Parsimonious time series modeling for high frequency climate data

Paul L. Anderson[*], Farzad Sabzikar[†], and Mark M. Meerschaert[‡]

June 7, 2020

Abstract

Climate data often provides a periodically stationary time series, due to seasonal variations in the mean and covariance structure. Periodic ARMA models, where the parameters vary with the season, capture the nonstationary behavior. High frequency data collected weekly or daily results in a large number of model parameters. In this paper, we apply discrete Fourier transforms to the parameter vectors, and develop a test for the statistically significant harmonics. An example of daily high temperatures illustrates the method, whereby a periodic autoregressive model with 1095 parameters is reduced to a parsimonious 12 parameter version without any apparent loss of fidelity.

Keywords: Periodic autoregressive moving average, discrete Fourier transform, climate data, parsimony.

1 Introduction

A stochastic process $\{\tilde{X}_t\}$ is called periodically stationary in the wide sense if $\mu_t = E\tilde{X}_t$ and $\gamma_t(h) = \mathrm{Cov}(\tilde{X}_t, \tilde{X}_{t+h})$ for $h = 0, \pm1, \pm2, \ldots$ are all periodic functions of time t with the same period $\nu > 1$. The process is stationary if all of the parameter functions are constant functions of the season. Periodically stationary processes manifest themselves in such fields as hydrology, climatology, geophysics, and economics, where the observed time series are characterized by seasonal variations in both the mean and covariance structure. Periodic ARMA time series, which allow the model parameters in the classical ARMA model to vary with the season, are useful for modeling such behavior. A periodic ARMA process $\{\tilde{X}_t\}$ with period ν, denoted by $\mathrm{PARMA}_\nu\,(p, q)$, takes the form

$$X_t - \sum_{j=1}^{p} \phi_t(j) X_{t-j} = \varepsilon_t - \sum_{j=1}^{q} \theta_t(j) \varepsilon_{t-j} \tag{1}$$

where $X_t = \tilde{X}_t - \mu_t$, and $\{\varepsilon_t\}$ is a sequence of noise random variables with mean zero and scale σ_t such that $\{\delta_t = \sigma_t^{-1}\varepsilon_t\}$ is independent and identically distributed. The notation in (1) is consistent with that of Box and Jenkins (1976). The model parameters $\phi_t(j)$, $\theta_t(j)$, μ_t, and σ_t are all periodic with the same period $\nu \geq 1$. Periodic time series models and their practical applications are discussed for example in Adams and Goodwin (1995), Anderson and Vecchia (1993), Anderson and Meerschaert (1997,1998), Anderson, Meerschaert and Veccia (1999), Basawa, Lund and Shao (2004), Dudek, Hurd and Wojtowicz (2016), Franses and Paap (2011), Jones and Brelsford (1967), Lund and Basawa (1999,2000), Mei, Shao and Liu (2017), Pagano (1978), Salas, Boes, and Smith (1982), Salas, Tabios, and Bartolini (1985), Shao and Lund (2004), Troutman (1979), Vecchia (1985a,1985b), Vecchia and Ballerini (1991), Ula (1990,1993), Ula and Smadi (1997,2003) and Tesfaye, Meerschaert, and Anderson (2005).

[*]Department of Mathematics and Computer Science, Albion College, USA.
[†]Department of Statistics, Iowa State University, USA.
[‡]Department of Statistics and Probability, Michigan State University, USA.

High-frequency PARMA models are plagued with a very large number of parameters. Discrete Fourier transforms (DFT) were applied by Anderson, Tesfaye and Meerschaert (2007) to develop a parsimonious model of weekly time series in hydrology. Asymptotics of the estimated autoregressive and moving average parameters in (1) were used to derive hypothesis tests for the DFT coefficients, to identify the statistically significant harmonics. In this paper, we complete this program by deriving the asymptotic distributions of the sample mean μ_t, autocovariance $\gamma_t(\cdot)$, and noise-variance σ_t^2. These results are then applied to test for the significant DFT harmonics for all the PARMA model parameters. The effectiveness of this modeling approach is illustrated using a time series of daily high temperatures.

In Section 2, asymptotic results are derived for a two-sided periodic moving-average process of infinite order. In particular, we derive the asymptotic distributions of moment estimates of the periodic mean μ_t, and the periodic autocovariance $\gamma_t(\cdot)$. In Section 3, asymptotic results are derived for estimators of the periodic noise-variance function σ_t^2, using the innovations algorithm. Section 4 develops a test to determine which of the real DFT harmonics are statistically significantly different from zero. We illustrate in detail the computations required for this test, in the case of a PARMA$_\nu(1,0)$ time series, i.e., a periodic autoregressive model of order 1. In Section 5, we apply the results of this paper to a time series of daily maximum temperatures. The data comes from a sophisticated NARCCAP climate model. A simple periodic autoregressive model of order 1 with Gaussian noise is found to be adequate to fit this data. Then our Fourier-PARMA methods are applied to reduce the $3 \times 365 = 1095$ model parameters to just 12 DFT harmonics, without any apparent loss of model fidelity.

2 Asymptotics of the periodic mean and autocovariance

We assume throughout this section that $\{X_t\}$ is a two-sided infinite order periodic moving average

$$X_t = \sum_{j=-\infty}^{\infty} \psi_t(j)\varepsilon_{t-j} \tag{2}$$

where $X_t = \tilde{X}_t - \mu_t$, $\psi_t(0) = 1$, and:

(i) $\sum_{j=-\infty}^{\infty} |\psi_t(j)| < \infty$ for all t; and

(ii) $\{\varepsilon_t\}$ is a sequence of noise variables with mean 0 and standard deviation σ_t such that $\{\delta_t = \sigma_t^{-1}\varepsilon_t\}$ is iid(0,1), i.e., independent and identically distributed with mean zero and variance 1.

Suppose we have N years of data given by $\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_{N\nu-1}$. Define the sample mean vector, $\hat{\mu} = (\hat{\mu}_0, \hat{\mu}_1, \dots, \hat{\mu}_{\nu-1})'$ where

$$\hat{\mu}_i = N^{-1} \sum_{j=0}^{N-1} \tilde{X}_{j\nu+i}. \tag{3}$$

Then, $E(\hat{\mu}_i) = \mu_i$, $i = 0, 1, \dots, \nu - 1$.

Proposition 2.1 *Let $\tilde{X}_t = \mu_t + \sum_{j=-\infty}^{\infty} \psi_t(j)\varepsilon_{t-j}$ where ε_t is as in (2). Then,*

$$\lim_{N\to\infty} N\mathrm{Var}(\hat{\mu}) = \Sigma_\mu \tag{4}$$

where the $(i,j)^{th}$ element of Σ_μ, $i, j = 0, 1, \dots, \nu - 1$, is

$$(\Sigma_\mu)_{ij} = \sum_{n=-\infty}^{\infty} \gamma_i(n\nu + j - i) \tag{5}$$

with $\gamma_i(\ell) = \mathrm{Cov}(\tilde{X}_{t\nu+i}, \tilde{X}_{t\nu+i+\ell})$.

PROOF. Note that

$$
\begin{aligned}
N\mathrm{Cov}(\hat{\mu}_i, \hat{\mu}_j) &= N[E(\hat{\mu}_i\hat{\mu}_j) - \mu_i\mu_j] \\
&= [N^{-1}\sum_{t=0}^{N-1}\sum_{u=0}^{N-1} E(\tilde{X}_{t\nu+i}\tilde{X}_{u\nu+j})] - N^{-1}\sum_{t=0}^{N-1}\sum_{u=0}^{N-1}\mu_i\mu_j \\
&= \sum_{|n|<N}\left(1 - \frac{|n|}{N}\right)\gamma_i(n\nu + j - i).
\end{aligned} \tag{6}
$$

Since $\sum |\gamma_i(k)| < \infty$ by the absolute summability of $\{\psi_t(j)\sigma_{t-j}\}$, the dominated convergence theorem gives

$$
\lim_{N\to\infty} N\mathrm{Cov}(\hat{\mu}_i, \hat{\mu}_j) = \sum_{n=-\infty}^{\infty}\gamma_i(n\nu + j - i),
$$

which finished the proof. □

Proposition 2.2 *If $\tilde{X}_t$ is the periodic moving average*

$$
\tilde{X}_t = \mu_t + \sum_{j=-\infty}^{\infty}\psi_t(j)\varepsilon_{t-j}, \tag{7}
$$

then $\hat{\mu}$ is $AN(\mu, N^{-1}\Sigma_\mu)$ where $\mu = (\mu_0, \mu_1, \ldots, \mu_{\nu-1})'$, meaning that $N^{1/2}(\hat{\mu} - \mu) \Rightarrow \mathcal{N}(0, \Sigma_\mu)$.

PROOF. Consider the truncated sequence

$$
\tilde{X}_{t,m} = \mu_{t,m} + \sum_{j=-m}^{m}\psi_t(j)\varepsilon_{t-j}.
$$

Define $\hat{\mu}_{i,m} = N^{-1}\sum_{t=0}^{N-1}\tilde{X}_{t\nu+i,m}$ and the ν-dimensional vector process, $Y_{t,m}$ by

$$
Y_{t,m} = (\tilde{X}_{t\nu,m}, \tilde{X}_{t\nu+1,m}, \ldots, \tilde{X}_{t\nu+\nu-1,m})'.
$$

For λ in $\mathbb{R}^\nu$ we have that $\{\lambda' Y_{t,m}\}$ is a strictly stationary $\{1 + \lfloor\frac{2m-1}{\nu}\rfloor\}$-dependent sequence. Since $N^{-1}\sum_{t=0}^{N-1} Y_{t,m} = \hat{\mu}_m$ where $\hat{\mu}_m = (\hat{\mu}_{0,m}, \hat{\mu}_{1,m}, \ldots, \hat{\mu}_{\nu-1,m})'$, by the Cramer-Wold Device it suffices to show that $\lambda'\hat{\mu}_m$ is AN$(\lambda'\mu_m, N^{-1}\lambda'(\Sigma_\mu)_m\lambda)$ for all λ such that $\lambda'(\Sigma_\mu)_m\lambda > 0$. Noting that $E(\hat{\mu}_m - \mu_m) = \underline{0}$, by Proposition 2.1, we have $\lim_{N\to\infty} N\mathrm{Var}[\lambda'(\hat{\mu}_m - \mu_m)] = \lambda'(\Sigma_\mu)_m\lambda > 0$ where $\mu_m = (\mu_{0,m}, \mu_{1,m}, \ldots, \mu_{\nu-1,m})'$, and the $(i,j)th$ element of $(\Sigma_\mu)_m$ is

$$
(\Sigma_\mu)_{ij,m} = \sum_{n=-\infty}^{\infty}\gamma_{i,m}(n\nu + j - i),
$$

where $\gamma_{i,m}(n\nu + j - i) = \mathrm{Cov}(\tilde{X}_{t\nu+i,m}, \tilde{X}_{(t+n)\nu+j,m})$. By the Central Limit Theorem for strictly stationary m-dependent sequences (Brockwell and Davis, 1991, Theorem 6.4.2), $\lambda'(\hat{\mu}_m - \mu_m)$ is $AN(\underline{0}, N^{-1}\lambda'(\Sigma_\mu)_m\lambda)$.

