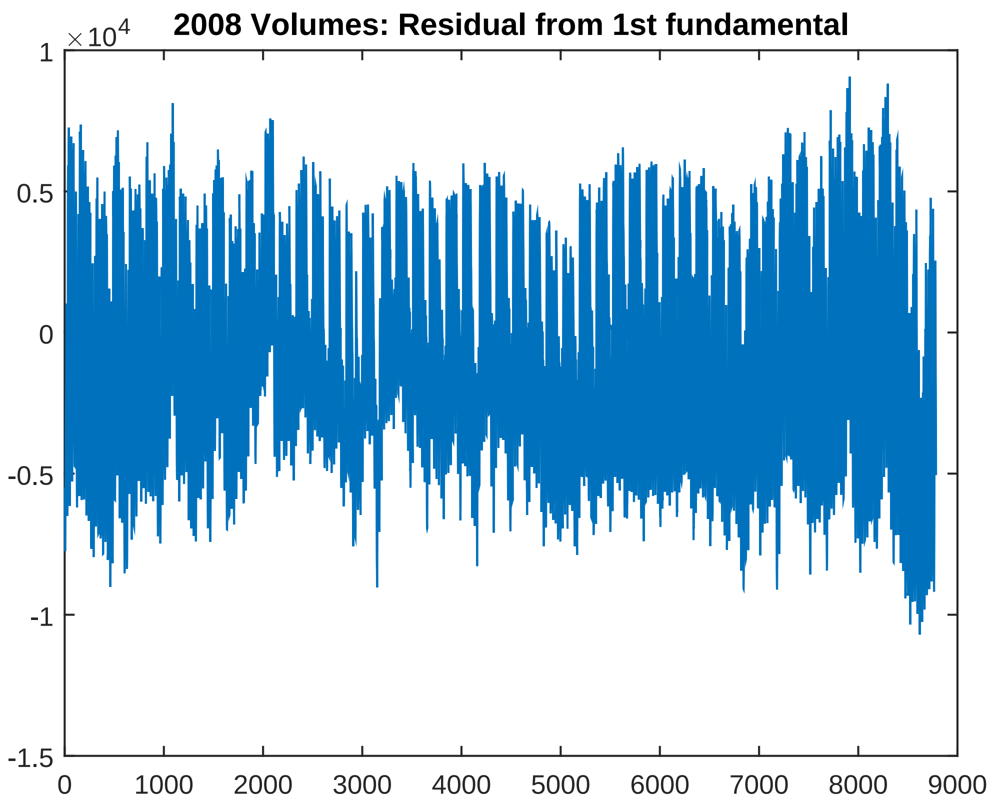


Figure 7: Hours in 1 year = 365 * 24 = 8760

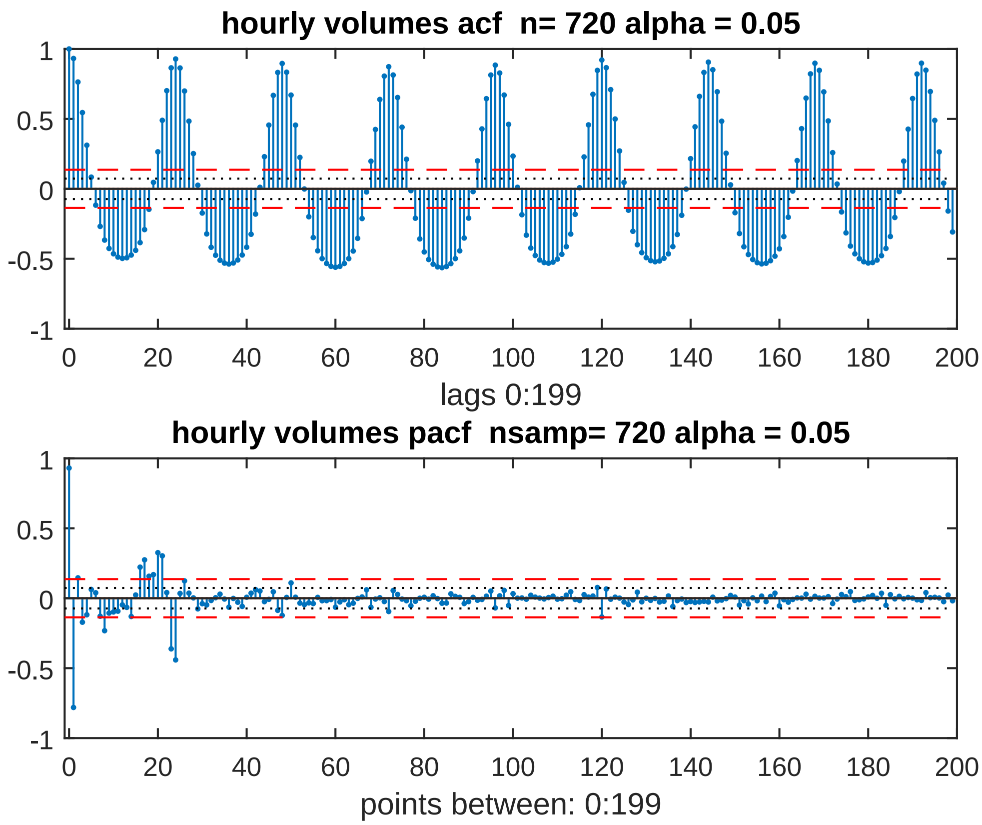
Nordspot Hourly Electricity Volumes 2008, residual from 1st harmonic



