PARMA methods based on Fourier Representation of Periodic Coefficients
Anna E. Dudek, Harry Hurd, Wioletta Wojtowicz

To cite this version:
<hal-01310918v1>

HAL Id: hal-01310918
https://hal.archives-ouvertes.fr/hal-01310918v1
Submitted on 25 May 2016 (v1), last revised 1 Jan 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
PARMA methods based on Fourier Representation of Periodic Coefficients

Anna Dudek
Institut de Recherche
Mathématique de Rennes,
CNRS umr 6625,
Université Rennes 2, Rennes, France

Harry Hurd
Department of Statistics and Operations Research
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3260

Wioletta Wójtowicz
Cracow University of Technology,
Institute of Applied Computer Science,
al. Jana Pawła II 37,
31-864 Cracow, Poland

Keywords
Periodically Correlated (PC) sequences, periodic ARMA models (PARMA models), Fourier-PARMA models, time series analysis

Abstract
In this paper we review the methods of solving PARMA systems when the periodic coefficients are represented by Fourier series. This method, introduced by Jones and Brelsford (1967), remains a method employed in practice yet still offers problem areas for future research. The primary benefit of the method is that for PARMA systems in which the periodic variations are smooth within the fundamental period, a substantial reduction in the number of estimated parameters may be realized by forcing many of the Fourier coefficients to be zero, hence restricting estimated solutions to a subspace. While reviewing the development of Fourier PARMA methods we naturally view many of the main advances in PARMA time series analysis under the usual parameterization. Two simulations are presented that demonstrate further potential and open problems associated with the methods.

There exist many natural random processes in which the probability structure has a periodic rhythm, which, in the strict sense means that the probability law
is invariant under shifts of length $\nu$. To be precise, a process $X_t(\omega) : \omega \rightarrow C$ or $R$ is called periodically stationary with period $\nu$ if for every $n$, collection of times $t_1, t_2, ..., t_n$ in $Z$ or $R$, collection of Borel sets $A_1, A_2, ..., A_n$ of $C$ or $R$,

$$
Pr[X_{t_1+\nu} \in A_1, X_{t_2+\nu} \in A_2, ..., X_{t_n+\nu} \in A_n] = Pr[X_{t_1} \in A_1, X_{t_2} \in A_2, ..., X_{t_n} \in A_n]
$$

(1)

and there are no smaller values of shift $\nu > 0$ for which (1) holds. Synonyms for periodically stationary include periodically non-stationary, cyclostationary (think of cyclically stationary), processes with periodic structure, and a few others. If $\nu = 1$, the process is strictly stationary. When the process is of second order, $X_t \in L_2(\omega, F, P)$ with $t \in Z$, it is called periodically correlated (or wide-sense cyclostationary) with period $\nu$ if

$$
m(t) = E\{X_t\} = m(t+\nu)\forall t, \quad (2)
$$

and

$$
R(s,t) = E\{X_sX_t\} = R(s+\nu, t+\nu)\forall s,t \in Z
$$

(3)

and there are no smaller values of $\nu > 0$ for which (2) and (3) hold. If $\nu = 1$, the process is weakly (or wide-sense) stationary. In many time series papers, $X_t$ is taken to be real valued and the correlation denoted as $\gamma_t(u) = R(t,t+u)$. Just as a broad class of stationary time series (or stochastic process) models may be realized by the autoregressive moving average (ARMA) models, there is a corresponding class, called periodic ARMA (or PARMA) which may be viewed as ARMA models having periodic coefficients. Precisely, a second order stochastic sequence $X_t$ is called PARMA(p,q) with period $\nu$ if it satisfies

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

(4)

where $\xi_t$ is an orthogonal sequence and $\phi_j(t) = \phi_j(t+\nu)$, $\theta_k(t) = \theta_k(t+\nu)$ and $\sigma(t) = \sigma(t+\nu)$ for every appropriate $j, k, t$. Sometimes we write $\theta_0(t) = \sigma(t)$. Just as ARMA sequences can be stationary under some constraints of the parameters, likewise PARMA sequences can be PC (periodically correlated) under constraints of their parameters. To obtain these conditions we first use the fact first noted by Gladyshev [16] that under the blocking of $X_t$ into vectors $X_n$ of length $\nu$,

$$
[X_n]_{j} = X_{j+\nu(n-1)},
$$

(5)

the condition for $X_t$ to be PC is identical to the condition for the vector sequence $X_n$ to be stationary. In the PARMA case, conditions for stationarity of $X_n$ may be derived from the parameters by noting that (4) becomes (see [45])

$$
LX_n - \sum_{j=1}^{p} U_jX_{n-j} = \Gamma \varepsilon_n - \sum_{k=1}^{q} V_k\varepsilon_{n-k},
$$

(6)
where

\[
L = \begin{bmatrix}
1 & 0 & 0 & \ldots & 0 \\
-\phi_1(1) & 1 & 0 & \ldots & 0 \\
-\phi_2(2) & -\phi_1(2) & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\phi_{\nu-1}(\nu-1) & -\phi_{\nu-2}(\nu-1) & -\phi_{\nu-3}(\nu-1) & \ldots & 1
\end{bmatrix}
\]

and

\[
[U_j]_{nn'} = \phi_{j+n-n'}(n),
\]

and similarly for \( \Gamma \) and \( V_k \). From Fuller, 1976, ch. 2 or (Hannan), \( X_n \) is stationary if and only if

\[
\det \left[ L - \sum_{j=1}^{p'} U_j \lambda^j \right] \neq 0, \text{ for } |\lambda| \leq 1. \quad (7)
\]

The condition (7) was expressed first by Pagano [35] for PAR, and then by Vecchia [45] for general PARMA; in the latter case, controlling the PAR part is sufficient to control explosive behavior of \( X_n \).

In the case of PAR, the vector sequence \( X_n \), could also be modeled by a vector AR, (VAR) model, but we note that the number of real autoregressive parameters for a VAR(p) is on the order of \( p\nu^2 \) because the autoregressive coefficients are \( \nu \times \nu \) matrices. But for PAR(p) the number is on the order of \( p\nu \), which can still be sizable when compared to the total length of the series available. See Pagano [35] p. 1316. For a full PARMA given by (4) the parameter count is seen to be \((p+q+1)\nu\). But an alternative parameterization of a PARMA system (suggested for PAR by Jones and Breslford [24]) can often substantially reduce the number of parameters by representing the periodically varying parameters by Fourier series

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

\[
\theta_k(t) = b_{k,1} + \sum_{n=1}^{\lfloor \nu/2 \rfloor} b_{k,2n} \cos(2\pi nt/\nu) + b_{k,2n+1} \sin(2\pi nt/\nu)
\]

for \( t = 0, 1, \ldots, \nu-1, j = 1, \ldots, p, k = 0, 1, \ldots, q \); the further reduction in parameters occurs when the number of nonzero Fourier coefficients in (8) is small. In the preceding and subsequently, \( \lfloor \cdot \rfloor \) denotes the floor function. The inverse for
the $a_{j,n}$ coefficients is given by

\begin{align*}
  a_{j,1} &= \frac{1}{\nu} \sum_{t=0}^{\nu - 1} \phi_j(t) \\
  a_{j,2n} &= \frac{2}{\nu} \sum_{t=0}^{\nu - 1} \phi_j(t) \cos\left(\frac{2\pi nt}{\nu}\right) \\
  a_{j,2n+1} &= \frac{2}{\nu} \sum_{t=0}^{\nu - 1} \phi_j(t) \sin\left(\frac{2\pi nt}{\nu}\right)
\end{align*}

for $n = 2, \ldots, \lfloor \nu/2 \rfloor, j = 1, \ldots, p$ and similarly for the $b_{k,2n}$.

This paper is primarily a review and summary of the existing work that has addressed the time series analysis of PARMA models when the periodic parameters are expressed as Fourier series. However, we naturally view many of the main advances in PARMA time series analysis under the usual parameterization. We include two simulations that demonstrate further potential and open problems associated with the methods.

**Chronological Review**

We note that contemporaneously with the papers reviewed here, other work was appearing on spectral theory, representations and non parametric time series analysis for periodically correlated processes; see [23] for a summary. For a thorough review of research in both statistics and engineering, where the term *cyclostationary* is often used synonymously with *periodically correlated*, see the review of Gardner, Napolitano and Paura [15].

In the following, we will adopt the notation used in many of the papers on time series analysis for PC sequences, namely we consider $X_t$ to be real valued and use the notation

$$
\rho_t(u) = \text{corr}\{X_t, X_{t+u}\} = \gamma_t(u) / \sigma(t) \sigma(t+u)
$$

where $\gamma_t(u) = \text{Cov}\{X_t, X_{t+u}\} = R(t, t+u)$. ([47, 34]). Thus with $\nu$ designated as the period, (3) becomes $\gamma_{t+\nu}(u) = \gamma_t(u)$ for all $t, u$. In the case of non-zero mean, we denote $m(t) = E\{X_t\}$. We shall use $t$ for time, $u$ for lag, $\nu$ for period, $j$ for indices of the PAR part of (4) and $k$ for indices of PMA part. We will say that $X_t$ is causal if, for every $t$, it can be written as an infinite moving average (IMA)

$$
\sum_{k=0}^{\infty} \psi_k(t) \epsilon_{t-k} \text{ where } \sum_{k=0}^{\infty} |\psi_k(t)| < \infty,
$$

(10)

with $\text{Var}[\epsilon_t] = 1$ and $\psi_k(t) = \psi_k(t+\nu)$ for all $k \geq 0$. Further, we will say $X_t$ is invertible if, for every $t$, $\epsilon_t$ can be written as an infinite regression on $\{X_s, s \leq t\}$,

$$
\sum_{j=0}^{\infty} \pi_j(t) X_{t-j} \text{ where } \sum_{j=0}^{\infty} |\pi_j(t)| < \infty,
$$

(11)
with \( p_0(t) = 1 \) for all \( t \), \( p_j(t) = p_j(t + \nu) \), for all \( t \) and \( j \geq 0 \).

In the paper entitled "Time series with periodic structure" [24] (1967), Jones and Brelsford introduced the basic idea of applying the representation (8) to PAR systems, i.e., expressed by (4) but with \( \theta_k(t) \equiv 0 \) for \( k \geq 1 \). They further note that in many physical situations the periodic structure can be adequately expressed (modeled) by only a few lower order harmonic terms in (8). They also addressed prediction of a bivariate PAR by considering an auxiliary series

\[
U(t) = [x_1(t), x_2(t), x_1(t) \sin(2\pi t/T), x_2(t) \sin(2\pi t/T), x_1(t) \cos(2\pi t/T), ...]'
\]

onto which \( x(t) = (x_1(t), x_2(t))' \) is regressed and the coefficients of regression are determined by a Yule-Walker approach. Although solution to the prediction problem has given way to more direct implementation of Yule-Walker methods, the representations of parameters by low order Fourier series remains a pivotal contribution.