Since $(\Sigma_\mu)_m \to \Sigma_\mu$ as $m \to \infty$, it suffices to show, by Proposition 6.3.9 of Brockwell and Davis (1991), that for every $\varepsilon > 0$

$$
\lim_{m\to\infty}\limsup_{N\to\infty} P[|N^{1/2}(\hat{\mu}_i - \mu_i) - N^{1/2}(\hat{\mu}_{i,m} - \mu_{i,m})| > \varepsilon] = 0. \tag{8}
$$

By Chebyschev's inequality,

$$
P[|N^{1/2}(\hat{\mu}_i - \mu_i) - N^{1/2}(\hat{\mu}_{i,m} - \mu_{i,m})| > \varepsilon] \le \varepsilon^{-2}\mathrm{Var}[N^{1/2}(\hat{\mu}_i - \mu_i) - N^{1/2}(\hat{\mu}_{i,m} - \mu_{i,m})].
$$

The right-hand side of the above inequality is

$$\varepsilon^{-2}\left[\mathrm{Var}[N^{1/2}(\hat{\mu}_i-\mu_i)]+\mathrm{Var}[N^{1/2}(\hat{\mu}_{i,m}-\mu_{i,m})]-2\mathrm{Cov}\left(N^{1/2}(\hat{\mu}_i-\mu_i),N^{1/2}(\hat{\mu}_{i,m}-\mu_{i,m})\right)\right]. \quad (9)$$

Taking the limit as $N\to\infty$ in (9) yields $\varepsilon^{-2}[(\Sigma_\mu)_{ii}+(\Sigma_\mu)_{ii,m}-2\lim_{N\to\infty}N\mathrm{Cov}(\hat{\mu}_i,\hat{\mu}_{i,m})]$ where $(\Sigma_\mu)_{ii}$ is the $(i,i)th$ element of Σ_μ and $(\Sigma_\mu)_{ii,m}$ is the $(i,i)th$ element of $(\Sigma_\mu)_m$. By calculations analogous to those in the proof of Proposition 2.1, it can be shown that as $m\to\infty$, $\lim_{N\to\infty}\mathrm{Cov}(\hat{\mu}_i,\hat{\mu}_{i,m})=(\Sigma_\mu)_{ii}$ and $(\Sigma_\mu)_{ii,m}\to(\Sigma_\mu)_{ii}$. Consequently

$$\lim_{m\to\infty}\varepsilon^{-2}[(\Sigma_\mu)_{ii}+(\Sigma_\mu)_{ii,m}-2\lim_{N\to\infty}N\mathrm{Cov}(\hat{\mu}_i,\hat{\mu}_{i,m})]=\varepsilon^{-2}[(\Sigma_\mu)_{ii}+(\Sigma_\mu)_{ii}-2(\Sigma_\mu)_{ii}]=0,$$

establishing (8) and the proposition. □

Let Π be the orthogonal $\nu\times\nu$ cyclic permutation matrix defined as

$$\Pi=\begin{pmatrix}0&1&0&0&\cdots&0\\0&0&1&0&\cdots&0\\\vdots&\vdots&\vdots&\vdots&&\vdots\\0&0&0&0&\cdots&1\\1&0&0&0&\cdots&0\end{pmatrix}. \quad (10)$$

Note that $\Pi^{-n}=(\Pi')^n$. Also, $\Pi^0=\Pi^\nu=I$. We use the Π-matrix in the following theorem and in Sections 3 and 4.

Theorem 2.3 *Let $\tilde{X}_t=\mu_t+\sum_{j=-\infty}^{\infty}\psi_t(j)\varepsilon_{t-j}$. Then,*

$$N^{1/2}(\hat{\mu}-\mu)\Rightarrow\mathcal{N}(0,\Sigma_\mu) \quad (11)$$

where $\Sigma_\mu=\sum_{n=-\infty}^{\infty}B_n\Pi^n$, $B_n=\mathrm{diag}\,(\gamma_0(n),\gamma_1(n),\ldots,\gamma_{\nu-1}(n))$, and Π^n is defined in (10).

PROOF. By Proposition 2.2 the theorem follows. Knowing that $(\Sigma_\mu)_{ij}=\sum_{n=-\infty}^{\infty}\gamma_i(n\nu+j-i)$ is the $(i,j)^{th}$ element of Σ_μ, the representation of Σ_μ given in the theorem can be shown through algebraic manipulation. □

The estimated periodic noise-variance (EPNV) function is directly related to the estimated periodic autocovariance function. We thus establish the asymptotic properties of estimators of the periodic autocovariance function, $\gamma_t(h)=\mathrm{Cov}(X_t,X_{t+h})$, at various lags h. Consider first the case when $\{\mu_t\}$ is known, and define

$$C_i(\ell)=N^{-1}\sum_{t=0}^{N-1}X_{t\nu+i}X_{t\nu+i+\ell}. \quad (12)$$

Proposition 2.4 *Let $X_t=\sum_{j=-\infty}^{\infty}\psi_t(j)\varepsilon_{t-j}$. Let $E\delta_t^4=\eta<\infty$, where we recall that $\{\delta_t=\sigma_t^{-1}\varepsilon_t\}$ is* iid*(0,1), and $\sum|\psi_i(j)|<\infty$ for all $i=0,1,\ldots,\nu-1$. Then, if $m_1\geq0$ and $m_2\geq0$, $\lim_{N\to\infty}N\mathrm{Cov}(C_i(m_1),C_\ell(m_2))=(W_{m_1m_2})_{i,\ell}$ where*

$$\begin{aligned}(W_{m_1m_2})_{i,\ell} &= (\eta-3)\textstyle\sum_{\ell_1=-\infty}^{\infty}\psi_i(\ell_1)\psi_{i+m_1}(\ell_1+m_1)\\ &\quad\cdot\textstyle\sum_{n=-\infty}^{\infty}\psi_\ell(\ell_1+n\nu+\ell-i)\psi_{\ell+m_2}(\ell_1+n\nu+\ell-i+m_2)\sigma_{i-\ell_1}^4\\ &+\textstyle\sum_{n=-\infty}^{\infty}[\gamma_i(n\nu+\ell-i)\gamma_{i+m_1}(n\nu+\ell-i-m_1+m_2)\\ &\quad+\gamma_i(n\nu+\ell-i+m_2)\gamma_{i+m_1}(n\nu+\ell-i-m_1)]\end{aligned} \quad (13)$$

for all $0\leq i,\ell\leq\nu-1$.

PROOF. We first observe that

$$E(\delta_s\delta_t\delta_u\delta_v)=\begin{cases}\eta & s=t=u=v\\ 1 & s=t\neq u=v\\ 0 & s\neq t, s\neq u, s\neq v.\end{cases} \tag{14}$$

Also,

$$\begin{aligned}&E(X_{t\nu+i}X_{t\nu+i+j}X_{t\nu+i+h+j}X_{t\nu+i+h+j+k})\\ &=\sum_{\ell_1}\sum_{\ell_2}\sum_{\ell_3}\sum_{\ell_4}\psi_i(\ell_1)\psi_{i+j}(\ell_2+j)\psi_{i+h+j}(\ell_3+h+j)\psi_{i+h+j+k}(\ell_4+h+j+k)\\ &\times\sigma_{i-\ell_1}\sigma_{i-\ell_2}\sigma_{i-\ell_3}\sigma_{i-\ell_4}E(\delta_{t\nu+i-\ell_1}\delta_{t\nu+i-\ell_2}\delta_{t\nu+i-\ell_3}\delta_{t\nu+i-\ell_4})\end{aligned}$$

and the sum can be rewritten using (14) as

$$\begin{aligned}&(\eta-3)\textstyle\sum_{\ell_1}\psi_i(\ell_1)\psi_{i+j}(\ell_1+j)\psi_{i+h+j}(\ell_1+h+j)\psi_{i+h+j+k}(\ell_1+h+j+k)\sigma^4_{i-\ell_1}\\ &\qquad+\gamma_i(j)\gamma_{i+h+j}(k)+\gamma_i(h+j)\gamma_{i+j}(h+k)+\gamma_i(h+j+k)\gamma_{i+j}(h).\end{aligned}$$

Hence,

$$\begin{aligned}&E(C_i(m_1)C_\ell(m_2))=N^{-2}E\Big[\textstyle\sum_{s=0}^{N-1}\sum_{t=0}^{N-1}X_{t\nu+i}X_{t\nu+i+m_1}X_{s\nu+\ell}X_{s\nu+\ell+m_2}\Big]\\ &=N^{-2}\textstyle\sum_{s=0}^{N-1}\sum_{t=0}^{N-1}\Big[(\eta-3)\sum_{\ell_1}\Big\{\psi_i(\ell_1)\psi_{i+m_1}(\ell_1+m_1)\psi_\ell(\ell_1+(s-t)\nu+\ell-i)\\ &\qquad\cdot\psi_{\ell+m_2}(\ell_1+(s-t)\nu+\ell-i+m_2)\sigma^4_{i-\ell_1}\Big\}\\ &+\gamma_i(m_1)\gamma_\ell(m_2)+\gamma_i((s-t)\nu+\ell-i)\gamma_{i+m_1}((s-t)\nu+\ell-i-m_1+m_2)\\ &\qquad+\gamma_i((s-t)\nu+\ell-i+m_2)\gamma_{i+m_1}((s-t)\nu+\ell-i-m_1)\Big].\end{aligned}$$

Letting $n=s-t$ we have

$$\begin{aligned}\mathrm{Cov}(C_i(m_1),C_\ell(m_2)) &= E[C_i(m_1)C_\ell(m_2)]-\gamma_i(m_1)\gamma_\ell(m_2)\\ &= N^{-1}\sum_{|n|<N}(1-N^{-1}|n|)T_n\end{aligned} \tag{15}$$

where

$$\begin{aligned}T_n&=(\eta-3)\textstyle\sum_{\ell_1}\psi_i(\ell_1)\psi_{i+m_1}(\ell_1+m_1)\psi_\ell(\ell_1+n\nu+\ell-i)\psi_{\ell+m_2}(\ell_1+n\nu+\ell-i+m_2)\sigma^4_{i-\ell_1}\\ &\quad+\gamma_i(n\nu+\ell-i))\gamma_{i+m_1}(n\nu+\ell-i-m_1+m_2)\\ &\qquad+\gamma_i(n\nu+\ell-i+m_2)\gamma_{i+m_1}(n\nu+\ell-i-m_1).\end{aligned}$$