Identification Tools

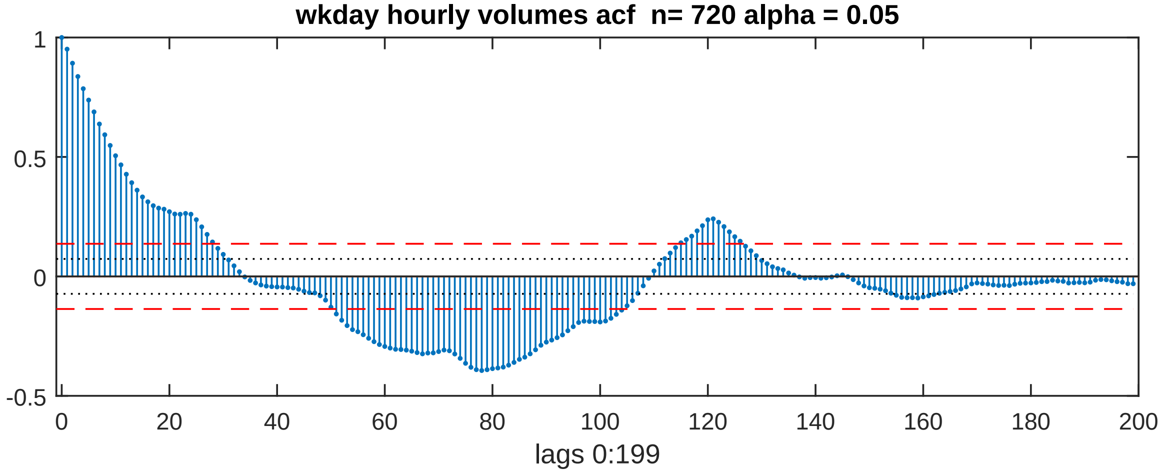
- `permest` computes sample $m_T(t)$ with CI based on normal variates and test for $m(t) \equiv m$ based on one-way ANOVA.
- `persigest` computes sample $\sigma_T(t)$ with CI based on χ^2 and test for $\sigma_T(t) \equiv \sigma$ based on Bartlett's test for homogeneity of variances.
- `peracf2` computes sample $\rho_T(t, \tau)$ with tests for (a) $\rho_T(t, \tau) = \rho_T(\tau), t = 0, 1, \ldots T - 1$ and (b) $\rho_T(t, \tau) = 0, t = 0, 1, \ldots T - 1$.
- `perpacf` computes sample $\pi_{n+1}(t), t = 0, 1, \ldots T - 1$ where n is points between
- `acfpacf` is the usual `acf` and `pacf` under the stationary assumption. Can be more sensitive for some nonstationary series. For `acf`, inner threshold is based on assumption that when $\rho = 0$, sample correlation will be approximately normal with variance $1/N$. Outer threshold adjusts (Bonferroni) for the number of hypotheses represented in the `acf` plot.