In the 1978 paper "On Periodic and Multiple Autoregressions" [35], Pagano made several important contributions to PAR time series, including one to our main theme. First, using Gladyshev’s mapping (5), the relation between \( \nu \)-variate VAR and PAR is given, namely that \( \mathbf{X}_n \) is a \( \nu \)-variate VAR with positive definite \( \Sigma \) (full rank) if the corresponding periodically correlated \( \mathbf{X}_t \) is a periodic autoregression of period \( \nu \) given by

\[
\mathbf{X}_t + \sum_{j=1}^{p} \phi_j(t) \mathbf{X}_{t-j} = \mathbf{\varepsilon}_t
\]

where \( \mathbf{\varepsilon}_t \) are uncorrelated with mean zero and for all \( t \), \( E\mathbf{\varepsilon}_t^2 = \sigma^2(t) > 0 \), \( p_t = p_{t+\nu} \), \( \sigma^2(t) = \sigma^2(t + \nu) \) and \( \phi_j(t) = \phi_j(t + \nu), j = 1, \ldots, p_t \) (note \( p_t \) can be made as large as necessary to realize all the necessary lags in the regression). Sufficient condition for stationarity of \( \mathbf{X}_n \) can be expressed (see Hannan [19]) from the VAR \( \mathbf{X}_n \) and its infinite MA. That is, if \( \det(\mathbf{I}_\nu + \sum_{j=1}^{p} \mathbf{A}(j)z^j) \neq 0 \) for \( |z| \leq 1 \), the VAR sequence \( \mathbf{X}_n \) has an infinite MA representation

\[
\mathbf{X}_n = \sum_{k=0}^{\infty} B_k \mathbf{e}_{n-k} \text{ where } \sum_{k=0}^{\infty} |B_k| < \infty.
\]

It follows that \( \mathbf{X}_n \) is stationary and causal, and hence that the associated PC sequence is causal and can be written as (10). An expression is given for the least MS predictor of \( \mathbf{X}_{t+h} \) given the past \( \mathbf{X}_t, \mathbf{X}_{t-1}, \ldots \) and then estimation of correlation is addressed, where it is proved that if \( \mathbf{X}_t \) is PC-\( \nu \) and Gaussian, then for \( t = 1, ..., \nu, u = 0, 1, ..., N\nu - t - 1 \),

\[
R_N(t, u) = \frac{1}{N} \sum_{n=0}^{m} X_{t+nu} X_{u+nu}^{N \to \infty} R(t, u)
\]

almost surely and in mean square, where \( m = \lfloor N - \max(t, u)/\nu \rfloor \); values of \( R(t, u) \) for other pairs of indices are determined by (3). The asymptotic limit of \( N \text{Cov}\{R_N(t_1, u_1), R_N(t_2, u_2)\} \) is given and it is shown that \( N^{1/2}\{R_N(t, u) - R(t, u)\} \) is asymptotically Normal. The Yule-Walker or normal equations for PAR sequences are, as usual, based on the
The idea that prediction errors must be orthogonal (normal) to the space of observations. Specifically, if $X_t$ is $PAR(p_1, p_2, \ldots, p_\nu)$ (with $\det(I_\nu + \sum_{j=1}^{p} A(j)z^j) \neq 0$ for $|z| \leq 1$), then the regression coefficients $\phi_j(t)$ and error variances $\sigma^2(t)$ are related to the correlations $R(\cdot, \cdot)$ through the normal equations

$$R(t, t-u) + \sum_{j=1}^{p} \phi_j(t) R(t-j, t-u) = \delta_{u0} \sigma^2(t)$$

(16) for $t = 1, 2, \ldots, \nu$ and $u \geq 0$. If $R$ is replaced by the sample correlations $R_N$, the solutions of (16) $\hat{\phi}_j(t), \hat{\sigma}(t)$ are called the Yule-Walker parameter estimators, and these estimators are shown to be a.s consistent and asymptotically normal, with the limiting covariance related to the appropriate part of the Fisher information matrix of the vector sequence $X_n$. The spectral density of the VAR associated with the PAR system (13) is found via the VAR part of (6)

$$LX_n + \sum_{j=1}^{p} A_j' X_{n-j} = \varepsilon_n'$$

where $L$ is unit lower triangular and $E\{\varepsilon'_n \varepsilon_n''\} = D = diag(\sigma^2(1), \ldots, \sigma^2(\nu))$. Defining $G(z) = L + \sum_{j=1}^{p} A_j' z^j$ the spectral density is then found in the usual way

$$f(\omega) = \frac{1}{2\pi} G^{-1}(e^{i\omega}) D(e^{i\omega})^\nu - 1.$$  

Finally, it is noted that the asymptotic normality of the estimators $\hat{\phi}_j(t)$ will imply it for $\hat{a}_{j,n}$ because of the continuity of the transformation (8) for PAR.

Contemporaneously with Pagano’s paper and in the years following, several related topics were addressed. Parzen and Pagano [36] (1979) addressed the issue of decomposing seasonal/periodic series into simpler parts. Ansley [6] (1979) developed a method for computing parameter likelihoods for samples from ARMA sequences; this method was subsequently used in other applications including PARMA (see below). In [43] (1979), Troutman addressed the representation of a PAR sequence $X_t$ as an infinite linear combination of independent, periodically distributed random variables, constraints on the parameters which permit such a representation, the covariance and spectral properties, and the asymptotic behaviour of cumulative sums of $X_t$ and of functions of these sums. In [42] (1980), Tiao and Grupe cover some of the same ground as [35] and [43], obtaining, via Gladyshev, the relationship between PAR and VAR. An important contribution was their discussion of the cost (to prediction error) for misspecification of models, such as assuming a stationary model when the data are truly a PAR. Hasselman and Barnett [20] (1981) address prediction when statistics are periodically time varying, but not in the specific context of PARMA models. They observe that representing the time dependence of the solution by Fourier series may give more significance at the price of “predictor skill” which is synonymous with normalized variance explained, so error free prediction has skill of unity. Vecchia, Obeysekera and Salas in [48] (1983) studied the ARMA sequence that is obtained from the yearly average (aggregate) of monthly values from a $PARMA(1, 1)$ series. The main finding
was that rather than estimating the parameters of the resulting ARMA, it is better to estimate the periodic parameters first and estimate the aggregated models based on the periodic estimates and the functional relationship between the parameters.

The book, *Applied Modeling of Hydrologic Time Series* by Salas, Delleur, Yevjevich and Lane, [50] (1980) contained several early treatments of problems connected to periodically varying ARMA time series. Included are (sec. 3.3) the Fourier series representation for statistical characteristics such as the periodic mean, variance and correlation coefficients, and the idea that in hydrologic series, only a few terms in the Fourier series should adequately represent the periodic time variation. The problem of choosing the best (most significant) frequencies for use in the representation is discussed and a method based on the cumulative periodogram test is recommended. Periodically varying AR models are introduced and estimation of correlation coefficients and residual variances treated in Sec 4.3.1. A method for estimation of correlation coefficients, based on a system of equations arising from a difference equation for the correlation coefficient, is suggested, as is the use of (periodic) Yule Walker equations in the estimation of PAR coefficients (Sec. 5.3.2). Throughout, normalizing the series by subtracting out the periodic sample mean and dividing by periodic sample $\sigma$ is advocated, as is use of the AIC method for parsimonious parameter selection. Some closed form expressions for parameter estimates in terms of sample correlations are given in Sec. 5.3.2 for some simpler cases of PARMA, notably the PARMA(1,1) model, and reliability of estimated parameters is discussed for the constant parameter case but not for periodic $\phi_j(t), \theta_k(t)$ or of their Fourier coefficients. Finally, procedures are outlined for multivariate cases of PAR and PARMA (Sec. 7.3).

In 1985 A.V. Vecchia published two papers [45, 46] that treated maximum likelihood estimation for PARMA time series. In [45] the Gladyshev mapping is used to derive (6) and referencing Fuller, 1976, ch. 2, to give the condition (7) for $X_n$ to be stationary and therefore the associated univariate $X_t$ to be PC, Vecchia mentions that an open problem is to find a condition on the roots of $\phi(t, z) = 1 - \sum_{j=1}^{p} \phi_j(t)z^j, t = 1, 2, ..., \nu$ that is equivalent to (7).

Applying (7) to PAR(1) gives $\text{det}(L - U_1 \lambda) = 1 - A\lambda$ with $A = \prod_{n=1}^{\nu} \phi_1(n)$, so the condition for stationarity is $|A| < 1$, which shows that the process can be locally expanding $|\phi(n)| > 1$ for some $n$ and locally contracting for other $n$, and still $|A| < 1$. See [23, Thm. 8.1] for a proof that any two of the following three conditions implies the other one: (i) $|A| < 1$, (ii) $||X_t||$ is bounded, (iii) $X_i$ is causal with respect to $\xi_i$ in (4). The sampled (marginal) series at time (phase or season) $i$, namely $X_{n+1}, n \in \mathbb{Z}$ are shown to be ARMA and also the yearly aggregates are ARMA; in both cases, the orders $(p, q)$ are given. For estimating the coefficients of a PAR model using sample autocorrelations in the Yule-Walker method, parameters are estimated for each season $i$ individually whereas in Pagano [35] a larger matrix system was solved to estimate the entire collection $\phi_i, i = 1, 2, ..., \nu$ simultaneously. Vecchia recommends moment estimates for PAR as if the errors are Gaussian, then these estimates are asymptotically most efficient (Pagano, [35]). In addition, $\phi_\nu$ will automatically satisfy the stationarity assumption (7) and $\sigma^2(\nu)$ is real-valued and positive (Troutman, [43]). Maximum likelihood estimation is introduced using the idea of Ansley [6] applied to PARMA.
systems. Specifically, set

\[ W_t = \begin{cases} X_t & t \in I_{t_0;m} \\ X_t - \sum_{j=1}^{p} \phi_j(t)X_{t-j} & t \in I_{t_0+m;n-m} \end{cases} \]

for \( t \) on the set of integers beginning at \( t_0 \) and of length \( n \),

\[ I_{\{t_0;n\}} = \{t_0, t_1, \ldots, t_0 + n - 1\} = I_{t_0;m} \cup I_{t_0+m;n-m}, \]

and \( m = \max(p, q) \). Then \( W_{t_0;n} = A_{\Phi}X_{t_0;n} \) where matrix \( A_{\Phi} \) is upper triangular with \( \det A_{\Phi} = 1 \), hence leading to

\[ L(\Phi, \Theta|X_{t_0;n}) = L(\Phi, \Theta|W_{t_0;n}). \]  

Note here we are using \( \theta_0(t) = \sigma(t), t = 0, 1, \ldots, \nu - 1 \). Furthermore, considering the \( m \) first \( W_t \)'s as fixed parameters, the reduced vector \( W_{t_0+m-1;n} \) may be seen follow the periodic MA associated with the right side of (4). Finally, he gives an algorithm outline, that uses moment estimates of the \( \sigma^2(t) \) in a loop with estimates of the AR and MA parameters obtained by maximizing (18). Vecchia calls this an approximate likelihood and some call it conditional, as it treats the leading \( m \) ordinates of the series as fixed. For \( n \gg m \), the error of approximation can be expected to be negligible.