The covariance term, $\gamma_i(n\nu+\ell-i)$, in the expression for T_n as expressed in terms of the ψ-weights is

$$\gamma_i(n\nu+\ell-i)=\sum_{r=-\infty}^{\infty}\psi_i(r)\psi_\ell(r+n\nu+\ell-i)\sigma^2_{i-r} \tag{16}$$

and hence

$$\sum_{n=-\infty}^{\infty} \gamma_i(n\nu+\ell-i) = \sum_{n=-\infty}^{\infty}\sum_{r=-\infty}^{\infty} \psi_i(r)\psi_\ell(r+n\nu+\ell-i)\sigma_{i-r}^2. \tag{17}$$

Note that,

$$\sum_{n=-\infty}^{\infty} |\gamma_i(n\nu+\ell-i)| \le \sum_{i=0}^{\nu-1}\sum_{i'=0}^{\nu-1}\sum_{r}\sum_{r'} |\psi_i(r)||\psi_i(r')| \times \max(\sigma_i^2; i=0,1,\ldots,\nu-1). \tag{18}$$

The other covariance terms in T_n can be treated equally, hence

$$\sum_{n=-\infty}^{\infty}\left[\left|\gamma_i(n\nu+\ell-i))\gamma_{i+m_1}(n\nu+\ell-i-m_1+m_2)\right| + \left|\gamma_i(n\nu+\ell-i+m_2)\gamma_{i+m_1}(n\nu+\ell-i-m_1)\right|\right]$$
$$\le 2\Big[\sum_{i=0}^{\nu-1}\sum_{i'=0}^{\nu-1}\sum_{r}\sum_{r'} |\psi_i(r)||\psi_{i'}(r')|\Big]^2 \times \max(\sigma_i^4; i=0,1,\ldots,\nu-1).$$

Similarly, $(\eta-3)\sum_{\ell_1}\psi_i(\ell_1)\psi_{i+m_1}(\ell_1+m_1)\sum_{n=-\infty}^{\infty}\psi_\ell(\ell_1+n\nu+\ell-i)\psi_{\ell+m_2}(\ell_1+n\nu+\ell-i+m_2)\sigma_{i-\ell_1}^4$

$$\le |(\eta-3)|\Big[\sum_{i=0}^{\nu-1}\sum_{i'=0}^{\nu-1}\sum_{r}\sum_{r'} |\psi_i(r)||\psi_{i'}(r')|\Big]^2 \times \max(\sigma_i^4; i=0,1,\ldots,\nu-1).$$

Thus, $\sum_{n=-\infty}^{\infty} |T_n| \le M$, where

$$M = (2+|\eta-3|)\Big[\sum_{i=0}^{\nu-1}\sum_{i'=0}^{\nu-1}\sum_{r}\sum_{r'} |\psi_i(r)||\psi_{i'}(r')|\Big]^2 \times \max(\sigma_i^4; i=0,1,\ldots,\nu-1). \tag{19}$$

We can apply the dominated convergence theorem in (15) to show that

$\lim_{N\to\infty} N\mathrm{Cov}(C_i(m_1), C_\ell(m_2)) = \sum_{n=-\infty}^{\infty} T_n$

$= (\eta-3)\sum_{\ell_1}\psi_i(\ell_1)\psi_{i+m_1}(\ell_1+m_1)\sum_{n=-\infty}^{\infty}\psi_\ell(\ell_1+n\nu+\ell-i)\psi_{\ell+m_2}(\ell_1+n\nu+\ell-i+m_2)\sigma_{i-\ell_1}^4$

$$+\sum_{n=-\infty}^{\infty}\Big[\gamma_i(n\nu+\ell-i)\gamma_{i+m_1}(n\nu+\ell-i-m_1+m_2)+\gamma_i(n\nu+\ell-i+m_2)\gamma_{i+m_1}(n\nu+\ell-i-m_1)\Big].$$

which completes the proof. □

Proposition 2.5 *Under the assumptions of Proposition 2.4, we have for any non-negative integer j,*

$$\begin{pmatrix} C(0) \\ \vdots \\ C(j) \end{pmatrix} \ \ is \ \ \mathrm{AN}\left[\begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(j) \end{pmatrix}, N^{-1}W\right] \tag{20}$$

where $C(m) = (C_0(m), \ldots, C_{\nu-1}(m))'$, $\gamma(m) = (\gamma_0(m), \ldots, \gamma_{\nu-1}(m))'$, and

$$W = \begin{pmatrix} W_{00} & W_{01} & \cdots & W_{0j} \\ W_{10} & W_{11} & \cdots & W_{1j} \\ \vdots & \vdots & \cdots & \vdots \\ W_{j0} & W_{j1} & \cdots & W_{jj} \end{pmatrix},$$

where $W_{m_1m_2}$, $m_1, m_2 = 0, 1, \ldots, j$, is the $\nu\times\nu$ matrix with (i,ℓ)th element given by (13).

PROOF. We define the truncated sequence $X_{t,s} = \sum_{j=-s}^{s} \psi_t(j)\varepsilon_{t-j}$, and let $\gamma_{t,s}(\cdot)$ be the autocovariance function of $\{X_{t,s}\}$. Next, define the $(j+1)\nu$-vectors

$$\begin{aligned} Y_{t,s} \;&=\; (X_{t\nu,s}X_{t\nu,s}, X_{t\nu+1,s}X_{t\nu+1,s}, \ldots, X_{t\nu+\nu-1,s}X_{t\nu+\nu-1,s}, \ldots, \\ &\quad X_{t\nu,s}X_{t\nu+j,s}, X_{t\nu+1,s}X_{t\nu+1+j,s}, \ldots, X_{t\nu+\nu-1,s}X_{t\nu+\nu-1+j,s})'. \end{aligned}$$

Then, $\{Y_{t,s}\}$ is a strictly stationary $\{\lfloor \frac{2s+j}{\nu} \rfloor + 1\}$-dependent sequence. Also,

$$N^{-1}\sum_{t=0}^{N-1} Y_t = \begin{pmatrix} C_s(0) \\ \vdots \\ C_s(j) \end{pmatrix}$$

where $C_s(m) = (C_{0,s}(m), C_{1,s}(m), \ldots, C_{\nu-1,s}(m))'$, and for $0 \le m_1, m_2 \le j$ we have

$$C_{i,s}(m_1) = N^{-1}\sum_{t=0}^{N-1} X_{t\nu+i,s}X_{t\nu+i+m_1,s} \quad \text{and} \quad C_{\ell,s}(m_2) = N^{-1}\sum_{t=0}^{N-1} X_{t\nu+\ell,s}X_{t\nu+\ell+m_2,s}$$

for $0 \le i, \ell \le \nu - 1$. By the Cramer-Wold Device it suffices to show that, as $N \to \infty$,

$$\lambda' \begin{pmatrix} C_s(0) \\ \vdots \\ C_s(j) \end{pmatrix} \text{ is AN} \left[\lambda' \begin{pmatrix} \gamma_s(0) \\ \vdots \\ \gamma_s(j) \end{pmatrix}, N^{-1}\lambda' W_s \lambda \right] \tag{21}$$

for all λ in $\mathbb{R}^{(j+1)\nu}$ such that $\lambda' W_s \lambda > 0$. Here, $\gamma_s(m) = (\gamma_{0,s}(m), \gamma_{1,s}(m), \ldots, \gamma_{\nu-1,s}(m))'$ and

$$W_s = \begin{pmatrix} W_{00,s} & W_{01,s} & \cdots & W_{0j,s} \\ W_{10,s} & W_{11,s} & \cdots & W_{1j,s} \\ \vdots & \vdots & \cdots & \vdots \\ W_{j0,s} & W_{j1,s} & \cdots & W_{jj,s} \end{pmatrix}.$$

Here, $(W_{m_1 m_2,s})$, $m_1, m_2 = 0, 1, \ldots, j$, is the $\nu \times \nu$ matrix with $(i, \ell)th$ element given by

$$\begin{aligned} &(W_{m_1 m_2,s})_{i,\ell} = (\eta - 3)\sum_{\ell_1=-s}^{s} \psi_i(\ell_1)\psi_{i+m_1}(\ell_1 + m_1) \\ &\cdot \sum_{n=-\infty}^{\infty} \psi_{\ell,s}(\ell_1 + n\nu + \ell - i)\psi_{\ell+m_2,s}(\ell_1 + n\nu + \ell - i + m_2)\sigma_{i-\ell_1}^4 \\ &\cdot \sum_{n=-\infty}^{\infty} \Big[\gamma_{i,s}(n\nu + \ell - i)\gamma_{i+m_1,s}(n\nu + \ell - i - m_1 + m_2) \\ &+ \gamma_{i,s}(n\nu + \ell - i + m_2)\gamma_{i+m_1,s}(n\nu + \ell - i - m_1) \Big] \end{aligned}$$

where:

$$\begin{cases} \psi_{i,s}(m) = \psi_i(m) & |m| \le s; \text{ and} \\ \psi_{i,s}(m) = 0 & |m| > s. \end{cases} \tag{22}$$

Noting that

$$E \begin{pmatrix} C_s(0) - \gamma_s(0) \\ C_s(1) - \gamma_s(1) \\ \vdots \\ C_s(j) - \gamma_s(j) \end{pmatrix} = 0$$

we have by Proposition 2.4,

$$\lim_{N\to\infty} N\mathrm{Var}\left(\lambda' \begin{pmatrix} C_s(0)-\gamma_s(0) \\ C_s(1)-\gamma_s(1) \\ \vdots \\ C_s(j)-\gamma_s(j) \end{pmatrix}\right) = \lambda' W_s \lambda$$

and by the Central Limit Theorem for strictly stationary s-dependent sequences,

$$\lambda' \begin{pmatrix} C_s(0)-\gamma_s(0) \\ C_s(1)-\gamma_s(1) \\ \vdots \\ C_s(j)-\gamma_s(j) \end{pmatrix} \textit{is} \ \ \mathrm{AN}\left[0, N^{-1}\lambda' W_s \lambda\right]. \tag{23}$$

Since $W_s \to W$ as $s \to \infty$, by Proposition 6.3.9 of Brockwell and Davis (1991) it suffices to show that for every $\varepsilon > 0$

$$\lim_{s\to\infty} \limsup_{N\to\infty} P[|N^{1/2}(C_i(m)-\gamma_i(m)) - N^{1/2}(C_{i,s}(m)-\gamma_{i,s}(m))| > \varepsilon] = 0. \tag{24}$$