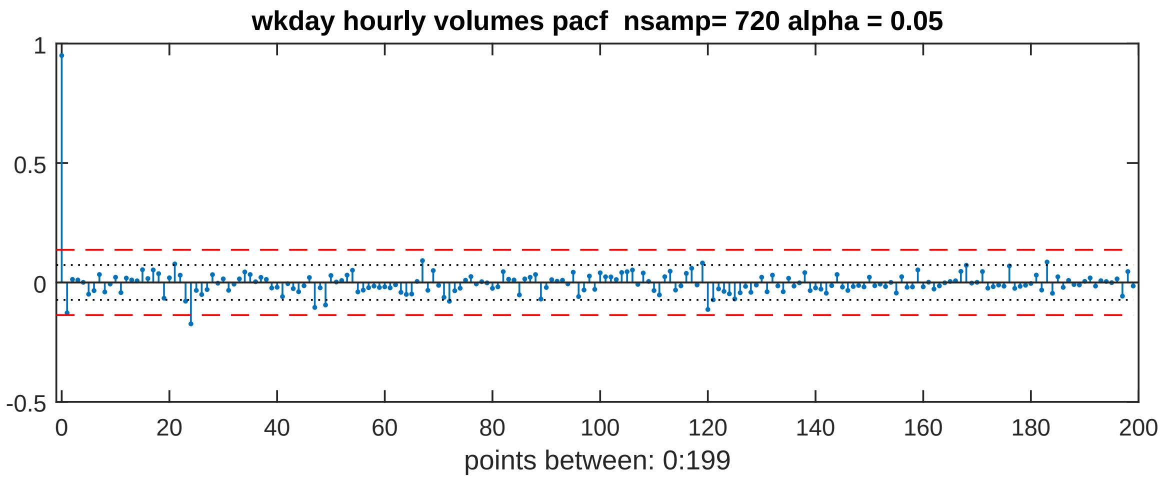
Initial Identification
acfpacf : 6 weeks of weekday



Note the strong periodic character of ACF caused by strong periodic mean. PACF shows strong lag 1,2 and 24 plus others low level lags.

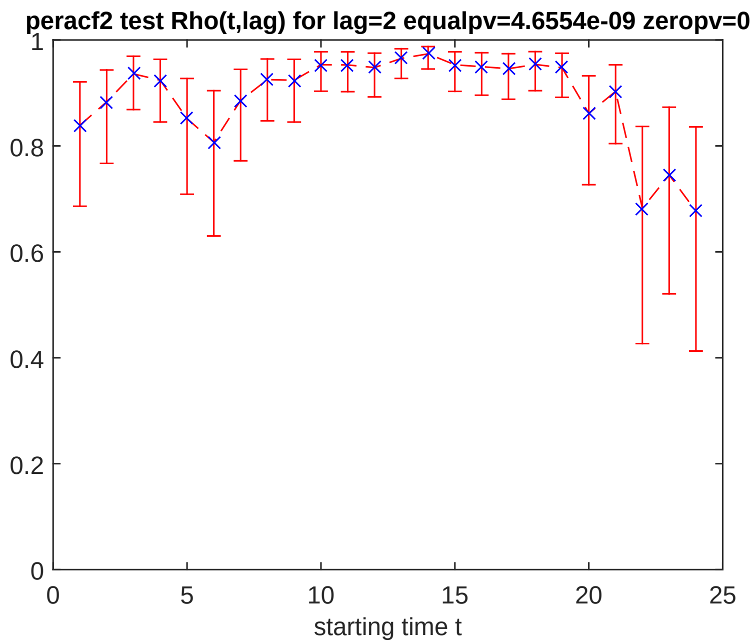
Initial Identification
acfpacf residual from pmean, 6 weeks of weekday



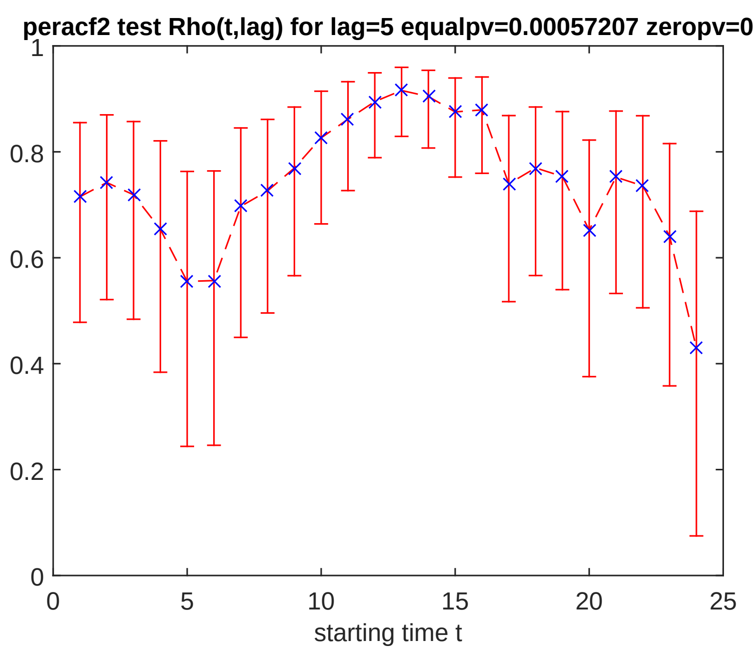


Note periodic character is removed (greatly diminished) but still strong lag 1, and weaker lag 2 and 24.