Concerning our current topic, for a \( PARMA_{12}(1, 1) \) model he expands \( \phi_1(t) \) and \( \theta_1(t) \) in a Fourier series, mentioning that only a slight change to a maximum likelihood procedure is required. We demonstrate it later via simulation. The important problem of determining the number of harmonics in representation (8) is introduced and an approach suggested that uses the AIC for choosing the number of parameters. The method is based on increasing number of harmonics in (8) to get a family of estimates and \( AIC^{(k)} \) values sequentially for \( k = 1, 2, \ldots, \lfloor \nu/2 \rfloor \), where

\[ AIC^{(k)} = -2 \ln L(\hat{\phi}^{(k)}, \hat{\theta}^{(k)}, \hat{\sigma}^{(k)}) + 8k. \]

At the \( k \)th step, initial estimates of the Fourier parameters are set to the ML estimates of the first \( k - 1 \) harmonic amplitudes determined in the \( k - 1 \)st step. Then the number of harmonics is determined to be the \( k \) that minimizes \( AIC^{(k)} \). The method is demonstrated by application to a streamflow series from the Rio Caroni River.

In [46], methods introduced in [45] are slightly amplified and simulations are included to illustrate the ideas. It is mentioned that identification for PARMA is an important unsolved problem (although much has been accomplished subsequently), and that parameter estimates based on moment methods often were not in an acceptable parameter space whereas ML methods proved better in this way.

In their 1988 paper, Li and Hui [25] also address the computation of the exact likelihood of a PARMA, providing a clear exposition of the ML method, although much like Vecchia’s [45, 46] description. The method of Ansley and the Cholesky decomposition are used for computing the orthogonal residuals and computing the likelihood.

H. Sakai [38] (1991) addressed the computation of spectral density of a PARMA sequence. A little additional background will help clarify the sense of spectral density.
From Gladyshev [16], PC sequences are harmonizable (strongly) and have the spectral representation (in the mean square sense)

\[ X_t = \int_0^{2\pi} e^{i\lambda t} z(d\lambda) \quad (19) \]

and this leads to the representation

\[ X_t = \sum_{k=0}^{\nu-1} z_k^t e^{i2\pi kt/\nu} \quad (20) \]

where

\[ z_k^t = e^{-i2\pi kt/\nu} \int_{[k2\pi/\nu,(k+1)2\pi/\nu]} e^{i\lambda t} z(d\lambda) \quad (21) \]

Note the spectral support of \( z_k^t \) is \([0, 2\pi/\nu)\), and the family \( \{z_0^t, z_1^t, \ldots, z_{\nu-1}^t\} \) is vector stationary with spectral distribution as follows. The harmonizability of \( X_t \) and condition (3) leads to the representation of the covariance

\[ E\{X_s \bar{X}_t\} = R(s, t) = \int_0^{2\pi} \int_0^{2\pi} e^{i(s\lambda_1 - t\lambda_2)} F(d\lambda_1, d\lambda_2) \quad (22) \]

where the support of \( F \) is contained in the union \( S_\nu = \cup_{k=-\nu+1}^{\nu-1} S_k \) of \( 2\nu - 1 \) diagonal lines

\[ S_k = \{ (\lambda_1, \lambda_2) \in [0, 2\pi) \times [0, 2\pi) : \lambda_2 = \lambda_1 - 2\pi k/\nu \}. \quad (23) \]

Since the components \( \{z_0^t, z_1^t, \ldots, z_{\nu-1}^t\} \) are jointly stationary, they have a matrix spectral distribution \( F \) which may be identified with the spectral measure \( F \) in (22) by partitioning the square \([0, 2\pi) \times [0, 2\pi)\) into subsquares of side \( 2\pi/\nu \) and then the cross spectral measure \( F_{pq}(\cdot) \) is obtained from the diagonal of subsquare \( pq \) by \( F_{pq}(A) = F(A + 2\pi p/\nu, A + 2\pi q/\nu) \). Gladyshev [16] also determined the transformation between \( F \) and the matrix spectral distribution \( \hat{F} \) of the blocked stationary sequence \( X_n \) (see (5)):

\[ F(d\lambda) = \frac{1}{\nu} V^{-1}(\nu\lambda) F(\nu d\lambda) V(\nu\lambda) \quad \lambda \in [0, 2\pi/\nu) \quad (24) \]

where \( V(\lambda) \) is a unitary matrix (a map \( \mathbb{C}^\nu \rightarrow \mathbb{C}^\nu \)) whose \((p, k)\)th element is given by

\[ v^{pk}(\lambda) = \frac{1}{\sqrt{T}} e^{i2\pi pk/T + i\lambda p/T}. \quad (25) \]

A more complete treatment may be found in [23] (see Figure 6.5 and surrounding text). Since the transformation (24) is invertible, the spectral distribution for the PARMA sequence may be considered to be either \( F \) or \( \hat{F} \). Beginning with \( F \) for a blocked PARMA sequence, from (6) denote

\[ A(z) = L - \sum_{j=1}^{p'} U_j z^j, \quad B(z) = \Gamma - \sum_{k=1}^{q'} V_k, \]
\[ \Sigma = \text{Cov}\{e_n\} = \text{diag}\{\sigma^2(0), \sigma^2(1), \ldots, \sigma^2(\nu - 1)\} \quad (26) \]
and then the matrix spectral density for $X_n$ if given by [10, p.431] (also Hannan [19])

$$f(\lambda) = \frac{dF(\lambda)}{d\lambda} = \frac{1}{2\pi} [A^{-1}(z)B(z)] \Sigma [A^{-1}(z)B(z)]' \big|_{z=\exp(-i\lambda)}, \quad 0 \leq \lambda < 2\pi \quad (2\pi/\nu??)$$

(27)

This gives the matrix spectral density of the blocked sequence in terms of the periodic parameters $\phi_j(\cdot), \theta_j(\cdot), \sigma^2(\cdot)$ which may be expressed in terms of the Fourier PARMA parameters as in (8). Sakai takes this further, and shows a similar formula for the spectral density of $F(\lambda)$,

$$\frac{dF(\lambda)}{d\lambda} = \frac{1}{2\pi} [A^{-1}(z)B(z)]Q[A^{-1}(z)B(z)]' \big|_{z=\exp(-i\lambda)}, \quad 0 \leq \lambda < 2\pi \quad (2\pi/\nu??)$$

(28)

except now $A$ and $B$ contain the complex Fourier coefficients determined as follows. Define

$$a_{jn} = \frac{1}{p} \sum_{t=0}^{\nu-1} \phi_j(t) \exp(i2\pi nt/\nu),$$

(29)

then $a_{j} = [a_{j0}, a_{j1}, \ldots, a_{j\nu-1}]$ and finally $A_j = [a_{j}, Ta_{j}, T^2a_{j}, \ldots, T^{\nu-1}a_{j}]D^j$ where $D = \text{diag}\{1, e^{-i2\pi/\nu}, \ldots, e^{-i2\pi(\nu-1)/\nu}\}$ and the transformation $T$ is right rotation with wrapping. Finally, $A(z) = I - A_1z - A_2z^2 - \cdots - A_{\nu}z^\nu$ and similarly for $B(z)$. Denoting $s$ as the complex Fourier coefficients of $\sigma^2(t)$, i.e., $s_n = \frac{1}{p} \sum_{t=0}^{\nu-1} \sigma^2(t) \exp(i2\pi nt/\nu)$, then $Q$ is determined by $Q_{mn} = s_{m-n}$.

Sakai gives other ways to express (27) and (28), but the main point, now, is that formulas exist for expressing the spectral densities in terms of the PARMA parameters, and methods exist (see e.g., [23]) for estimating the spectral densities. This gives a possible path for estimation of PARMA parameters via spectral matching as in [13].

In [47] (1991), Vecchia and Ballerini present statistical procedures for deciding if the PC property is present in the autocorrelation function of a seasonal time series after removal of seasonal means and normalizing by seasonal standard deviations. The procedures are based on (1) asymptotic properties of estimates of the periodic autocorrelation function $\hat{\rho}_t(u) = \gamma_t(u)/[\gamma_t(0)\gamma_t+u(0)]^{1/2}$, and of their Fourier coefficients and (2) on the null hypothesis that the observed series is stationary. It is clear that for $\rho_t(u) = \rho_{t+u}(u)$ to be properly periodic for fixed $u$, at least one of the Fourier coefficients having $m > 0$, in $\rho_t(u) = c_0(u) + \sum_{m>0} [c_m(u) \cos 2\pi mt/\nu + s_m(u) \sin 2\pi mt/\nu]$, must be nonzero. The tests are based on asymptotic properties of the estimated Fourier coefficients

$$\tilde{F}(u) = \left\{ \begin{array}{ll}
(\hat{c}_0(u), \hat{c}_1(u), \hat{s}_1(u), \ldots, \hat{c}_{(\nu-1)/2}(u), \hat{s}_{(\nu-1)/2}(u))' & (\nu \text{ odd}) \\
(\hat{c}_0(u), \hat{c}_1(u), \hat{s}_1(u), \ldots, \hat{s}_{(\nu/2-1)}(u), \hat{c}_{(\nu/2)}(u))' & (\nu \text{ even})
\end{array} \right.$$
where

\[ \hat{c}_r(u) = \nu^{-1/2} \sum_{m=0}^{\nu-1} \hat{\rho}_m(u) \cos \frac{2\pi rm}{\nu} \quad (r = 0 \text{ or } \nu/2) \]

\[ \hat{c}_r(u) = 2^{1/2} \nu^{-1/2} \sum_{t=0}^{\nu-1} \hat{\rho}_m(u) \cos \frac{2\pi rm}{\nu} \quad (r = 1, \ldots, (\nu - 1)/2) \] (31)

\[ \hat{s}_r(u) = 2^{1/2} \nu^{-1/2} \sum_{t=0}^{\nu-1} \hat{\rho}_m(u) \sin \frac{2\pi rm}{\nu} \quad (r = 1, 2, \ldots, (\nu - 1)/2) \]

for \( n = 2, \ldots, \lfloor \nu/2 \rfloor, j = 1, \ldots, p \).