By Chebyschev's inequality,

$$\begin{aligned} &P[|N^{1/2}(C_i(m)-\gamma_i(m)) - N^{1/2}(C_{i,s}(m)-\gamma_{i,s}(m))| > \varepsilon] \\ &\quad \le \varepsilon^{-2}\mathrm{Var}[N^{1/2}(C_i(m)-\gamma_i(m)) - N^{1/2}(C_{i,s}(m)-\gamma_{i,s}(m))]. \end{aligned}$$

The right-hand side of the above inequality is

$$\begin{aligned} &\varepsilon^{-2}\Big[\mathrm{Var}[N^{1/2}(C_i(m)-\gamma_i(m))] + \mathrm{Var}[N^{1/2}(C_{i,s}(m)-\gamma_{i,s}(m))] \\ &-2\mathrm{Cov}\Big(N^{1/2}(C_i(m)-\gamma_i(m)), N^{1/2}(C_{i,s}(m)-\gamma_{i,s}(m))\Big)\Big] \end{aligned} \tag{25}$$

Taking the limit as $N \to \infty$ in (25) we obtain

$$\varepsilon^{-2}[(W_{mm})_{ii} + (W_{mm})_{ii,s} - 2\lim_{N\to\infty} N\mathrm{Cov}\{C_i(m), C_{i,s}(m)\}],$$

where $(W_{mm})_{ii}$ is the (i,i)-element of W_{mm} and $(W_{mm,s})_{ii}$ is the (i,i)-element of $W_{mm,s}$. By calculations analogous to those in Proposition 2.4, it can be shown that

$$\begin{aligned} &\lim_{N\to\infty} N\mathrm{Cov}(C_i(m), C_{i,s}(m)) \\ &\quad = (\eta-3)\sum_{n=-\infty}^{\infty} \sum_{\ell_1=-s-n\nu}^{s-n\nu-m} \psi_i(\ell_1)\psi_{i+m}(\ell_1+m)\psi_i(\ell_1+n\nu)\psi_{i+m}(\ell_1+n\nu+m)\sigma^4_{i-\ell_1} \\ &\qquad + \sum_{n=-\infty}^{\infty} [\gamma_{i,s}(n\nu)\gamma_{i+m,s}(n\nu) + \gamma_{i,s}(n\nu+m)\gamma_{i+m,s}(n\nu-m)] \end{aligned}$$

Thus as $s \to \infty$, $\lim_{N\to\infty} N\mathrm{Cov}(C_i(m), C_{i,s}(m)) \to (W_{mm})_{ii}$ and $(W_{mm})_{ii,s} \to (W_{mm})_{ii}$ and consequently

$$\lim_{s\to\infty} \varepsilon^{-2}[(W_{mm})_{ii} + (W_{mm})_{ii,s} - 2\lim_{N\to\infty} N\mathrm{Cov}\{C_i(m), C_{i,s}(m)\}] = 0$$

establishing (25). $\square$

Theorem 2.6 *Let $\tilde{X}_t = \mu_t + \sum_{j=-\infty}^{\infty} \psi_t(j)\sigma_{t-j}\delta_{t-j}$ where δ_t is iid(0,1). Let $E\delta_t^4 = \eta < \infty$ and $\sum |\psi_i(j)| < \infty$ for all $i = 0, 1, \ldots, \nu - 1$. Then, for any non-negative integer j,*

$$\begin{pmatrix} \hat{\gamma}(0) \\ \vdots \\ \hat{\gamma}(j) \end{pmatrix} \quad is \quad \mathrm{AN}\left[\begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(j) \end{pmatrix}, N^{-1}W\right] \tag{26}$$

where $\hat{\gamma}(m) = (\hat{\gamma}_0(m), \hat{\gamma}_1(m), \ldots, \hat{\gamma}_{\nu-1}(m))'$ and given N years of data $\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_{N\nu-1}$, the estimated periodic autocovariance function at season i and lag m is defined by

$$\hat{\gamma}_i(m) = N^{-1} \sum_{j=0}^{N-1} (\tilde{X}_{j\nu+i} - \hat{\mu}_i)(\tilde{X}_{j\nu+i+m} - \hat{\mu}_{i+m}) \tag{27}$$

where $\hat{\mu}_i$ is defined by (3) and any terms involving $\tilde{X}_t$ are set equal to 0 whenever $t > N\nu - 1$. Also,

$$W = \begin{pmatrix} W_{00} & W_{01} & \cdots & W_{0j} \\ W_{10} & W_{11} & \cdots & W_{1j} \\ \vdots & \vdots & \cdots & \vdots \\ W_{j0} & W_{j1} & \cdots & W_{jj} \end{pmatrix}$$

where $W_{m_1 m_2}$, $m_1, m_2 = 0, 1, \ldots, j$, is the $\nu \times \nu$ matrix with (i, ℓ)th element given by (13).

PROOF. We want to show that for any season i, $i = 0, 1, \ldots, \nu - 1$ and any lag m, $0 \le m \le j$, that

$$N^{1/2}(\hat{\gamma}_i(m) - C_i(m)) = o_p(1).$$

Put $\hat{\mu}_i = N^{-1} \sum_{t=0}^{N-1} \tilde{X}_{t\nu+i}$ and observe that

$$\begin{aligned} &N^{1/2}(\hat{\gamma}_i(m) - C_i(m)) \\ &\quad = N^{1/2}\left[N^{-1} \sum_{t=0}^{N-1} (\tilde{X}_{t\nu+i} - \hat{\mu}_i)(\tilde{X}_{t\nu+i+m} - \hat{\mu}_{i+m}) - N^{-1} \sum_{t=0}^{N-1} (\tilde{X}_{t\nu+i} - \mu_i)(\tilde{X}_{t\nu+i+m} - \mu_{i+m})\right] \\ &\quad = N^{1/2}(\mu_i \hat{\mu}_{i+m} + \mu_{i+m}\hat{\mu}_i - \hat{\mu}_i \hat{\mu}_{i+m} - \mu_i \mu_{i+m}) \\ &\quad = N^{1/2}(\hat{\mu}_{i+m} - \mu_{i+m})(\mu_i - \hat{\mu}_i). \end{aligned}$$

Now $N^{1/2}(\hat{\mu}_{i+m} - \mu_{i+m}) \Rightarrow Y$ where Y is distributed $\mathcal{N}(0, V_{i+m,i+m})$ according to Theorem 2.3 where $V_{i+m,i+m} = \sum_{n=-\infty}^{\infty} \gamma_{i+m}(n\nu)$. Note that $(i + m)$ is modulo ν. This implies that $N^{1/2}(\hat{\mu}_{i+m} - \mu_{i+m}) = O_p(1)$, and by the weak law of large numbers, $\mu_i - \hat{\mu}_i = o_p(1)$, hence

$$N^{1/2}(\hat{\gamma}_i(m) - C_i(m)) = N^{1/2}(\hat{\mu}_{i+m} - \mu_{i+m})(\mu_i - \hat{\mu}_i) = o_p(1),$$

which completes the proof. □

Remark 2.7 *Tjøstheim and Paulsen (1982) derived asymptotics of the sample mean and covariance functions of a multivariate second-order stationary process. Since a periodic moving-average process can be expressed in terms of an equivalent multivariate stationary moving-average process, their results are relevant to PARMA modeling. However, the prediction problem is different for the two models. For example, a PARMA model of daily data uses observations from earlier days in the same year to make forecasts, whereas the vector model uses only observations from past years. Thus, our results are essential for complete model description, parsimony, and forecasting.*

3 Asymptotics of the estimated noise-variance

Using the innovations algorithm from Section 2 in Anderson, Meerschaert, and Vecchia (1999), the asymptotic distributions of the autoregressive parameters $\hat{\phi}_t(j)$ and moving average parameters $\hat{\theta}_t(i)$ for the PARMA$_\nu(p,q)$ process (1) were derived by Anderson and Meerschaert (2005), Theorems 1 and 2. The asymptotics of the seasonal sample mean $\hat{\mu}_t$ were laid out in the previous Section 2. Finally, we will develop the asymptotics of the noise-variance $\hat{\sigma}_t^2$ in this section. Assume:

(i) Finite Fourth Moment: $E\varepsilon_t^4 < \infty$.

(ii) The model admits a causal representation

$$X_t = \sum_{j=0}^{\infty} \psi_t(j)\varepsilon_{t-j} \tag{28}$$

where $\psi_t(0) = 1$ and $\sum_{j=0}^{\infty} |\psi_t(j)| < \infty$ for all t. Note that $\psi_t(j) = \psi_{t+k\nu}(j)$ for all j. Also, $X_t = \tilde{X}_t - \mu_t$ and $\varepsilon_t = \sigma_t\delta_t$ where $\{\delta_t\}$ is iid(0,1).

(iii) The model satisfies an invertibility condition

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_t(j)X_{t-j} \tag{29}$$

where $\pi_t(0) = 1$ and $\sum_{j=0}^{\infty} |\pi_t(j)| < \infty$ for all t. Again, $\pi_t(j) = \pi_{t+k\nu}(j)$ for all j.

(iv) The spectral density matrix $f(\lambda)$ of the equivalent vector ARMA process (see Anderson and Meerschaert, 1997, pg. 778) is such that for some $0 < c \le C < \infty$ we have

$$cz'z \le z'f(\lambda)z \le Cz'z, \quad -\pi \le \lambda \le \pi,$$

for all z in R^ν.

(v) The number of iterations k of the iterations algorithm satisfies $k \le N\nu - 1$, and $k^2/N \to 0$ as $N \to \infty$ and $k \to \infty$.

Proposition 3.1 *Define* $V_{N,k,j} = (V_{N,k,j}^{(0)}, V_{N,k,j}^{(1)}, \dots, V_{N,k,j}^{(\nu-1)})'$ *and* $V_j = (V_j^{(0)}, V_j^{(1)}, \dots, V_j^{(\nu-1)})'$ *where* $V_{N,k,j}^{(i)} = \sum_{\ell=0}^{j} \pi_i(\ell)N^{1/2}(\hat{\gamma}_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell))$, $(W_{m_1m_2})_{i-m_1,i-m_2}$ *is given by (13), and*

$$V_j^{(i)} \sim \mathcal{N}\left(0, \sum_{m_1=0}^{j}\sum_{m_2=0}^{j} \pi_i(m_1)\pi_i(m_2)(W_{m_1m_2})_{i-m_1,i-m_2}\right).$$

If assumptions (i) through (v) hold, then as $N \to \infty$ and $k \to \infty$, it follows that

$$V_{N,k,j} \Rightarrow V_j \sim \mathcal{N}(\underline{0}, A_jWA_j') \tag{30}$$

for fixed lag $j \ge 0$, where $\underline{0}$ is the ν-dimensional 0-vector, A_j is a $\nu \times (j+1)\nu$ block matrix defined by

$$A_j = \left[\ I \ \middle|\ F_1\Pi^{\nu-1} \ \middle|\ \cdots \ \middle|\ F_j\Pi^{\nu-j}\ \right] \tag{31}$$

where I is the $\nu \times \nu$ identity matrix, $F_j = \mathrm{diag}(\pi_0(j), \pi_1(j), \dots, \pi_{\nu-1}(j))$, Π is the the orthogonal $\nu \times \nu$ cyclic permutation matrix defined in Equation (10), and W is the matrix defined in Theorem 2.6.