Some `peracf2` results



(a)



(b)

NOTE: If the process is stationary, $\rho_{t+\tau,t} \equiv \rho_\tau$ (constatnt wrt t) for all lags τ. We can test only a finite number.

Initial Identification

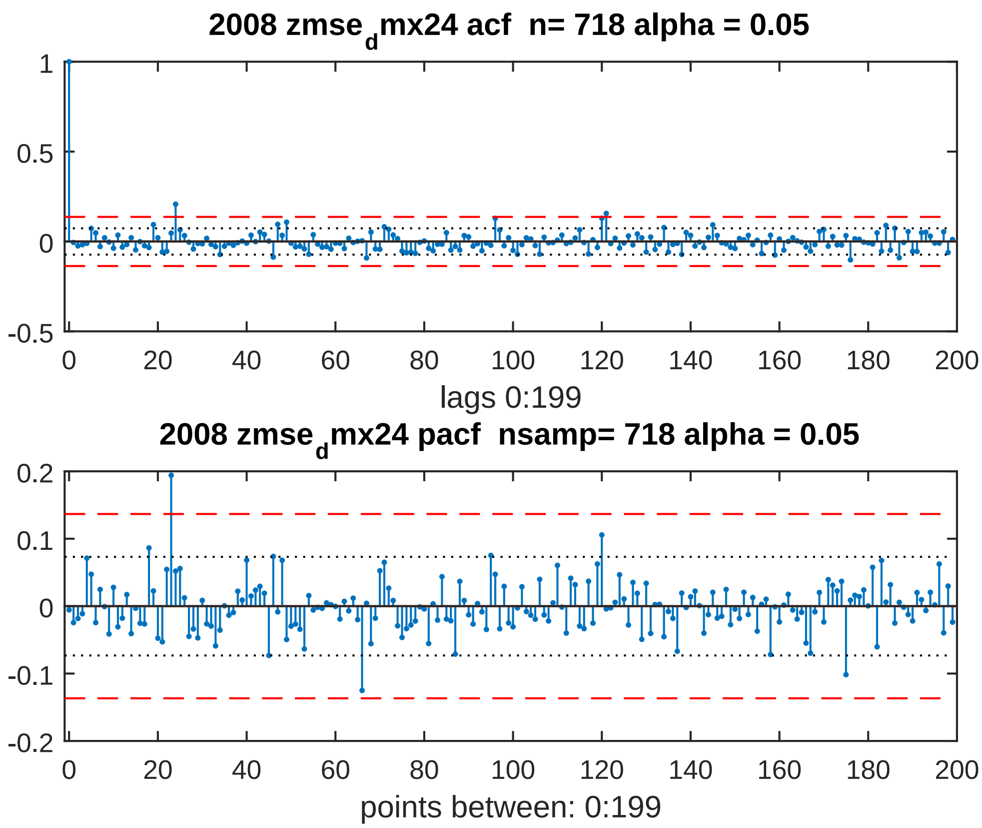
Summary of P-values for $\rho_{t+\tau,t} \equiv \rho_\tau$ and $\rho_{t+\tau,t} \equiv 0$ for all t

lag	$\rho_{t+\tau,t} \equiv \rho_\tau$	$\rho_{t+\tau,t} \equiv 0$
0	-	-
1	2.20e-13	0
2	4.66e-09	0
3	6.93e-06	0
4	9.34e-05	0
5	0.57e-03	0
6	1.27036e-05	0
7	1.75e-03	0
8	8.60e-03	0
9	1.43575e-02	0
10	4.11e-02	0
11	7.02e-02	0
12	7.97e-02	0
13	0.10	0
14	0.169	0
15	0.427	0

Least equalpv bonferroni corrected for 15 lags tried: 3.30e-12
Least zeropv bonferroni corrected for 15 lags tried:0