The assumptions for deriving the tests are that \( X_t - \mu_t \) is causal, as described in (10) with \( \text{Var} [\varepsilon_t] = 1 \), and \( E \{ \varepsilon_t^4 \} < \infty \). These assumptions lead to the asymptotic normality for the estimators of the periodic autocovariance \( \gamma_t(u) \) and for the correlation \( \rho_t(u) \). When \( X_t \) is stationary, the asymptotic covariances of the vector \( \hat{\rho}(u) = [\hat{\rho}_0(u), \hat{\rho}_1(u), \ldots, \hat{\rho}_{\nu-1}(u)]' \) are expressed in terms of the cross-spectral density function \( g_{mm'}(\cdot) \) between the lag \( m \) and lag \( m' \) product processes \( \{X_tX_{t+m}\} \) and \( \{X_tX_{t+m'}\} \). Finally, under the null hypothesis, the Fourier coefficients of \( \hat{r}(u) \) in (30) are asymptotically independent and normally distributed and furthermore, have asymptotic mean zero and asymptotic variance given by the asymptotic covariances of the vector \( \hat{\rho}(u) \).

Remark 1. Testing for nonzero Fourier coefficients for the mean or correlation, as in (30), leads to test statistics that can be expressed directly in terms of the data. On the other hand, since PARMA coefficients are indirectly expressed (4) it can be expected that estimating or testing for nonzero parameters involves test statistics that are also indirectly determined.

Remark 2. See Hurd and Gerr [22] (1991) for a method to determine the presence of periodic correlation that is based on the estimation of correlations (or coherency) in the sample spectrum of the time series. The method is also discussed in [23].

McLeod, in [32] (1993), advocates testing the residuals of fitting seasonal ARMA models in order to identify the presence of periodic correlation; a positive result indicates the model does not adequately incorporate periodic correlation. Inadequacy of usual ARMA models was found in several of the hydrologic and economic time series that were tested.

Anderson and Vecchia, in [5] (1993), address asymptotics for estimators of the autocovariance and autocorrelation, but through the use of Gladyshev’s mapping (5) so that results from stationary vector time series can be applied (see Hannan [19]). The working assumptions include a 2-sided version of the causality condition (10). Denoting \( \gamma(u) = [\gamma_0(u), \gamma_1(u), \ldots, \gamma_{\nu-1}(u)]' \) as a vector of covariances with \( u \) designating lag, the joint asymptotic result for two vectors \( \gamma(u), \gamma(v) \) is obtained; namely

\[ N^{1/2} \begin{pmatrix} \hat{\gamma}(u) - \gamma(u) \\ \hat{\gamma}(v) - \gamma(v) \end{pmatrix} \Rightarrow N \left( 0, \begin{bmatrix} V_{uu} & V_{uv} \\ V_{vu} & V_{vv} \end{bmatrix} \right) \]
where the $i, l$th term in the $\nu \times \nu$ matrix $V_{uv}$ is given by

$$
[V_{uv}]_{i,l} = \sum_{n=-\infty}^{\infty} \gamma_i(n\nu + l - i)\gamma_{i+u}(n\nu + l - i - u) + \gamma_i(n\nu + l + v)\gamma_{i+u}(n\nu + l - i - u).
$$

(32)

And then similarly, for the estimator $\hat{\rho}(u)$ of the periodic autocorrelations,

$$
N^{1/2} \left[ \begin{array}{c} \hat{\rho}(u) - \rho(u) \\ \hat{\rho}(v) - \rho(v) \end{array} \right] \Rightarrow \mathcal{N} \left( \begin{array}{cc} 0 & W_{uv} \\ W_{vu} & W_{vv} \end{array} \right)
$$

where $\rho(u) = [\rho_0(u), \rho_1(u), \ldots, \rho_{\nu-1}(u)]^T$, and $W_{uv}$ are doubly infinite sums of products of permutation matrices and matrices $F_n = \text{diag}(\rho_0(n), \rho_1(n), \ldots, \rho_{\nu-1}(n))$. The asymptotics for the Fourier coefficients ($\hat{r}(u)$ in (30)) are obtained by application of the continuous mapping theorem to the $\rho(u)$ above (as in [47]). Under the null hypothesis that the sequence is stationary, so $\rho_1(u) \equiv \rho(u)$, it follows that Fourier coefficients $\hat{r}(u)$ are asymptotically independent and normally distributed:

$$
N^{1/2} [\hat{c}_m(u) - \mu_m(u)] \text{ and } N^{1/2} [\hat{s}_m(u) - \mu_m(u)] \Rightarrow \mathcal{N}(0, R_m(u))
$$

(33)

for all $u \geq 1$ where $\mu_m(u) = \nu^{1/2} \rho(u), m = 0$ and $\mu_m(u) = 0, m > 0$. Denoting $\rho(u) = E\{Y_tY_{t+u}\}$ where $Y_t = (X_t - \mu_t)/\gamma_t^{1/2}$ is the standardized process, the limiting variances are

$$
R_m(u) = \begin{cases} 
V_m(u) & (m = 0 \text{ or } \nu/2) \\
2^{-1}V_m(u) & (0 < m < \nu/2) 
\end{cases}
$$

(34)

with

$$
V_m(u) = \sum_{n=-\infty}^{\infty} \cos(2\pi mn/\nu)[\rho^2(n) + \rho(n - u)\rho(n + u) + 2\rho(u)\rho(n + u) + \rho(n)] + \frac{1}{2}\rho^2(u)\rho^2(n - u) + 2\rho(n) + \rho^2(n + u)].
$$

(35)

Two examples given where the asymptotic results can be applied to PARMA fitting. First is to identification of frequencies in the Fourier parameterization of a PAR(1) model. Specifically, one is to decide if the parameters $a_1$ and $b_1$ are significantly non-zero in the model

$$
\phi_1(t) = a_0 + 2a_1\cos(2\pi t/\nu) + 2b_1\sin(2\pi t/\nu).
$$

Since the correlation $\rho_1(u)$ is given so simply in terms of $\phi(t)$ for the PAR(1) model, the machinery set up to obtain the asymptotic results for $\hat{r}(u)$ can be applied to obtain asymptotic results for $\phi_1(t)$, its estimators and for the Fourier coefficients of the estimators, yielding also standard errors of these, from which decision thresholds may be
computed. This may be seen as a foretelling of the more general results to appear later in [34, 40].

The second application is to determine which harmonic terms appear in the time variation of the autocorrelation \( \rho_t(u) \) for a Salt River streamflow series. The standard errors \( R_m(u) \) were estimated from an estimator of average correlations \( \hat{\rho}_m(u) = \hat{c}_0(u)/\sqrt{12} \), when fitted to \( u \), yielding approximately \( \hat{\rho}(u) \approx 0.68 \); and finally \( R_m(u) \) were approximated via (34) and (35). Using Bonferroni\(^1\) corrections for 12 seasons, the significant \( \hat{c}_m \), \( \hat{s}_m \), at \( p = .01 \), suggested a maximum of \( m = 2 \) harmonics in the Fourier series for \( \phi_1(t), \theta_1(t), \sigma(t) \). The final estimates were based on the maximum likelihood techniques of Vecchia [45, 46].

Also in 1993, Ula [44] addressed minimum MSE forecasts and their errors for multivariate PARMA. Recursive evaluations of these quantities were shown to follow from the conditional expectation approach, update equations and prediction ellipsoids and for future values of the process were given for the obtained forecasts. The Fourier parameterization was not addressed.

The paper [33] (1994) of McLeod addresses mainly identification, estimation and diagnostic checking for PAR models. To test for remaining PC structure in residuals, a modified (for the periodic case) portmanteau statistic,

\[
\hat{Q}_{L,m} = \sum_{l=1}^{L} \frac{\hat{r}_{l,m}^2}{\text{Var}(r_{l,m})} \tag{36}
\]

is presented, where \( \hat{r}_{l,m} \) is the residual autocorrelation for season \( m \) and lag \( l \), and \( \text{Var}(r_{l,m}) \) is a theoretical variance based on periodic correlations of white noise.

Lund and Basawa [26] (1999) make several important advances to the analysis of causal invertible PARMA time series, beginning with the recursive computation of the coefficients \( \psi_k(t) \) in (10) and coefficients \( \pi_j(t) \) in (11) in terms of parameters \( \theta_k(t) \) and \( \phi_j(t) \). When (7) holds, then also does (10) and then the covariance can be expressed for all \( t \) and \( u \geq 0 \) as

\[
\gamma_t(u) = \text{Cov}\{X_t, X_{t+u}\} = \sum_{k=0}^{\infty} \psi_{k+u}(t+u)\psi_k(t-u)\sigma^2(t-k). \tag{37}
\]

The recursion in the variable \( t \) is given for determining the coefficients \( \{\theta_{t,k}, k = 1, 2, \ldots, t\} \) and the error variances \( v_t \) of the LMS 1-step predictor \( \hat{W}_{t+1} \) based on previous errors,

\[
\hat{W}_{t+1} = \sum_{k=1}^{t} \theta_{t,k}(W_{t+1-k} - \hat{W}_{t+1-k}) \tag{38}
\]

where \( W_t \) is given by (17), and for which the covariance matrix \( R_N = \text{Cov}\{X_{1,N}X'_{1,N}\} \) of the sample \( X_{1,N} = \{X_1, X_2, \ldots, X_N\}' \) is invertible for all \( N > 0 \). This recursion,

\(^1\)If \( N \) independent and identically distributed random variables are tested, the Bonferroni adjusted threshold is established by controlling, in the null case, the probability of at least one threshold exceedance occurring in the \( N \) tests.
called the innovation algorithm, is presented for the stationary case with similar notation in [10].

In the following, the prediction is to time \( t + 1 \), where \( t = n\nu + s \) for \( 0 \leq s \leq \nu - 1 \).