PROOF. Since $V_{N,k,j} = A_j N^{1/2}\Gamma_j$ where

$$\Gamma_j = \begin{pmatrix} \hat{\gamma}(0) - \gamma(0) \\ \hat{\gamma}(1) - \gamma(1) \\ \vdots \\ \hat{\gamma}(j) - \gamma(j) \end{pmatrix} \quad \text{and} \quad \hat{\gamma}(\ell) - \gamma(\ell) = \begin{pmatrix} \hat{\gamma}_0(\ell) - \gamma_0(\ell) \\ \hat{\gamma}_1(\ell) - \gamma_1(\ell) \\ \vdots \\ \hat{\gamma}_{\nu-1}(\ell) - \gamma_{\nu-1}(\ell) \end{pmatrix}$$

the result follows from Theorem 2.6 and the continuous mapping theorem, Proposition 6.4.2 in Brockwell and Davis (1991). □

Remark 3.2 *Matrix multiplication shows that*

$$A_j W A_j' = \begin{pmatrix} s_{0,0,j} & \cdots & s_{0,\nu-1,j} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0,j} & \cdots & s_{\nu-1,\nu-1,j} \end{pmatrix}$$

where $s_{i,\ell,j} = \sum_{m_1=0}^{j}\sum_{m_2=0}^{j} \pi_i(m_1)\pi_\ell(m_2)(W_{m_1 m_2})_{i-m_1,\ell-m_2}$, $0 \le i,\ell \le \nu-1$, *and* $(i-m_1)$ *and* $(\ell-m_2)$ *are modulo* ν.

Proposition 3.3 *From Proposition 3.1 and Remark 3.2 we have, as* $j \to \infty$,

$$V_j \Rightarrow V \sim \mathcal{N}\left(\underline{0}, \begin{pmatrix} s_{0,0} & \cdots & s_{0,\nu-1} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0} & \cdots & s_{\nu-1,\nu-1} \end{pmatrix}\right) \tag{32}$$

where for all $0 \le i,\ell \le \nu - 1$ *we have*

$$s_{i,\ell} = \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} \pi_i(m_1)\pi_\ell(m_2)(W_{m_1 m_2})_{i-m_1,\ell-m_2}. \tag{33}$$

PROOF. In Remark 3.2, we see that

$$s_{i,\ell,j} = \sum_{m_1=0}^{j}\sum_{m_2=0}^{j} \pi_i(m_1)\pi_\ell(m_2)(W_{m_1 m_2})_{i-m_1,\ell-m_2}.$$

Noting that

$$|s_{i,\ell,j}| \le M \sum_{m_1=0}^{\infty} |\pi_i(m_1)| \sum_{m_2=0}^{\infty} |\pi_\ell(m_2)|$$

for all non-negative lags j and all $0 \le i,\ell \le \nu - 1$, where M is the bound given in Equation (19), we can write $s_{i,\ell} = \lim_{j\to\infty} s_{i,\ell,j}$. Let $\phi_{V_j^{(0)},\ldots,V_j^{(\nu-1)}}(t_1,\ldots,t_\nu)$ be the characteristic function of the vector $(V_j^{(0)},\ldots,V_j^{(\nu-1)})'$. Then

$$\lim_{j\to\infty} \phi_{V_j^{(0)},\ldots,V_j^{(\nu-1)}}(t_1,\ldots,t_\nu) = \lim_{j\to\infty} \exp\left(-\frac{1}{2}(t_1,\ldots,t_\nu)\begin{pmatrix} s_{0,0,j} & \cdots & s_{0,\nu-1,j} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0,j} & \cdots & s_{\nu-1,\nu-1,j} \end{pmatrix}\begin{pmatrix} t_1 \\ \vdots \\ t_\nu \end{pmatrix}\right)$$

$$= \exp\left(-\frac{1}{2}(t_1,\ldots,t_\nu)\lim_{j\to\infty}\begin{pmatrix} s_{0,0,j} & \cdots & s_{0,\nu-1,j} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0,j} & \cdots & s_{\nu-1,\nu-1,j} \end{pmatrix}\begin{pmatrix} t_1 \\ \vdots \\ t_\nu \end{pmatrix}\right)$$

$$= \exp\left(-\frac{1}{2}(t_1,\ldots,t_\nu)\begin{pmatrix} s_{0,0} & \cdots & s_{0,\nu-1} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0} & \cdots & s_{\nu-1,\nu-1} \end{pmatrix}\begin{pmatrix} t_1 \\ \vdots \\ t_\nu \end{pmatrix}\right) = \phi_{V^{(0)},\ldots,V^{(\nu-1)}}(t_1,\ldots,t_\nu),$$

which is the characteristic function of

$$V \sim \mathcal{N}\left(\underline{0}, \begin{pmatrix} s_{0,0} & \cdots & s_{0,\nu-1} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0} & \cdots & s_{\nu-1,\nu-1} \end{pmatrix}\right),$$

and this completes the proof. □

Lemma 3.4 *As $N \to \infty$ we have*

$$\lim_{j\to\infty} \limsup_{k\to\infty} P(|V^{(i)}_{N,k} - V^{(i)}_{N,k,j}| > \varepsilon) = 0 \tag{34}$$

for every $\varepsilon > 0$, $i = 0, 1, \ldots, \nu - 1$.

PROOF. By Chebyshev's inequality with $r = 2$ (e.g., see Proposition 6.2.1 in Brockwell and Davis, 1991), we have

$$\begin{aligned} P(|V^{(i)}_{N,k} - V^{(i)}_{N,k,j}| > \varepsilon) &\leq \varepsilon^{-2} E|V^{(i)}_{N,k} - V^{(i)}_{N,k,j}|^2 \\ &= \varepsilon^{-2} \left\| \sum_{\ell=j+1}^{k} \pi_i(\ell) N^{1/2} (\hat{\gamma}_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell)) \right\|^2 \\ &\leq \varepsilon^{-2} \left(\sum_{l=j+1}^{k} \sqrt{\pi_i^2(\ell) N \mathrm{Var}(\hat{\gamma}_{i-\ell}(\ell))} \right)^2 \\ &= \varepsilon^{-2} \left(\sum_{l=j+1}^{k} |\pi_i(\ell)| \right)^2 M \end{aligned} \tag{35}$$

where M is defined in Equation (19). Then, as $N \to \infty$, $M(\sum_{l=j+1}^{k} |\pi_i(\ell)|)^2 \to 0$ as $j, k \to \infty$ by the absolute summability of the π-weights. Hence, as $N \to \infty$, (34) holds for every $\varepsilon > 0$, $i = 0, 1, \ldots, \nu - 1$. □

Lemma 3.5 *As $N \to \infty$ we have*

$$\lim_{j\to\infty} \limsup_{k\to\infty} P(|\lambda_1 V^{(0)}_{N,k} + \cdots + \lambda_\nu V^{(\nu-1)}_{N,k} - (\lambda_1 V^{(0)}_{N,k,j} + \cdots + \lambda_\nu V^{(\nu-1)}_{N,k,j})| > \varepsilon) = 0$$

for every $\varepsilon > 0$ where $\lambda = (\lambda_1, \ldots, \lambda_\nu)'$ is any vector in $\mathbb{R}^\nu$.

PROOF. Write

$$\left|\lambda_1 V^{(0)}_{N,k} + \cdots + \lambda_\nu V^{(\nu-1)}_{N,k} - (\lambda_1 V^{(0)}_{N,k,j} + \cdots + \lambda_\nu V^{(\nu-1)}_{N,k,j})\right| = \left|\lambda_1 (V^{(0)}_{N,k} - V^{(0)}_{N,k,j}) + \cdots + \lambda_\nu (V^{(\nu-1)}_{N,k} - V^{(\nu-1)}_{N,k,j})\right|.$$

If $\left|\lambda_1 (V^{(0)}_{N,k} - V^{(0)}_{N,k,j}) + \cdots + \lambda_\nu (V^{(\nu-1)}_{N,k} - V^{(\nu-1)}_{N,k,j})\right| > \varepsilon$, then either

$\left|\lambda_1\right|\left|V^{(0)}_{N,k} - V^{(0)}_{N,k,j}\right| > \varepsilon/\nu$, or, …, or, $\left|\lambda_\nu\right|\left|V^{(\nu-1)}_{N,k} - V^{(\nu-1)}_{N,k,j}\right| > \varepsilon/\nu$.

Hence,

$P(|\lambda_1 V^{(0)}_{N,k} + \cdots + \lambda_\nu V^{(\nu-1)}_{N,k} - (\lambda_1 V^{(0)}_{N,k,j} + \cdots + \lambda_\nu V^{(\nu-1)}_{N,k,j})| > \varepsilon)$

$$\leq P\Big(\big|\lambda_1\big|\Big|V_{N,k}^{(0)} - V_{N,k,j}^{(0)}\Big| > \varepsilon/\nu\Big) + \cdots + P\Big(\big|\lambda_\nu\big|\Big|V_{N,k}^{(\nu-1)} - V_{N,k,j}^{(\nu-1)}\Big| > \varepsilon/\nu\Big) \to 0$$

as $N \to \infty$, $k \to \infty$ and $j \to \infty$ according to Lemma 3.4. $\Box$

Proposition 3.6 *As $N \to \infty$, we have*

$$\begin{pmatrix} V_{N,k}^{(0)} \\ \vdots \\ V_{N,k}^{(\nu-1)} \end{pmatrix} \Rightarrow \begin{pmatrix} V^{(0)} \\ \vdots \\ V^{(\nu-1)} \end{pmatrix}$$

as $k \to \infty$, where $V_{N,k}^{(i)} = \sum_{\ell=0}^{k} \pi_i(j) N^{1/2}(\hat{\gamma}_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell))$, $s_{i,\ell}$ is given by (33), and

$$\begin{pmatrix} V^{(0)} \\ \vdots \\ V^{(\nu-1)} \end{pmatrix} \sim \mathcal{N}\left(\underline{0}, \begin{pmatrix} s_{0,0} & \cdots & s_{0,\nu-1} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0} & \cdots & s_{\nu-1,\nu-1} \end{pmatrix}\right),$$