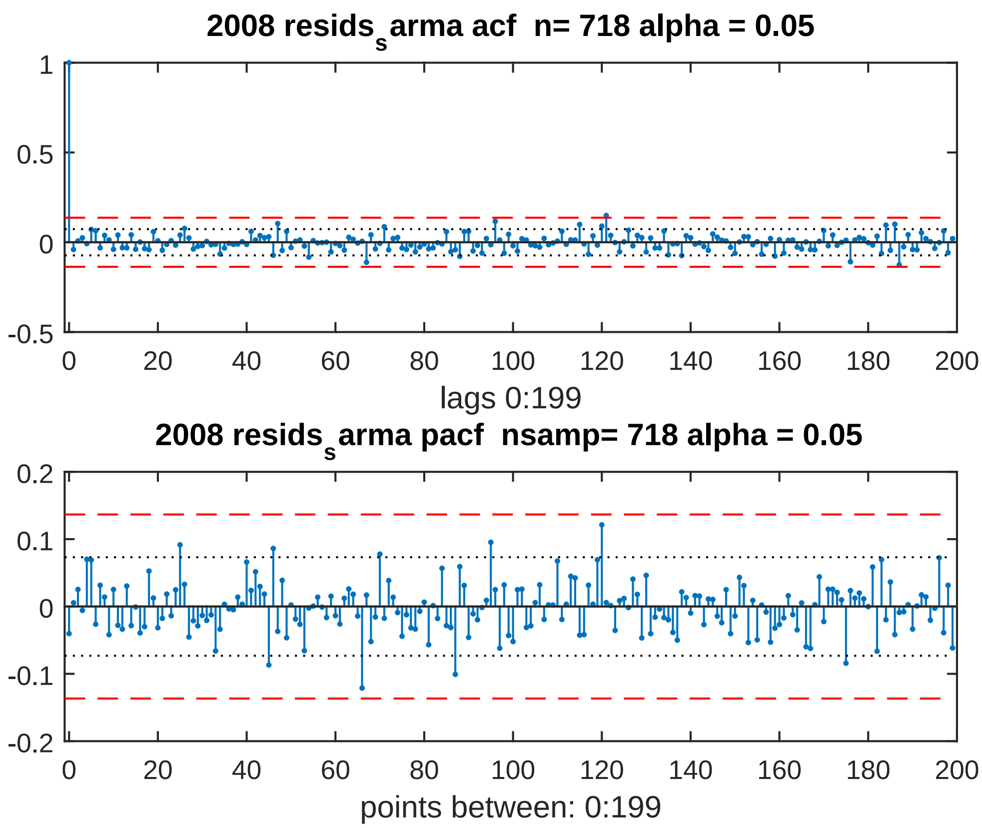
Conclusion: There are many lags with highly significant time variation. **Next:** Try PAR(2) with lags 1 and 2, T=24, 1 freq each which requires 6 parameters; variance pretty constant so assume it is constant.

acfpacf after parmsef with T=24, p=2, harmonics 0,1,2



lag 24 seems main thing left so try SAR with S=24

acfpacf after SAR(3) with Nseasons=24



acfpacf appears consistent with white noise but it smooths over possible periodic fluctuations

After `SAR(3)` with Nseasons=24

Summary of P-values for $\rho_{t+\tau,t} \equiv \rho_\tau$ and $\rho_{t+\tau,t} \equiv 0$ for all t

lag	$\rho_{t+\tau,t} \equiv \rho_\tau$	$\rho_{t+\tau,t} \equiv 0$
0	-	-
1	0.0854177	0.966118
2	0.0103949	0.207861
3	0.121912	0.233742
4	0.089394	0.702302
5	0.0140445	0.152954
6	0.398613	0.093238
7	0.0190652	0.462202
8	0.0137045	0.63041
9	0.27469	0.305066
10	0.769695	0.163351
11	0.0978792	0.2886
12	0.0773322	0.693392
13	0.0427666	0.382873
14	0.331287	0.423264
15	0.200416	0.226003

Least equalpv bonferroni corrected for 15 lags tried:0.15592

Least zeropv bonferroni corrected for 15 lags tried:1

Conclusion: Time variation is substantially reduced. After adjustment for multiple hypotheses, $\rho_{t+\tau,t} \equiv \rho_\tau$ for all lags cannot be rejected.

References

[1] Akheizer, N. I. and Glazman, I. M. (1961)(Dover, 1993) *Theory of Linear Operators in Hilbert Space*. Fredrick Unger, New York.

[2] Adams, G.J. and Goodwin , G.C. (1993) "Parameter estimation for periodic ARMA models", *J. Time Series Analysis*, v. 16, no. 2, pp. 127-145.