Some facts derived are (a) for causal and invertible PARMA’s, \( v_{n\nu+s} \to \sigma(t+1) = \sigma(s+1) \) as \( n \to \infty \) for each season \( s \); (b) for any \( \text{PC}(\nu) \) series, \( v_{n\nu+s} \) is nonincreasing w.r. \( n \) for each \( s \); (c) for causal and invertible PARMA’s, \( \theta_{n\nu+s-1,k} \to \theta_k(s) \) as \( n \to \infty \) for each season \( s \) and \( 1 \leq k \leq q \). Note these results follow only from the structure of the covariance, a non-random quantity. The use of sample covariance in the innovation algorithm is treated in the paper by Anderson, Meerschaert and Vecchia [4] discussed subsequently.

Once the \( \theta_{i,k} \) and \( \hat{e}_i \) are computed, the likelihood for the sample \( \{X_1, X_2, \ldots X_t\} \) for each season \( \nu \) where \( t = N \), can be computed (see Brockwell and Davis [10, Eq. (8.7.4),p.256]) by using \( X_j - \hat{X}_j = W_j - \hat{W}_j \) in

\[
L(\hat{\psi}, \hat{\theta}, \hat{\sigma}; \hat{X}) = \frac{1}{(2\pi)^{N/2}} \left( \prod_{j=0}^{N-1} v_j \right)^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} \sum_{j=1}^{N} \frac{(X_j - \hat{X}_j)^2}{v_{j-1}} \right],
\]

provided \( R_N \) is invertible. By their proposition 4.1, \( R_N \) is indeed invertible for each \( N \geq 1 \) provided \( X_t \) is causal and \( \sigma(t) > 0 \) for \( 0 \leq t \leq \nu - 1 \).

Assuming causality and invertibility as in (10) and (11), Anderson, Meerschaert and Vecchia, [4, 1999] also give the innovations algorithm plus results on convergence of quantities arising from the algorithm or derived from those quantities. In the following, the objective is to find the LMS predictor of \( X_{i+n} \) based on the observations \( \{X_i, X_{i+1}, \ldots, X_{i+n-1}\} \) in terms of the past prediction errors

\[
\hat{X}_{i+n} = \begin{cases} 0 & n = 0 \\ \sum_{k=1}^{n} \theta_{n,k}^{(i)} (X_{i+n-k} - \hat{X}_{i+n-k}) & n \geq 1 \end{cases}
\]

for \( i = 0, 1, \ldots, \nu - 1 \), the starting time (season) of data vector, \( n \) is the length of the data vector and \( \hat{X} \) is the index of the orthogonal components \( \{X_{i+n-k} - \hat{X}_{i+n-k}\} \) of the predictor. Denote \( v_{n,i} = E\{[X_{i+n} - \hat{X}_{i+n}]^2\} \) as the mean square prediction error. Specifically, with \( <k> = k \mod \nu \) as the season associated with \( k \), they establish that (1) for each \( i = 0, 1, \ldots, \nu - 1 \), predictor error variance \( \sigma_m^{<i-m>} \) converges to shock variance \( \sigma^2(i) \) as \( m \to \infty \); (2) prediction error converges to the shock, \( ||X_{i+m} - \hat{X}_{i+m}^{(i)} - \varepsilon_{i+m}|| \to 0 \) as \( m \to \infty \); (3) prediction coefficients \( \theta_{i-m}^{<i-m>} \to \psi_k(i) \), the IMA coefficients, as \( m \to \infty \); (4) Y-W coefficients: solving \( \Gamma_{n,i} \phi_n^{(i)} = \gamma_n^{(i)} \) yields \( \phi_m^{<i-m>} \to -\pi_k(i) \), the negative of the inversion coefficients, as \( m \to \infty \).

Addressing now use of sample autocovariances \( \hat{\gamma}_t(u) \) in place of \( \gamma_t(u) \), the following are obtained: (5) The coefficient estimates from the innovations algorithm are weakly consistent in the sense that, as \( k \to \infty \),

\[
(\hat{\theta}_{k,1}^{<i-k>} - \psi_1(i), \hat{\theta}_{k,2}^{<i-k>} - \psi_2(i), \ldots, \hat{\theta}_{k,k}^{<i-k>} - \psi_k(i), 0, \ldots) \to 0;
\]

(6) [Thm 3.1] If \( X_t \) is PARMA with period \( \nu \), \( E\varepsilon_t^2 < \infty \), and \( f(\lambda) \) is spectral density of the associated VARMA such that there exist \( m \leq M < \infty \) with \( mzz' \leq zf(\lambda)z' \leq
To all \( z \in \mathbb{R}^v \), and if \( k(N) \) s.t. \( k^2(N)/N \to 0 \) as \( N \to \infty \), then \( \| \hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1} \|_2 \to 0 \) as \( k \to \infty \); (7) [Thm 3.2] Under the same hypotheses, and using sample autovariances, \( \hat{\Gamma}_{k,i} \) and \( \hat{\gamma}_k^{(i)} \), to solve the Yule Walker equations, \( \hat{\phi}_k = \hat{\gamma}_k \), leads to \( \hat{\phi}_k^{(i)} - \hat{\phi}_k^{(i)} \to P \) as \( k \to \infty \); (8) [Thm 3.3] Under the same hypotheses, for all \( \nu \) and starting times \( \nu \), \( \hat{\phi}_k^{(i-k)} \to -\pi_j(i) \) as \( k \to \infty \); (9) [Thm 3.4] Under the same hypotheses, for all starting times \( \nu \), the error variances \( \hat{\psi}_{k,\nu} \) from the innovations algorithm satisfy \( \hat{\psi}_{k,\nu} \to \sigma^2(i) \). The remainder of paper treats \( \alpha \)-stable shocks, a subject outside our current scope.

In the paper ”Parameter estimates for periodically stationary time series” [3], Anderson and Meerschaert make a further advance by giving conditions under which the convergence of \( \hat{\phi}_k^{(i-k)} \), \( \hat{\phi}_k^{(i-k)} \), \( \hat{\psi}_{k,\nu} \) are asymptotically normal. The assumptions needed for the result include, causality, invertibility, \( E\hat{\epsilon}_i^2 < \infty \) and the spectral condition on the equivalent vector moving average (6). In addition, it is assumed that

\[
N^{1/2} \sum_{j=0}^{\infty} |\pi_k(N) + \nu(s)| \to 0 \quad \text{for } s = 0, 1, 2, \ldots, \nu - 1 \tag{41}
\]

where \( k(N) \to \infty \) and \( k^3/N \to 0 \) as \( N \to \infty \). Then for every finite \( D \),

\[
N^{1/2} \left( \pi_u(i) + \hat{\phi}_k^{(i-k)} : 1 \leq u \leq D, 0 \leq \nu \leq \nu - 1 \right) \Rightarrow \mathcal{N}(0, W),
\]

and

\[
N^{1/2} \left( \hat{\phi}_k^{(i-k)} - \psi_u(i) : 1 \leq u \leq D, 0 \leq \nu \leq \nu - 1 \right) \Rightarrow \mathcal{N}(0, V),
\]

See [3] for the details of the \( D \nu \times D \nu \) covariance matrices \( W \) and \( V \). A useful corollary is given for fixed \( u \) and \( i \), namely

\[
N^{1/2} \left[ \hat{\phi}_k^{(i-k)} - \psi_u(i) \right] \Rightarrow \mathcal{N} \left( 0, \sum_{n=0}^{u-1} \frac{\sigma^2(i - n)}{\sigma^2(i - u)} \psi_u(i) \right). \tag{42}
\]

In [34] (2007), Anderson, Tesfaye and Meerschaert present a major milestone on statistical methods for Fourier-PARMA models by establishing asymptotic normality of the discrete Fourier transform coefficients (see (9)), and then developing tests for significantly non-zero coefficients. The results may be viewed as an extension of the method presented by Vecchia and Ballerini [47] and Anderson and Vecchia [5], where the Fourier coefficients of \( \tilde{\rho}_k(\tau) \) were used to detect presence of PC structure.

Using the assumptions of causality (10) and invertibility (11), the innovations algorithm is used to produce estimates \( \hat{\psi}_u(i) = \hat{\phi}_k^{(i-k)} \), \( \hat{\psi}_u(i) = \psi_k(i-k) \), and then the causal representation (10) is used in (4) to produce a system of equations involving the
parameters $\psi_k(t), \phi_j(t), \theta_k(t)$

\[
\begin{align*}
\psi_0(t) &= 1 \\
\psi_1(t) - \phi_1(t)\psi_0(t-1) &= -\theta_1(t) \\
\psi_2(t) - \phi_1(t)\psi_1(t-1) - \phi_2(t)\psi_0(t-2) &= -\theta_2(t) \\
\psi_3(t) - \phi_1(t)\psi_2(t-1) - \phi_2(t)\psi_1(t-2) - \phi_3(t)\psi_0(t-3) &= -\theta_3(t)
\end{align*}
\]

subject to the $\text{PARMA}(p,q)$ conditions that for all $t$, $\theta_k(t) = 0$ for $k > q$ and $\phi_j(t) = 0$ for $j > p$. This must be solved for all $t$ but the periodicity, $\psi_j(t) = \psi_j(t+\nu)$ for all $j$, implies that the solutions for $t = 0, 1, \ldots, \nu - 1$ is enough. Solving these equations is generally difficult, but simple enough for the important case of $p = q = 1$ which produces $\psi_1(t) - \phi_1(t) = -\theta_1(t)$ and $\psi_2(t) - \phi_1(t)\psi_1(t-1) = 0$. Thus $\phi_1(t), \theta_1(t)$ can be solved in terms of $\psi_u(t), u = 1, 2$, giving estimates $\hat{\phi}_1(t)$ and $\hat{\theta}_1(t)$ in terms of $\hat{\psi}_u(t), u = 1, 2$. Using the result (42) from [3], it follows that $\mathbb{N}_t^\phi[\phi_i - \phi_i] \Rightarrow \mathcal{N}(0, w_{\phi_i}^2)$ and $\mathbb{N}_t^\phi[\theta_i - \theta_i] \Rightarrow \mathcal{N}(0, w_{\theta_i}^2)$. Expressions for the variances $w_{\phi_i}^2$ and $w_{\theta_i}^2$ in terms of some $\sigma(i)$ and $\psi_u(i)$ are given in [34]. The authors first present the testing methodology, including Bonferroni threshold adjustments to correct for multiple frequencies, for the PMA case. This is followed by the discussion of the $\text{PARMA}(p,q)$ case, but the full solution is limited to $\text{PARMA}(1,1)$ where the system (43) can be solved. Here we summarize the latter.