PROOF. By Propositions 3.1, 3.3, and Lemma 3.5, as well as the Cramer-Wold Device, we have by Proposition 6.3.9 of Brockwell and Davis (1991) that $\lambda' V_{N,k} \Rightarrow \lambda' V$ for all $\lambda = (\lambda_1, \ldots, \lambda_\nu)'$ in $\mathbb{R}^\nu$. By the Cramer-Wold Device we can conclude that $V_{N,k} \Rightarrow V$ as $N \to \infty$ and $k \to \infty$. $\Box$

Let $\hat{X}_{i+k}^{(i)} = P_{\mathcal{H}_{k,i}} X_{i+k}$ denote the one-step predictor of X_{i+k}, where the orthogonal projection

$$\hat{X}_{i+k}^{(i)} = \phi_{k,1}^{(i)} X_{i+k-1} + \cdots + \phi_{k,k}^{(i)} X_i, \quad k \geq 1, \tag{36}$$

onto the space $\mathcal{H}_{k,i} = \overline{\text{sp}}\{X_i, ..., X_{i+k-1}\}$ minimizes the mean squared error

$$v_{k,i} = \langle X_{i+k} - \hat{X}_{i+k}^{(i)}, X_{i+k} - \hat{X}_{i+k}^{(i)} \rangle = \|X_{i+k} - \hat{X}_{i+k}^{(i)}\|^2 = E(X_{i+k} - \hat{X}_{i+k}^{(i)})^2. \tag{37}$$

The error $v_{k,i}$ estimates the noise variance. The vector of coefficients $\phi_k^{(i)} = (\phi_{k,1}^{(i)}, \ldots, \phi_{k,k}^{(i)})'$ solves the prediction equations $\Gamma_{k,i}\phi_k^{(i)} = \gamma_k^{(i)}$, where $\gamma_k^{(i)} = (\gamma_{i+k-1}(1), \gamma_{i+k-2}(2), \ldots, \gamma_i(k))'$ and

$$\Gamma_{k,i} = \Big[\gamma_{i+k-\ell}(\ell - m)\Big]_{\ell,m=1,\ldots,k} \tag{38}$$

is the covariance matrix of $(X_{i+k-1}, ..., X_i)'$ for each $i = 0, ..., \nu - 1$. Proposition 4.1 of Lund and Basawa (1999) shows that if $\sigma_i^2 > 0$ for $i = 0, \ldots, \nu - 1$, then for a causal PARMA$_\nu(p, q)$ process the covariance matrix $\Gamma_{k,i}$ is nonsingular for every $k \geq 1$ and each i. Hence $\phi_k^{(i)} = \Gamma_{k,i}^{-1}\gamma_k^{(i)}$ for every $k \geq 1$ and each i. Given N years of data, $\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_{N\nu-1}$, if we replace the autocovariances in $\Gamma_{k,i}$ and $\gamma_k^{(i)}$ with sample autocovariances defined in (27), then we get $\hat{\phi}_k^{(i)} = \hat{\Gamma}_{k,i}^{-1}\hat{\gamma}_k^{(i)}$. Writing

$$\hat{X}_{i+k}^{(i)} = \sum_{j=1}^{k} \theta_{k,j}^{(i)} (X_{i+k-j} - \hat{X}_{i+k-j}^{(i)}) \tag{39}$$

yields the one-step predictors in terms of the *innovations* $X_{i+k-j} - \hat{X}_{i+k-j}^{(i)}$. Proposition 2.2.1 in Anderson, Meerschaert, and Vecchia (1999) shows that if $EX_t = 0$ and $\Gamma_{k,i}$ is nonsingular for each $k \geq 1$, then $\theta_{k,j}^{(i)}$ and $v_{k,i}$ can be computed using the *innovations algorithm.* If we have N years of data, $\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_{N\nu-1}$ and we replace the autocovariances in $\gamma_k^{(i)}$ with sample autocovariances defined in (27), we obtain the innovations estimates $\hat{\theta}_{k,j}^{(i)}$ and $\hat{v}_{k,i}$. If k is chosen as a function of the sample size N such that $k^2/N \to 0$ as $N \to \infty$ and $k \to \infty$, then the results in Section 3 of Anderson, Meerschaert, and Vecchia (1999) show that

$$\hat{\theta}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} \psi_i(j), \quad \hat{v}_{k,(\langle i-k \rangle)} \xrightarrow{P} \sigma_i^2 \quad \text{and} \quad \hat{\phi}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} -\pi_i(j) \tag{40}$$

for all i, j where "$\xrightarrow{P}$" denotes convergence in probability, $\langle t \rangle = t - \nu[t/\nu]$ for $t = 0, 1, \ldots$ and $\langle t \rangle = \nu + t - \nu[t/\nu + 1]$ for $t = -1, -2, \ldots$ so that $\langle t \rangle$ denotes the season associated with time t.

Lemma 3.7 *As $N \to \infty$ and $k \to \infty$, we have that*

$$\sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)}) N^{1/2} (\hat{\gamma}_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell)) = o_p(1). \tag{41}$$

PROOF. Since $\hat{\gamma}_{i-\ell}(\ell)$ and $C_{i-\ell}(\ell)$, defined in Equation (12), have the same asymptotic distribution it suffices to show that $\sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)}) N^{1/2} (C_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell)) = o_p(1)$. From Chebychev's inequality,

$$\begin{aligned}
&P\Big(\Big| \sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)}) N^{1/2} (C_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell)) \Big| > \varepsilon \Big) \\
&\quad \leq \varepsilon^{-2} E \Big(\sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)}) N^{1/2} (C_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell)) \Big)^2 \\
&\quad = \varepsilon^{-2} \Big[\sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)})^2 N \mathrm{Var}(C_{i-\ell}(\ell)) \\
&\qquad\qquad + 2 \sum_{m<n} (\pi_i(m) + \phi_{k,m}^{(\langle i-k \rangle)}) (\pi_i(n) + \phi_{k,n}^{(\langle i-k \rangle)}) \times N \mathrm{Cov}(C_{i-m}(m), C_{i-n}(n)) \Big].
\end{aligned}$$

Note that $N\mathrm{Var}(C_{i-\ell}(\ell) \leq M$ where M is the upper bound given in equation (19). Similarly, by the Cauchy-Schwarz inequality, $N\mathrm{Cov}(C_{i-m}(m), C_{i-n}(n)) \leq M$. Hence,

$$\begin{aligned}
&\varepsilon^{-2} \Big[\sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)})^2 N \mathrm{Var}(C_{i-\ell}(\ell)) \\
&\qquad\qquad + 2 \sum_{m<n} (\pi_i(m) + \phi_{k,m}^{(\langle i-k \rangle)}) (\pi_i(n) + \phi_{k,n}^{(\langle i-k \rangle)}) \times N \mathrm{Cov}(C_{i-m}(m), C_{i-n}(n)) \Big] \\
&\quad \leq \varepsilon^{-2} M \Big[\sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)}) \Big]^2 = \varepsilon^{-2} M \cdot 0 = 0
\end{aligned}$$

since from the proof of Corollary 2.2.4 in Anderson, Meerschaert, and Vecchia (1999), we have $\sum_{\ell=0}^{k} (\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k \rangle)}) \to 0$, as $N \to \infty$ and $k \to \infty$. □

Lemma 3.8 *In addition to assumptions (i–v) in this section, suppose that*

$$N^{3/4} \sum_{\ell > k} |\pi_i(\ell)| \to 0 \text{ as } N \to \infty \text{ and } k \to \infty. \tag{42}$$

Then

$$\sum_{\ell=0}^{k} N^{1/2}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})\gamma_{i-\ell}(\ell) \to 0 \tag{43}$$

as $N \to \infty$ and $k \to \infty$.

PROOF. It is easily shown that $\gamma_{i-\ell}(\ell) = \sum_{k=0}^{\infty} \psi_{i-\ell}(k)\psi_i(\ell+k)\sigma_{i-\ell-k}^2$. Hence

$$|\gamma_{i-\ell}(\ell)| \le \max(\sigma_i^2; i = 0, 1, \ldots, \nu-1) \times \sum_{i'=0}^{\nu-1}\sum_{i=0}^{\nu-1}\Big|\sum_{k=0}^{\infty}\psi_{i'}(k)\Big|\Big|\sum_{k=0}^{\infty}\psi_i(k)\Big|$$

where the right-hand side of the above inequality is a constant, call it B. Hence, $|\gamma_{i-\ell}(\ell)|$ is bounded for $0 \le \nu - 1$ and all ℓ. Next, apply Corollary 2.2.4 in Anderson, Meerschaert, and Vecchia (1999) to see that

$$\begin{aligned} \sum_{\ell=0}^{k} N^{1/2}|\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)}| &\le N^{1/2}\sum_{\ell=0}^{k}|\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)}| \\ &\le N^{1/2}k^{1/2}\sqrt{\sum_{\ell=0}^{k}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})^2}. \end{aligned} \tag{44}$$

From assumption (v), we have $(k^2/N)^{1/4} \to 0$, i.e., $k^{1/2}/N^{1/4} \to 0$, so for $N > N_0$ for some integer N_0, we have

$$\begin{aligned} N^{1/2}k^{1/2}\sqrt{\sum_{\ell=0}^{k}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})^2} &\le N^{1/2}N^{1/4}\sqrt{\sum_{\ell=0}^{k}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})^2} \\ &\le N^{3/4}\sqrt{\sum_{\ell=0}^{k}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})^2}. \end{aligned} \tag{45}$$

The proof of Corollary 2.2.4 in Anderson, Meerschaert, and Vecchia (1999) contains the inequality $\sum_{\ell=0}^{k}(\pi_i(\ell)+\phi_{k,\ell}^{(\langle i-k\rangle)})^2 \le \frac{2K}{\pi C}\Big(\sum_{\ell>k}|\pi_i(\ell)\Big)^2$ where $K = \max(\gamma_i(0); i = 0, 1, \ldots, \nu-1)$ and the constant C is as in assumption (iv) of this section. Hence,

$$\sum_{\ell=0}^{k}\Big[N^{3/4}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})\Big]^2 \le \frac{2K}{\pi C}\Big(\sum_{\ell>k}N^{3/4}|\pi_i(\ell)|\Big)^2.$$

If $N^{3/4}\sum_{\ell>k}|\pi_i(\ell)| \to 0$ as $N \to \infty$ and $k \to \infty$, then $\frac{2K}{\pi C}\Big(\sum_{\ell>k}N^{3/4}|\pi_i(\ell)|\Big)^2 \to 0$, therefore $\sum_{\ell=0}^{k}N^{1/2}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})\gamma_{i-\ell}(\ell) \to B \cdot 0 = 0$ and the Lemma is proved. □

Lemma 3.9 *If (42) holds along with assumptions (i–v), we have that*

$$\sum_{\ell=k+1}^{\infty} N^{1/2}\pi_i(\ell)\gamma_{i-\ell}(\ell) \to 0 \tag{46}$$

as $N \to \infty$ and $k \to \infty$.