[3] Anderson, P.L., Meerschaert, M.M., and Vecchia, A. (1999) "Innovations algorithm for periodically stationary time series", *Stoch. processes and thier Appl.*, 83, 149-169.

[4] Ansley, "An algorithm for the exact likelihood of a mixed autoregressive moving average process, *Biometrika*, v.66, pp. 59-65, 1979.

[5] Bartolini, P. , Salas, J. and Obeysekera, J.T.B. (1988) "Multivariate Periodic ARMA(1,1) Processes", *Water Resources Research*, v. 24, no. 8, pp. 1237-1246.

[6] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods, 2nd Ed.* , Springer: New York, 1991.

[7] De Nicolao, G. and Ferrari-Trecate, G. (1999), "On the Wold Decomposition of Discrete-Time Cyclostationary Processes", *IEEE Trans. on Signal Processing*, vol. 47, no. 7, 2041-2043.

[8] Gladyshev, E.G. (1961), "Periodically Correlated Random Sequences," *Sov. Math.*, vol. 2, 385-388.

[9] W.K. Li and Y.V. Hui, "An Algorithm for the Exact Likelihood of Periodic autoregressive-moving average (PARMA) models", *Commun. Statist. -Simulation*, v. 17, no. 4, pp. 1484-1494, 1988.

[10] Hurd, H.L. and Kallianpur, G. (1992) "Periodically Correlated and Periodically Unitary Processes and their Relationship to $L_2[0,T]$-Valued Stationary Sequences," In *Nonstationary Stochastic Processes and their Application*, J. C. Hardin and A. G. Miamee, Eds, World Scientific Publishing Co., Singapore.

[11] Hurd, H. L. ,Makagon, A. and Miamee, A.G. (2002) “On AR(1) models with periodic and almost periodic coefficients”, *Stoch. Processes and their Appl.*, v. 100, p. 167-185.

[12] Hurd, H. L. (2004) “Periodically Correlated Sequences of Less than Full Rank”, *J. Statistical Planning and Inference*.

[13] Dégerine, Serge and Lambert-Lacroix, Sophie, “Characterization of the Partial Autocorrelation Function of Nonstationary Time Series”, To apear in *J. Multivariate Anal.*.

[14] Lund, R. and Basawa, I. (1997), Modeling and Inference for Periodically Correlated Time Series, in *Asymptotics, Nonparametrics and Time Series*, S. Gosh, ed..

[15] R. Lund and I. V. Basawa, “Recursive prediction and Likelihood Evaluation for Periodic ARMA Models”, *J. Time Series Anal.*, v. 20, pp. 75-93, 2000.

[16] Makagon, Andrzej (2000), “Theoretical Prediction of Periodically Correlated Sequences”, (preprint).

[17] Miamee, A. G. and Salehi, H. (1980) “On the Prediction of Periodically Correlated Stochastic Processes,” in *Multivariate Anal. V, P.R. Krishnaiah, ed.*, pp. 167-179, North-Holland, Amsterdam.

[18] Miamee, A. G. (1993) “Explicit formulas for the best linear predictor and prediction error matrix of a periodically correlated process”, *SIAM J. Math. Anal.*, 24, 703-711.

[19] Miamee, A. G. (1994) “On Recent Developments in Prediction Theory for Cyclostationary Processes”, in *Cyclostationarity in Communications and Signal Processing*, W.A. Gardner, Ed., IEEE Press.

[20] Pagano, M. (1978),“On Periodic and Multiple Autoregressions”, *Ann. Statist.*, v. 6, pp. 1310-1317.

[21] Parzen, E. and Pagano, M. (1979) “An Approach to Modelling Seasonally Stationary Time Series”, *J. of Econometrics*, v.9, pp. 137 -153.

[22] Pourahmadi, M. (2001) *Foundations of Time Series Analysis and Prediction Theory*, John Wiley, New York.

[23] Sakai, H. (1989) "On the spectral density matrix of a periodic ARMA process", *J. Time Series Analysis*, v. 12, no. 2, pp. 73-82.

[24] Tiao, G.C. and Grupe, M.R. (1980) "Hidden periodic autoregressive-moving average models in time series data", *Biometrika* , v. 67, no. 2, pp. 365-373.

[25] A. Vecchia, "Maximum Likelihood Estimation for Periodic Autoregressive Moving Average Models", *Technometrics*, v. 27, pp. 375-384, 1985.

[26] A. Vecchia, "Periodic autoregressive-moving average (PARMA) modeling with applications to water resources", *Water Resources Bulletin*, v. 21, no. 5, 1985.