Adapting to our notation, the parameters $\phi_j(t), \theta_k(t)$ and $\theta_0(t) = \sigma(t)$ are assumed to be given by (8) with inverse relation (9). It follows that under the null hypothesis ($X_t$ is stationary), then $\phi_1(t) = \phi, \theta_1(t) = \theta$ and $\sigma(t) = \sigma$ are constant with respect to $t$ (thus producing null Fourier coefficients for non-zero frequencies) and so the resulting Fourier coefficients are asymptotically normal and mutually independent; that is, $\mathbb{N}_t^{\phi}[\hat{\phi}_{1,n} - \phi] \Rightarrow \mathcal{N}(0, V_{\phi})$ and $\mathbb{N}_t^{\phi}[\hat{\phi}_{1,n}] \Rightarrow \mathcal{N}(0, V_{\phi})$ for $n > 1$ and similarly $\mathbb{N}_t^{\phi}[\hat{\theta}_{1,n} - \theta] \Rightarrow \mathcal{N}(0, V_{\theta})$ and $\mathbb{N}_t^{\phi}[\hat{\theta}_{1,n}] \Rightarrow \mathcal{N}(0, V_{\theta})$ for $n > 1$. The asymptotic variance can be computed from some $\sigma_i$ and $\psi_i(u)$. In our notation (8) the set of $\{a_{1,2n}, n = 1, \ldots, \lfloor \nu/2 \rfloor \}$ are cosine coefficients and $\{a_{1,2n+1}, n = 1, \ldots, \lfloor \nu/2 \rfloor \}$ are sine coefficients. The thresholds (see their Eq. (27) and (28)) for testing if $E[\hat{a}_{1,n}] \neq 0$ or $E[\hat{b}_{1,n}] \neq 0$ for some $n > 1$ are based on the Bonferroni correction to $\alpha$ for the $n - 1$ coefficients, excluding frequency zero, to be tested.

The results are demonstrated by a simulation of a $\text{PARMA}_{12}(0,1)$ and a $\text{PARMA}_{12}(1,1)$. Here we summarize the latter, where $c_{a0} = 0.35, c_{a1} = 0.15, s_{a1} = 0.40, c_{a2} = 0.25, s_{a2} = 0.35, c_{a3} = 0.35, s_{b1} = 0.25, s_{b2} = 0.45, s_{b3} = -0.15, \sigma_t^2 \equiv 1, N_g = 500$. Using the Bonferroni corrected thresholds, their tests for non-zero coefficients agree perfectly with the model, only coefficients at harmonics 1 and 2 are significant. We use this same model in a simulation for smaller $N_g$ later to be discussed later.

These methods were applied to a 72 year stream flow series (Fraser River at Hope BC) using a $\text{PARMA}_{12}(1,1)$ model suggested by previous studies. Tests for significant frequencies were applied, using the Bonferroni correction with $\nu = 12, \alpha = .01$. Coefficient estimates at the identified frequencies, presumably computed by estimating
\( \hat{\psi}_i(u) \) via the innovations algorithm, solution of (43) to obtain \( \phi_1(1), \theta_1(1) \), and then computing coefficients by (9), yielded \( c_{a0} = 0.304, s_{a1} = -0.426, c_{a3} = 0.665, c_{a6} = 0.408, s_{a2} = 0.355, c_{b3} = -0.649 \). Subsequent identification procedures found the residuals to be consistent with white noise. Similar results were found for a weekly series.

In addition to the importance of the methods developed, this paper also shows that, as previously proposed, the Fourier parameterization can be extremely effective in reducing parameters while still producing statistically satisfactory explanation of the data. But of course, the extent to which a Fourier PARMA model adequately explains any particular data depend on the individual application.

Agnieszka Wyłomanska, in [49] (2008) obtains an expression for the spectral measures of a PARMA sequence in terms of the parameters. Using the notation from [23], the periodicity \( \gamma(t) = E\{X_{t+u}X_t\} = \gamma_{t+\nu}(u) \) leads to the Fourier series representation \( \gamma(t) = \sum_{l=0}^{\nu-1} e^{2\pi jl/\nu} B_j(u) \) where the Fourier coefficients can be expressed as \( B_j(u) = \int_0^{2\pi} e^{\lambda j u} \mu_j(\lambda) d\lambda \), with \( \mu_j(\cdot), j = 0, 1, \ldots, \nu - 1 \), each finite measures. This collection of measures is (sometimes) identified as the spectrum of a PC sequence.

The main result (Theorem 2) is that for bounded PARMA sequences the measures \( \mu_j(\cdot) \) are absolutely continuous with respect to normalized Lebesgue measure on \([0, 2\pi)\) and

\[
\frac{d\mu_j(\lambda)}{d\lambda} = \sum_{l=0}^{\nu-1} \hat{g}_l(\lambda + 2\pi l/\nu) \overline{\hat{g}}_{-l}(\lambda + 2\pi l/\nu) \quad (44)
\]

where

\[
\hat{g}_j(\lambda) = \frac{1}{\nu} \sum_{n=0}^{\nu-1} g_n(\lambda) e^{2\pi j n/\nu} \quad (45)
\]

and \( g_n(\lambda) \) are expressed in terms of the PARMA parameters.

Questions arising from this paper include the connection between condition (7) for a PC solution to conditions I and II given in the paper and and comparison or utility of result of Theorem 2 (of the paper) with the results of Sakai [38].

In [40] (2010) Tesfayea, Anderson and Meerchaert develop the asymptotics for the Fourier-PARMA methods introduced in [34]. Assumptions used are finite second moments, causality (see (10) and invertibility (see (11), finite fourth moments and, in addition, some cases of infinite fourth moments; we will discuss only the finite fourth moment case here. In addition the spectral density condition (see item (6) in our discussion of [4]) on the equivalent VARMA process is assumed along with using the function \( k(N) \) such that \( k(N) \to \infty \) and \( k^{3/2}/N \to 0 \) as \( N \to \infty \).

The results of the paper are: (1) (Thm 1) For PMA sequence, set \( \hat{\psi}_u(i) = \hat{\psi}_u(i - k) \), and denote \( \psi(u) = [\hat{\psi}_u(0), \hat{\psi}_u(1), \ldots, \hat{\psi}_u(\nu - 1)]' \) as a vector of IMA coefficients with \( u \) designating lag (term in the IMA), then the joint asymptotic result for two vectors \( \psi(u), \psi(u') \) is obtained; namely, as \( k \to \infty \),

\[
N^{1/2} \begin{bmatrix} \hat{\psi}(u) - \psi(u) \\ \hat{\psi}(u') - \psi(u') \end{bmatrix} \Rightarrow \mathcal{N}(0, \begin{bmatrix} V_{uu} & V_{uu'} \\ V_{u'u} & V_{u'u'} \end{bmatrix})
\]
where $V_{u'u'}$ matrices are given in terms of $\sigma(0), \sigma(1), \ldots, \sigma(\nu - 1)$ and the $\{\psi_n(i), n = 1, 2, \ldots, \min(u, u'), i = 0, 1, \ldots, \nu - 1\}$.

(2) For two times $(i,i')$ and two lags $(u, u')$ there is Corr. 1: Under the conditions of Thm. 1, for $0 \leq i, i' \leq \nu - 1, 0 \leq u \leq u'$, as $k \to \infty$,

$$N^{1/2} \begin{bmatrix} \hat{\psi}_u(i) - \hat{\psi}_u(i) \\ \hat{\psi}_{u'}(i') - \hat{\psi}_{u'}(i') \end{bmatrix} \Rightarrow \mathcal{N} \left( 0, \begin{bmatrix} V_{iui'u'} & V_{iui'u'} \\ V_{iui'u'} & V_{iui'u'} \end{bmatrix} \right)$$

where $V_{iui'u'} = \sum_{n=1}^{u} \sigma_{i-u+n}^{2} \sigma_{i-u}^{2} \psi_{u-n}(i) \psi_{u-n}(i - u + u')$ if $k = i + u' - u \bmod \nu$ and $V_{iui'u'} = 0$ otherwise. Note this provides a subset of the elements of $V_{uu'}$.

(3) Complete difference equations for $\psi$-weights (see [34] and (43) above for the first few), are given, utilizing the constraints of PARMA(p,q) and hence producing

$$\psi_j(t) - \sum_{k=1}^{j} \phi_k(t) \psi_{j-k}(t-k) = 0 \quad j \geq \max(p, q + 1)$$
$$\psi_j(t) - \sum_{k=1}^{p} \phi_k(t) \psi_{j-k}(t-k) = -\theta_j(t) \quad 0 \leq j \leq \max(p, q + 1).$$

Three specific cases are addressed, general PMA(q), general PAR(p), and PARMA(1,1), since the solution is difficult for the general PARMA(p,q) case. Continuing for the PARMA(1,1) case, under assumptions of Thm. 1, (4) asymptotic normality is obtained (Thm. 2) for $\hat{\phi}_1$, specifically $N^{1/2}(\hat{\phi}_1 - \phi_1) \Rightarrow \mathcal{N}(0, Q)$ where $\phi_1 = [\phi_1(0), \phi_1(1), \ldots, \phi_1(\nu - 1)]$, similarly $\phi_1 = [\hat{\phi}_1(0), \hat{\phi}_1(1), \ldots, \hat{\phi}_1(\nu - 1)]$, and $\nu \times \nu$ matrix $Q$ is given in terms of parameters $\{\psi_n(i)\}$. In particular (Corr. 2), (5) $N^{1/2}(\hat{\phi}_1(i) - \phi_1(i)) \Rightarrow \mathcal{N}(0, W_{\hat{\phi}_1})$ for $0 \leq i \leq \nu - 1$, and $W_{\hat{\phi}_1}$ is given in terms of $\sigma(i)s$ and $\psi_j(n)s$. Still under the same assumptions, (6) asymptotic normality is obtained for $\hat{\theta}_1$, specifically $N^{1/2}(\hat{\theta}_1 - \theta_1) \Rightarrow \mathcal{N}(0, S)$ where $\theta_1 = [\theta_1(0), \theta_1(1), \ldots, \theta_1(\nu - 1)]'$, similarly $\hat{\theta}_1 = [\hat{\theta}_1(0), \hat{\theta}_1(1), \ldots, \hat{\theta}_1(\nu - 1)]'$, and $\nu \times \nu$ matrix $S$ is given in terms of parameters $\psi_j(n)$. In particular (Corr. 3), (7) $N^{1/2}(\hat{\theta}_1 - \theta_1) \Rightarrow \mathcal{N}(0, W_{\theta_1})$ for $0 \leq i \leq \nu - 1$, and $W_{\theta_1}$ is given in terms of $\sigma(i)s$ and $\psi_j(n)s$.