PROOF. Note that

$$\Big|\sum_{\ell=k+1}^{\infty} N^{1/2}\pi_i(\ell)\gamma_{i-\ell}(\ell)\Big| \leq \sum_{\ell=k+1}^{\infty} N^{1/2}|\pi_i(\ell)||\gamma_{i-\ell}(\ell)| \leq B\cdot\sum_{\ell=k+1}^{\infty} N^{1/2}|\pi_i(\ell)| \to 0, \tag{47}$$

where B is as in Lemma 3.8. $\square$

Proposition 3.10 *Given assumptions (i–v) as well as the condition (42), we have*

$$V_{N,k}^{(i)} - N^{1/2}(\hat{v}_{k,\langle i-k\rangle} - \sigma_i^2) = o_p(1).$$

PROOF. Let $\mathcal{H}_t = \overline{sp}\{X_j, -\infty < j \leq t\}$. Then

$$E(\varepsilon_t^2) = \sigma_t^2 = E[X_t + \sum_{j=1}^{\infty}\pi_t(j)X_{t-j}]^2 = \langle X_t - P_{\mathcal{H}_{t-1}}X_t, X_t - P_{\mathcal{H}_{t-1}}X_t\rangle$$

where $P_{\mathcal{H}_{t-1}}X_t$ is the projection of X_t onto the closed subspace $\mathcal{H}_{t-1}$. Since $P_{\mathcal{H}_{t-1}}X_t \perp (X_t - P_{\mathcal{H}_{t-1}}X_t)$, and

$$\langle X_t, X_t - P_{\mathcal{H}_{t-1}}X_t\rangle = \gamma_t(0) + \pi_t(1)\gamma_t(-1) + \cdots$$

we have $\sigma_t^2 = \sum_{j=0}^{\infty}\pi_t(j)\gamma_t(-j) = \sum_{j=0}^{\infty}\pi_t(j)\gamma_{t-j}(j)$. Hence,

$$\hat{v}_{k,\langle i-k\rangle} - \sigma_i^2 = \sum_{\ell=0}^{k} -\phi_{k,\ell}^{(\langle i-k\rangle)}\hat{\gamma}_{i-\ell}(\ell) - \sum_{\ell=0}^{\infty}\pi_i(\ell)\gamma_{i-\ell}(\ell)$$

and

$$N^{1/2}(\hat{v}_{k,\langle i-k\rangle} - \sigma_i^2) = \sum_{\ell=0}^{k} -\phi_{k,\ell}^{(\langle i-k\rangle)}N^{1/2}\hat{\gamma}_{i-\ell}(\ell) - \sum_{\ell=0}^{\infty}\pi_i(\ell)N^{1/2}\gamma_{i-\ell}(\ell).$$

Then,

$$\begin{aligned}
&V_{N,k}^{(i)} - N^{1/2}(\hat{v}_{k,\langle i-k\rangle} - \sigma_i^2)\\
&= \sum_{\ell=0}^{k}\pi_i(\ell)N^{1/2}(\hat{\gamma}_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell)) + \sum_{\ell=0}^{k}\phi_{k,\ell}^{(\langle i-k\rangle)}N^{1/2}\hat{\gamma}_{i-\ell}(\ell) + \sum_{\ell=0}^{\infty}\pi_i(\ell)N^{1/2}\gamma_{i-\ell}(\ell)\\
&= \sum_{\ell=0}^{k}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})N^{1/2}(\hat{\gamma}_{i-\ell}(\ell) - \gamma_{i-\ell}(\ell)) + \sum_{\ell=0}^{k}N^{1/2}(\pi_i(\ell) + \phi_{k,\ell}^{(\langle i-k\rangle)})\gamma_{i-\ell}(\ell)\\
&\quad + \sum_{\ell=k+1}^{\infty}N^{1/2}\pi_i(\ell)\gamma_{i-\ell}(\ell).
\end{aligned}$$

As $N\to\infty$ and $k\to\infty$, the first summand above approaches 0 in probability by Lemma 3.7, the second summand approaches 0 by Lemma 3.8, and the last approaches 0 by Lemma 3.9. $\square$

Proposition 3.11 *Given assumptions (i) through (v) as well as the condition (42), we have*

$$N^{1/2}\begin{pmatrix}\hat{v}_{k,\langle 0-k\rangle} - \sigma_0^2\\ \hat{v}_{k,\langle 1-k\rangle} - \sigma_1^2\\ \vdots\\ \hat{v}_{k,\langle \nu-1-k\rangle} - \sigma_{\nu-1}^2\end{pmatrix} - V_{N,k} = o_p(1). \tag{48}$$

where $V_{N,k} = (V_{N,k}^{(0)}, V_{N,k}^{(1)}, \ldots, V_{N,k}^{(\nu-1)})'$ *and*

$$V_{N,k}^{(i)} = \sum_{j=0}^{k} \pi_i(j) N^{1/2} (\hat{\gamma}_{i-j}(j) - \gamma_{i-j}(j)). \tag{49}$$

PROOF. The Proposition follows from Proposition 3.6 and from Definition 6.1.4 in Brockwell and Davis (1991). □

Theorem 3.12 *Suppose that the PARMA time series defined in (1) is causal, invertible, $E\varepsilon_t^4 < \infty$, and that for some $0 < c \le C < \infty$ we have $cz'z \le z'f(\lambda)z \le Cz'z$, $-\pi \le \lambda \le \pi$, for all z in $\mathbb{R}^\nu$, where $f(\lambda)$ is the spectral density matrix of the equivalent vector moving average process (see Anderson and Meerschaert, 1997, pg. 778). Then for any sequence of positive integers $\{k(N), N = 1, 2, \ldots\}$ such that $k \le N\nu - 1$, $k \to \infty$ and $k^2/N \to 0$ as $N \to \infty$ and (42) holds, we have that*

$$\begin{pmatrix} \hat{v}_{k,\langle 0-k\rangle} \\ \hat{v}_{k,\langle 1-k\rangle} \\ \vdots \\ \hat{v}_{k,\langle \nu-1-k\rangle} \end{pmatrix} \quad is \ \ \text{AN} \left[\begin{pmatrix} \sigma_0^2 \\ \sigma_1^2 \\ \vdots \\ \sigma_{\nu-1}^2 \end{pmatrix}, N^{-1} \begin{pmatrix} s_{0,0} & \cdots & s_{0,\nu-1} \\ \vdots & \cdots & \vdots \\ s_{\nu-1,0} & \cdots & s_{\nu-1,\nu-1} \end{pmatrix} \right] \tag{50}$$

where $s_{i,\ell}$ is given by (33) and $(W_{m_1 m_2})_{i,\ell}$ is given by (13), with $\eta = E(\delta_t^4)$ and $\delta_t = \sigma_t^{-1}\varepsilon_t$.

PROOF. Theorem 3.12 follows from Propositions 3.6 and 3.11 and from Proposition 6.3.3 in Brockwell and Davis (1991). □

Corollary 3.13 *The convergence in Theorem 3.12 can be rewritten as*

$$N^{1/2} \begin{pmatrix} \hat{v}_{k,\langle 0-k\rangle} - \sigma_0^2 \\ \hat{v}_{k,\langle 1-k\rangle} - \sigma_1^2 \\ \vdots \\ \hat{v}_{k,\langle \nu-1-k\rangle} - \sigma_{\nu-1}^2 \end{pmatrix} \Rightarrow \mathcal{N}(0, \lim_{j\to\infty} A_j W A_j') \tag{51}$$

where we recall from equation (31) that A_j is the $\nu \times (j+1)\nu$ block matrix given by

$$A_j = \left[\ I \ \middle|\ F_1 \Pi^{\nu-1} \ \middle|\ \cdots \ \middle|\ F_j \Pi^{\nu-j}\ \right] \tag{52}$$

and

$$A_j' = \begin{bmatrix} I \\ \hline \Pi F_1 \\ \hline \vdots \\ \hline \Pi^j F_j \end{bmatrix}. \tag{53}$$

where I is the $\nu \times \nu$ identity matrix, $F_j = \operatorname{diag}(\pi_0(j), \pi_1(j), \ldots, \pi_{\nu-1}(j))$, and Π is the the orthogonal $\nu \times \nu$ cyclic permutation matrix in Equation (10), and the matrix W is defined in Theorem 2.6.

PROOF. The corollary follows by Propositions 3.1 and 3.3 and Remark 3.2.

4 Fourier-PARMA modeling

The general PARMA$_\nu(p,q)$ model has $p+q+2$ parameter vectors, meaning that there are $(p+q+2)\nu$ parameter values, a large number for high frequency data. Each parameter in an ARMA model becomes a periodic function in the corresponding PARMA model, written as a vector of parameters. In this section, we detail a hypothesis test to identify the statistically significant DFT harmonics. The remaining harmonics can be then zeroed out, leading to a parsimonious Fourier-PARMA model.

4.1 The discrete Fourier transform

For any given parameter vector $\boldsymbol{x}$, the DFT of the periodic function $\boldsymbol{x} = [x_t : 0 \le t \le \nu - 1]$ is

$$x_t = c_{x0} + \sum_{r=1}^{k} \left\{ c_{xr} \cos\left(\frac{2\pi r t}{\nu}\right) + s_{xr} \sin\left(\frac{2\pi r t}{\nu}\right) \right\} \tag{54}$$

where c_{xr} and s_{xr} are the DFT harmonics, r is the Fourier frequency, and k is the total number of frequencies.Write the vector of DFT harmonics in the form

$$f_{\boldsymbol{x}} = \begin{cases} \left[c_{x0}, c_{x1}, s_{x1}, \ldots, c_{x(\nu-1)/2}, s_{x(\nu-1)/2}\right]' & (\nu \text{ odd}) \\ \left[c_{x0}, c_{x1}, s_{x1}, \ldots, s_{x(\nu/2-1)}, c_{x(\nu/2)}\right]' & (\nu \text{ even}). \end{cases} \tag{55}$$

The complex DFT and its inverse

$$x_r^* = \nu^{-1/2} \sum_{t=0}^{\nu-1} \exp\left(\frac{-2i\pi r t}{\nu}\right) x_t \quad \text{and} \quad x_t = \nu^{-1/2} \sum_{r=0}^{\nu-1} \exp\left(\frac{2i\pi r t}{\nu}\right) x_r^* \tag{56}$$

can be expressed in matrix form $\boldsymbol{x}^* = U\boldsymbol{x}$ using the $\nu \times \nu$ matrix

$$U = \nu^{-1/2} \left(e^{\frac{-i2\pi r t}{\nu}}\right)_{r,t=0,1,\ldots,\nu-1}. \tag{57}$$