To begin the treatment of asymptotics for the discrete Fourier transforms (DFTs) of $\psi_u(t)$, define $f(j)$ as

$$f(u) = \begin{cases} [c_0(u), c_1(u), s_1(u), \ldots, c_{(\nu-1)/2}(u), s_{(\nu-1)/2}(u)]' & (\nu \text{ odd}) \\ [c_0(u), c_1(u), s_1(u), \ldots, s_{(\nu/2-1)}(u), c_{\nu/2}(u)]' & (\nu \text{ even}) \end{cases}$$

where

$$\psi_u(t) = c_0(u) + \sum_{r=1}^{\lfloor \nu/2 \rfloor} c_r(u) \cos 2\pi rt/\nu + s_r(u) \sin 2\pi rt/\nu.$$
(see [40], Eq. 57). The estimates \( \hat{f}(u) \) are determined by the same transformation 
\( \hat{f}(u) = LPU \hat{\psi}(u) \). The final result, using the continuous mapping theorem applied to \( B = LPU \) is that (8) 
\[ N^{1/2}[\hat{f}(u) - f(u)] \Rightarrow \mathcal{N}(0, R_V) \]
where \( R_V = BV_{uu}B' \), where \( V_{uu} \) is from item (1) above. Using this result under the null hypothesis that the process is stationary, so that \( \psi_u(t) = \psi(t) \) and \( \sigma(t) = \sigma \), the covariance \( R_V \) becomes diagonal, resulting in (9) 
\[ N^{1/2}[\hat{c}_m(u) - \mu_m(u)] \mbox{ and } N^{1/2}[\hat{s}_m(u) - \mu_m(u)] \Rightarrow \mathcal{N}(0, R_m(u)) \]
for all \( u \geq 1 \) where \( \mu_m(u) = \psi(u), m = 0 \) and \( \mu_m(u) = 0, m > 0 \). The limiting variances are
\[ R_m(u) = \begin{cases} 
\nu^{-1} \eta_V(u) & (m = 0 \mbox{ or } \nu/2) \\
2\nu^{-1} \eta_V(u) & (0 < m < \nu/2) 
\end{cases} \]
where \( \eta_V(u) = \sum_{n=1}^{\nu} \psi_n^2 \), meaning that the elements of \( f(u) \) are asymptotically independent and the nonzero frequencies have asymptotic mean of zero. Since the asymptotic means are zero, the asymptotic variances can be used to construct tests for presence of a nonzero mean, interpreted as presence of a periodic component. Expressions for the test thresholds, with Bonferroni correction, are given in [40].

Finally, addressing the PARMA(1, 1) case, the Fourier series representation for \( \theta_u(t) \) is
\[ \theta_u(t) = c_{a0}(u) + \sum_{r=1}^{\lfloor \nu/2 \rfloor} c_{ar}(u) \cos 2\pi rt/\nu + s_{ar}(u) \sin 2\pi rt/\nu. \]  
(50)
with \( c_{br}(u), s_{br}(u) \) and \( c_{dr}, s_{dr} \) similarly corresponding to Fourier coefficients for \( \phi_u(t) \) and \( \sigma(t) \). Denoting \( \theta_1 = [\theta_1(0), \theta_1(1), \ldots, \theta_1(\nu-1)]' \), \( \phi_1 = [\phi_1(0), \phi_1(1), \ldots, \phi_1(\nu-1)]' \), and \( \sigma = [\sigma(0), \sigma(1), \ldots, \sigma(\nu-1)]' \), the same process used above leads to \( f_{\theta_1} = LPU \theta_1 \) and \( f_{\phi_1} = LPU \phi_1 \) and then to (10) (Thm. 6)
\[ N^{1/2}[\hat{f}_{\theta_1} - f_{\theta_1}] \Rightarrow \mathcal{N}(0, R_{\theta}) \]  
(51)
\[ N^{1/2}[\hat{f}_{\phi_1} - f_{\phi_1}] \Rightarrow \mathcal{N}(0, R_{\phi}) \]  
(52)
where, as above, \( R_{\theta} = BS\theta B' \) and \( R_{\phi} = BQ\phi B' \) with \( S \) the asymptotic covariance of \( N^{1/2}(\hat{\theta}_1 - \theta_1) \) and \( S \) the asymptotic covariance of \( N^{1/2}(\hat{\phi}_1 - \phi_1) \).

Finally, (Thm. 7) (11) for mean standardized PARMA(1,1), under null hypothesis of stationarity, the elements of \( f_{\theta_1} \) are asymptotically independent and the nonzero frequencies have asymptotic mean of zero. Similarly, under the same hypotheses, (12) the elements of \( f_{\phi_1} \) are asymptotically independent and the nonzero frequencies have asymptotic mean of zero. As before, since the asymptotic means are zero, the asymptotic variances can be used to construct tests for presence of significant periodic components, and expressions for the Bonferroni corrected thresholds are given in the paper.

An application to streamflow time series is the subject of the earlier paper [34].

Subsequent to [40] further progress in PARMA time series analysis includes the following. Franq, Roy and Saidi [12] address the asymptotics for various least squares
estimation methods for PARMA models in which the shocks are assumed to be only uncorrelated; these are denoted as weak PARMA whereas the case with IID shocks are denoted as strong. They make the important point that the asymptotic covariance can be much different for weak PARMA models than in the case of strong ones. Anderson, Meerschaert and Zhang [2] address forecasting errors for PARMA models, provide a formula for the asymptotic error variance, thus permitting prediction intervals to be computed in the Gaussian case. The innovation algorithm and often used method of Ansley [6] are employed along with results from Lund and Basawa [26]. Future research using reduced models, including those realized by Fourier parameterization, is suggested. Tang and Shao [39] address parameter estimation for a series composed of the sum of a PAR and a non-linear trend. A two-step estimation method is proposed, essentially estimation of the PAR followed by estimation of the trend residual.

Two simulations

In this section we present/document two recent simulations that illustrate the estimation of Fourier PARMA parameters and suggests avenues for further study. The first is for the PARMA(1,1) system described by Meerschaert, Anderson and Tesfaye in [34], and the second is for PAR systems when the period \( \nu \) equals the observation size \( N \) where \( N \) is large. The goals of the simulation were (1) to determine if the parameter estimates were close to Normal and how this depends on sample size, and (2) to determine the sample means and variances and how they depend on true parameter value and on sample size.

The R package perARMA [11] contains the code to run these simulations.

PARMA(1,1) Simulation

To illustrate the ML method for Fourier parameterized PARMA, we conducted a \( N = 100 \) sample simulation of a \( PARMA_{1,2}(1, 1) \) which was determined in [34] to provide a good model for Fraser River streamflow measurements. The true parameters are given in the following expressions.

\begin{align*}
\phi_1(t) &= 0.35 + 0.25 \cos \frac{2\pi t}{\nu} + 0.35 \sin \frac{2\pi t}{\nu} + 0.45 \cos \frac{4\pi t}{\nu} - 0.15 \sin \frac{4\pi t}{\nu} \\
\theta_1(t) &= 0.35 + 0.15 \cos \frac{2\pi t}{\nu} + 0.40 \sin \frac{2\pi t}{\nu} + 0.25 \cos \frac{4\pi t}{\nu} + 0.35 \sin \frac{4\pi t}{\nu} \\
\sigma(t) &= 1.
\end{align*}

We used the maximum likelihood method that is based on the method of Ansley [6] and adapted to PARMA by Vecchia [45, 46], Li and Hui [25] and alternatively formulated by Lund and Basawa [26]. Geneally, given the data \( X_{t_0:n} = (X_{t_0}, X_{t_0+1}, \ldots, X_{t_0+n-1})' \) and Fourier coefficients \( A, B \) from (8),(9), the likelihood \( L(\Phi(A), \Theta(B)|X_{t_0:n}) \) is computed in two main steps: (S1) first, using (8), \( W \) is computed according to (17) and then (18) becomes

\[ L(\Phi(A), \Theta(B)|X_{t_0:n}) = L(\Phi(A), \Theta(B)|W_{t_0:n}). \] (53)
The usual simplification leading to Vecchia’s approximate likelihood is utilized, basing the likelihood calculation on \( W = W_{t_0 + m,n-m} \), which is seen to be a PMA(q) sequence given by the right side of (4) and whose covariance \( R_W \) is simply computed. See [25, 26]. So then (S2) the likelihood \( L(\Phi(A), \Theta(B)|W_{t_0 + m,n-m}) \) is computed and used to approximate \( L(\Phi(A), \Theta(B)|X_{t_0,n}) \). This process, S1 followed by S2, is placed in a loop that seeks to maximize \( L \) as a function of \( A, B \).

As we have seen above, Anderson et.al. in their collection of interesting papers [3, 34, 40] give conditions under which the estimators for Fourier coefficients of a PARMA(1, 1) converge to Normal as the sample length increases. In this study we performed \( NSAMP = 100 \) simulations for each sample length \( NLEN \) in the set \( \{300, 600, 1200, 2400\} \) corresponding to 25, 50, 100, 200 periods of length 12. We note that the simulation reported in [34] was of length 6000, or 500 periods of length 12.

For each value of \( NLEN \) we checked the closeness of the sample distribution to Normality by performing the Lilliefors test, a Kolmogorov-Smirnov type of test in which the null is Normal with parameters estimated from the data. Figures 1 and 2 are histograms from the estimates of parameters \( a_{21} \) and \( b_{21} \) for the simulation \( NSAMP = 100, NLEN = 600 \). The Lilliefors p-values are both \( \geq 0.5 \), indicating that the Normality is strongly not rejected. Table 1 summarizes the true parameter values, the sample means and deviations for the four values of \( NLEN \). An asterisk placed on the sample mean indicates a Lilliefors p-value that is less than 0.05, indicating a rejection of Normality. Only a few of the estimated parameters indicate rejection of Normality.

For each parameter, the dependence of \( \hat{\sigma} \) on \( NLEN \) can be seen by fitting a straight line, \( y = mx + b \), to the pairs \( (NLEN, \hat{\sigma}) \), where both \( NLEN \) and \( \hat{\sigma} \) are transformed to a log scale, so the expected \( (NLEN)^{-1/2} \) dependence becomes \( m = -1/2 \). Figures 3 and 4 illustrate this fitting, producing \( m = -0.525 \) and \( m = -0.552 \) in the two cases; the dotted lines connect the observed data and the solid line is the result of the ordinary least squares fit. The empirical dependence on \( NLEN \) is slightly steeper than the expected \( m = -1/2 \). The computed standard errors for parameters \( a_{21} \) and \( b_{21} \) (corresponding to parameters labeled \( c_b1 \) and \( c_a1 \) in [34]) for \( NLEN = 6000 \) are
Table 1  Summary of simulation results for $NSAMP = 100$ showing $\hat{\mu}$, $\hat{\sigma}$ and * denotes Lilliefors p-value < 0.05.

<table>
<thead>
<tr>
<th>true values</th>
<th>mean 300</th>
<th>std</th>
<th>mean 600</th>
<th>std</th>
<th>mean 1200</th>
<th>std</th>
<th>mean 2400</th>
<th>std</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{11}$</td>
<td>0.35</td>
<td>0.35</td>
<td>0.098</td>
<td>0.36*</td>
<td>0.058</td>
<td>0.35</td>
<td>0.041</td>
<td>0.35</td>
</tr>
<tr>
<td>$a_{21}$</td>
<td>0.25</td>
<td>0.28</td>
<td>0.113</td>
<td>0.25</td>
<td>0.087</td>
<td>0.25</td>
<td>0.055</td>
<td>0.25</td>
</tr>
<tr>
<td>$a_{31}$</td>
<td>0.35</td>
<td>0.34</td>
<td>0.145</td>
<td>0.33</td>
<td>0.092</td>
<td>0.34</td>
<td>0.061</td>
<td>0.34</td>
</tr>
<tr>
<td>$a_{41}$</td>
<td>0.45</td>
<td>0.46</td>
<td>0.133</td>
<td>0.44</td>
<td>0.077</td>
<td>0.45</td>
<td>0.054</td>
<td>0.45</td>
</tr>
<tr>
<td>$a_{51}$</td>
<td>-0.15</td>
<td>-0.16</td>
<td>0.122</td>
<td>-0.15</td>
<td>-0.072</td>
<td>-0.16</td>
<td>-0.050</td>
<td>-0.15</td>
</tr>
<tr>
<td>$a_{61}$</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.095</td>
<td>-0.01</td>
<td>-0.058</td>
<td>-0.01</td>
<td>0.040</td>
<td>0.00</td>
</tr>
<tr>
<td>$a_{71}$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.100</td>
<td>0.00</td>
<td>0.063</td>
<td>0.00</td>
<td>0.040</td>
<td>0.00</td>
</tr>
<tr>
<td>$c_{11}$</td>
<td>1.00</td>
<td>0.98</td>
<td>0.039</td>
<td>0.99</td>
<td>0.030</td>
<td>1.00</td>
<td>0.018</td>
<td>1.00</td>
</tr>
<tr>
<td>$b_{11}$</td>
<td>0.35</td>
<td>0.35*</td>
<td>0.101</td>
<td>0.34</td>
<td>0.062</td>
<td>0.35</td>
<td>0.042</td>
<td>0.35</td>
</tr>
<tr>
<td>$b_{21}$</td>
<td>0.15</td>
<td>0.13</td>
<td>0.115</td>
<td>0.14</td>
<td>0.079</td>
<td>0.14</td>
<td>0.051</td>
<td>0.15</td>
</tr>
<tr>
<td>$b_{31}$</td>
<td>0.40</td>
<td>0.40</td>
<td>0.164</td>
<td>0.42*</td>
<td>0.104</td>
<td>0.41*</td>
<td>0.066</td>
<td>0.41</td>
</tr>
<tr>
<td>$b_{41}$</td>
<td>0.25</td>
<td>0.23</td>
<td>0.152</td>
<td>0.25</td>
<td>0.080</td>
<td>0.25</td>
<td>0.055</td>
<td>0.25</td>
</tr>
<tr>
<td>$b_{51}$</td>
<td>0.35</td>
<td>0.36</td>
<td>0.139</td>
<td>0.35</td>
<td>0.087</td>
<td>0.36</td>
<td>0.065</td>
<td>0.35*</td>
</tr>
<tr>
<td>$b_{61}$</td>
<td>0.00</td>
<td>0.02</td>
<td>0.130</td>
<td>0.02</td>
<td>0.083</td>
<td>0.01</td>
<td>0.057</td>
<td>0.00</td>
</tr>
<tr>
<td>$b_{71}$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.128</td>
<td>-0.01</td>
<td>-0.078</td>
<td>-0.01</td>
<td>-0.045</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure 3: parameter $a_{21}, NLEN = 300, 600, 1200, 2400$, $NSAMP = 100$ (solid) and OLS fit $(\hat{m}, \hat{b}) = (-0.525, 0.368)$ (dashed).

Figure 4: parameter $b_{21}, NLEN = 300, 600, 1200, 2400$, $NSAMP = 100$ (solid) and OLS fit $(\hat{m}, \hat{b}) = (-0.552, 0.425)$ (dashed).
The empirical values for $NLEN = 2400$ are extrapolated by the factor $\left[\frac{(2400)}{(6000)}\right]^{1/2}$ to give .0247 and .0234, showing reasonable agreement with the computed values. In order to show the variability of parameter estimates when parameter values are zero, we added $a_{61}, a_{71}, b_{61}, b_{71}$ to the list of parameters to be estimated but the true values of these parameters were all zero. Figure 5 shows the boxplots of all 15 parameters estimated. The ability to perceive non-nullity of parameters is visually clear.

**Numerical Estimation of Fourier-PAR Models when $N = \nu$.**

The usual method for estimation of PAR parameters is derived through the Yule-Walker equations and use of the presumed property of periodic correlation, see (3), in the observed time series. (Pagano[35], Vecchia[45], maybe [21, 23]). This solution relies on the length $N$ of the observed series $X_t$ being large enough relative to the period $\nu$ that there are sufficiently many pairs of times $(t_1 + j\nu, t_2 + j\nu)$ to allow satisfactory estimation of $R(t_1, t_2)$ by use of the PC property. But sometimes there may be only very few periods (perhaps just one) available for estimating the parameters of an hypothesized PAR model. If we use the direct parameterization given in (4), then even for a PAR(1) model there are $N$ parameters to estimate (the values $\{\phi_1(t), t = 1, 2, \ldots \nu\}$) from a series of length $N = \nu$. But the Fourier parameterization, to the extent it is an accurate model, means a potentially much smaller parameter set may give an adequate model (fit) of the data. For example, if a PAR(1) model has only the constant term and the fundamental harmonic term, then there only three non-zero parameters appearing in (8), namely $a_{1,1}, a_{1,2}, a_{1,3}$. When the ratio of $N$ to the number of parameters is large, it is no surprise that the parameters can be successfully estimated from the series.
In this section we give a brief demonstration of a procedure that can effectively estimate the parameters of a PAR model that is parameterized by the Fourier coefficients of $\phi_j(t)$ as in (8). The procedure can be motivated by considering the negative log of the likelihood $W = W_{t_0+m:n-m}$ with $m = p$, $q = 0$ and the shock weights are constant and unity, $\theta_0(t) = 1$, $t = 0, 1, \ldots, \nu - 1$; for then $R_W = I$ and $\det(R_W) = 1$, leading finally to

$$- \log L(\Phi(A), \Theta(B) | W) = \frac{N - p}{2} + \frac{1}{2} \sum_{t=p+1}^{N} \left[ X_t - \sum_{j=1}^{p} \phi_j A_j(t) X_{t-j} \right]^2$$

(54)

where $A = \{A_1, A_2, \ldots, A_p\}$,

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

and we can take $N = \nu$. So under these conditions, the maximization of likelihood is achieved by minimization of the rightmost term in (54), which we denote as $Q(A)$ and
which is the mean square error in the ordinary least squares (OLS) sense. Even without this connection to the maximum likelihood solution, the best OLS fit is of interest. We have coded the minimization of $Q(A)$ into a procedure entitled parmsef. As in the more general procedure parmaf, the user can easily specify which frequencies in the Fourier series of $\phi_j(t)$ are to be estimated. The following figures document a simulation of NSAMP=100 realizations of a PAR(2) sequence having the parameters

$$\phi_1(t) = 1.1 + 0.6 \cos \frac{2\pi t}{\nu}$$
$$\phi_2(t) = -0.3450 - 0.330 \cos \frac{2\pi t}{\nu} + 0.0 \sin \frac{2\pi t}{\nu}$$
$$-0.045 \cos 2 \cdot \frac{2\pi t}{\nu} + 0.0 \sin 2 \cdot \frac{2\pi t}{\nu}$$

(55)

so the pertinent true $a$ values are $a_{1,1} = 1.1$, $a_{1,2} = 0.6$, $a_{1,3} = 0.0$, $a_{2,1} = -0.345$, $a_{2,2} = -0.330$, $a_{2,3} = 0.0$, $a_{2,4} = -0.045$, $a_{2,5} = 0.0$. All remaining $a$ values are null.

Figures 6 and 7 show the resulting empirical distributions of the parameter estimates for $a_{1,1}$ and $a_{1,2}$. As indicated at the top of each figure, the p-value of the Lilliefors test for normality exceeds 0.5, meaning there is strong evidence in favor of the null (normality). Figures 8 and 9 illustrate that for $\nu = N$ in the range 512-4096, the $\hat{\sigma}$ for $a_{1,1}$ and $a_{1,2}$ both diminish approximately as $N^{-1/2}$ as discussed above.

Finally, the boxplots of Figure 10 show the perceptability of non-zero parameters when there are some null parameters included in the fit. The only non-zero parameters, $a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}, a_{2,4}$ are defined by (55) and text following it.

Figure 10: Boxplots of parmsef Parameter Estimates, NSAMP = 100 realizations, $N = \nu = 1024$, 17 parameters estimated; the only non-null parameters those with indices 1,2,10,11,13.
Conclusion

The parameterization of PARMA models by Fourier coefficients can substantially reduce the number of parameters to be estimated while providing satisfactory fit to many instances of PARMA data. It can be expected to be especially effective when data are derived from physical systems in which the periodic changes are slowly varying throughout the period, for example in data connected to daily, monthly or yearly variations caused by the earth’s motions. Substantial progress has been made on this topic as summarized in the preceding paragraphs, but especially we note the series of papers by Vecchia, Anderson, Lund, Meerschaert and Tesfaye [4, 3]. Some remaining topics of interest include: (1) extending the methods of [34] beyond the PARMA(1,1) model, (2) new methods for identification of frequencies to be included in the model, (3) improved numerical methods to facilitate parameter estimation, maximum likelihood and otherwise, (4) use of simulation and bootstrapping, (5) $N = \nu$ problems, (6) multiple rhythms (day, month, year) and differing models operating at each period, (7) time series analysis of deficient rank PARMA.

Acknowledgement

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 655394.

References