Since U is a unitary matrix, the complex DFT inverse $\hat{\boldsymbol{x}} = \tilde{U}^{-1}\hat{\boldsymbol{x}}^* = \tilde{U}'\hat{\boldsymbol{x}}^*$ where the tilde denotes taking the complex conjugate of each entry. Define

$$[P]_{\ell j} = \begin{cases} 1 & \text{if } \ell = j = 0; \\ 2^{-1/2} & \text{if } \ell = 2r-1 \text{ and } j = r \text{ for some } 1 \le r \le [(\nu-1)/2]; \\ 2^{-1/2} & \text{if } \ell = 2r-1 \text{ and } j = \nu - r \text{ for some } 1 \le k \le \lfloor(\nu-1)/2\rfloor; \\ i2^{-1/2} & \text{if } \ell = 2r \text{ and } j = r \text{ for some } 1 \le r \le [(\nu-1)/2]; \\ -i2^{-1/2} & \text{if } \ell = 2r \text{ and } j = \nu - r \text{ for some } 1 \le k \le \lfloor(\nu-1)/2\rfloor; \\ 1 & \text{if } \ell = \nu - 1 \text{ and } j = \nu/2 \text{ and } \nu \text{ is even; and} \\ 0 & \text{otherwise.} \end{cases} \tag{58}$$

The matrix P is also unitary, and if we let

$$L = \begin{cases} \operatorname{diag}(\nu^{-1/2}, \sqrt{2/\nu}, \ldots, \sqrt{2/\nu}) & (\nu \text{ odd}) \\ \operatorname{diag}(\nu^{-1/2}, \sqrt{2/\nu}, \ldots, \sqrt{2/\nu}, \nu^{-1/2}) & (\nu \text{ even}) \end{cases} \tag{59}$$

it follows that (see Section 5 of Tesfaye, Anderson and Meerschaert (2011) for complete details)

$$f_{\boldsymbol{x}} = LPU\boldsymbol{x} \quad \text{and} \quad \hat{f}_{\boldsymbol{x}} = LPU\hat{\boldsymbol{x}}. \tag{60}$$

Now suppose that the parameter vector $\hat{\boldsymbol{x}}$ has normal asymptotics

$$N^{1/2}(\hat{\boldsymbol{x}} - \boldsymbol{x}) \Rightarrow \mathcal{N}(0, \Sigma_{\boldsymbol{x}}), \tag{61}$$

so that $\hat{\boldsymbol{x}}$ is an asymptotically unbiased estimator of $\boldsymbol{x}$ with asymptotic variance-covariance matrix $\Sigma_{\boldsymbol{x}}$. Let $G = LPU$ so that $f_{\boldsymbol{x}} = G\boldsymbol{x}$ and $\hat{f}_{\boldsymbol{x}} = G\hat{\boldsymbol{x}}$. Then it follows by a straightforward application of the Continuous Mapping Theorem that

$$N^{1/2}\left[\hat{f}_{\boldsymbol{x}} - f_{\boldsymbol{x}}\right] \Rightarrow \mathcal{N}(0, R_{\boldsymbol{x}}) \tag{62}$$

where

$$R_{\boldsymbol{x}} = G\Sigma_{\boldsymbol{x}}\tilde{G}' = LPU\Sigma_{\boldsymbol{x}}\tilde{U}'\tilde{P}'L. \tag{63}$$

The goal of this section is to determine which real DFT harmonics are statistically significantly different from zero. Then by zeroing out the insignificant terms, a more parsimonious PARMA model can be obtained. The null hypothesis of the test is therefore $H_0 : f_{\boldsymbol{x}} = (c_{x0}, 0, \ldots, 0)'$, in which case it follows from (54) that $x_t = c_{x0}$ is a constant function. If all the parameter vectors are constant functions, then the PARMA process reduces to a stationary ARMA process, with mean $\mu = \mu_t$ for all $0 \le t \le \nu - 1$ and likewise for the remaining parameters. In this sense, the null hypothesis is that the model is stationary.

Now (62) leads directly to the desired test for significance. Letting $[M]_{ij}$ denote the ij entry of a matrix M, it follows from (62) that

$$N^{1/2}\left[[\hat{f}_{\boldsymbol{x}}]_i - [f_{\boldsymbol{x}}]_i\right] \Rightarrow \mathcal{N}(0, [R_{\boldsymbol{x}}]_{ii}). \tag{64}$$

Recalling that we index the vector $f_{\boldsymbol{x}}$ as $0, 1, 2, \ldots, \nu - 1$, for each cosine term $i = 1, 2, 3, \ldots, k$, we wish to test the null hypothesis $H_0 : [f_{\boldsymbol{x}}]_{2i-1} = c_{xi} = 0$ versus $H_a : [f_{\boldsymbol{x}}]_{2i-1} = c_{xi} \ne 0$. For each of these two-sided z-tests, the test statistic is

$$z_{2i-1} = \frac{[\hat{f}_{\boldsymbol{x}}]_{2i-1}}{\sqrt{[R_{\boldsymbol{x}}]_{2i-1,2i-1}/N}} \tag{65}$$

and then we reject the null hypothesis at level $\alpha = 0.5$ when the test statistic $|z_{2i-1}| > 3.81$. This Bonferroni test uses the $\alpha' = \alpha/(\nu - 1)$ tail quantile with $\nu = 365$, so that $P[Z > 3.81] = \alpha'/2$ when $Z \sim \mathcal{N}(0, 1)$. For each sine term $i = 1, 2, 3, \ldots, k$ we wish to test the null hypothesis $H_0 : [f_{\boldsymbol{x}}]_{2i} = s_{xi} = 0$ versus $H_a : [f_{\boldsymbol{x}}]_{2i} = s_{xi} \ne 0$. The test statistic is

$$z_{2i} = \frac{[\hat{f}_{\boldsymbol{x}}]_{2i}}{\sqrt{[R_{\boldsymbol{x}}]_{2i,2i}/N}} \tag{66}$$

and again we reject the null hypothesis at level $\alpha = 0.5$ when the test statistic $|z_{2i}| > 3.81$.

The Fourier-PARMA model for the real harmonics $f_{\boldsymbol{x}}$ of the DFT of this parameter vector is obtained by setting $[\hat{f}_{\boldsymbol{x}}]_i = 0$ whenever the null hypothesis is not rejected. This leaves us with a sparse parameter vector $\hat{f}_{x'}$ where (hopefully) most of the vector entries are zero. Inverting the DFT, we obtain the Fourier-PARMA model $\boldsymbol{x}' = G^{-1}\hat{f}_{\boldsymbol{x}'}$, which is a smoothed version of the original parameter estimates. Substituting all these $p+q+2$ smoothed parameter vectors into the PARMA model (1) yields the Fourier-PARMA model for this time series, a parsimonious alternative with a greatly reduced number of parameters. For the autoregressive and moving average parameters, this procedure for fitting a Fourier-PARMA model was laid out in Section 4 of Tesfaye, Anderson, and Meerschaert (2011), and illustrated with an application to weekly river flows in Section 5 of Anderson, Tesfaye and Meerschaert (2007). The Fourier-PARMA model for the seasonal mean and standard deviation is the subject of the present paper, and together with our previous results, it completes the construction of Fourier-PARMA modeling for high frequency data.

4.2 Periodic autoregression of order one

Here we describe the calculations needed to test for statistically significant real DFT harmonics in a PARMA$_\nu(1, 0)$ model with Gaussian innovations. This is the model we will apply to the climate data in the next section. We write this model in the form $\tilde{X}_t = X_t + \mu_t$ where

$$X_t = \phi_t X_{t-1} + \varepsilon_t$$

and $\sigma_t^{-1}\varepsilon_t$ are IID Gaussian. The three parameter vectors are $\boldsymbol{\mu}$, $\boldsymbol{\sigma}^2$, and $\boldsymbol{\phi}$. Section 3.2 in Tesfaye, Anderson, and Meerschaert (2011) shows that $\boldsymbol{\psi}(1) = \boldsymbol{\phi}$ in the PMA(∞) representation (28) of this

PAR(1) model. Corollary 1 in Anderson and Meerschaert (2005) shows that the variance-covariance matrix for the $\hat{\boldsymbol{\psi}}(1)$ vector from the innovations algorithm is I, the $\nu \times \nu$ identity matrix. Then $\Sigma_{\boldsymbol{\phi}} = I$ and $R_{\boldsymbol{\phi}} = LPUI\tilde{U}'\tilde{P}'L' = L^2$ since both P and U are unitary, and L is a diagonal matrix. Then:

$$R_{\boldsymbol{\phi}} = L^2 = \begin{cases} \operatorname{diag}(1/\nu, 2/\nu, \ldots, 2/\nu) & \text{for } \nu \text{ odd; and} \\ \\ \operatorname{diag}(1/\nu, 2/\nu, \ldots, 2/\nu, 1/\nu) & \text{for } \nu \text{ even.} \end{cases} \tag{67}$$

Under the null hypothesis we have $\phi_t = \phi$ and $\gamma_t(h) = \gamma(h)$ for all integers t. Then for a PAR(1) model, it is not hard to check using $\gamma_t(1) = E[X_t X_{t+1}] = E[X_t(\phi_{t+1}X_t + \varepsilon_{t+1})] = \phi_{t+1}\gamma_t(0)$ and so forth that the autocovariance function is

$$\gamma(h) = \phi^{|h|}\gamma(0) \quad \text{for any integer } h. \tag{68}$$

Then using Theorem 2.3 along with (5), if $i = j$, we have

$$(\Sigma_{\boldsymbol{\mu}})_{ii} \;=\; \sum_{n=-\infty}^{\infty} \gamma_i(n\nu) = \sum_{n=-\infty}^{\infty} \phi^{|n\nu|}\gamma(0) = \gamma(0)\left[\frac{1+r}{1-r}\right] \tag{69}$$

where $r = \phi^\nu$, using the formula for a geometric series (twice). If $j > i$ we have

$$(\Sigma_{\boldsymbol{\mu}})_{ij} = \phi^{j-i}\gamma(0) + \sum_{n=1}^{\infty} r^n \phi^{j-i}\gamma(0) + \sum_{n=1}^{\infty} r^{n-1}\phi^{\nu+i-j}\gamma(0) = \gamma(0)\left[\frac{\phi^{j-i} + \phi^{\nu+i-j}}{1-r}\right]. \tag{70}$$

Since $\Sigma_{\boldsymbol{\mu}}$ is a symmetric matrix, this completes the calculation.

To calculate $\Sigma_{\boldsymbol{\sigma}^2}$, we apply Theorem 3.12. Since ε_t are iid *Gaussian*, the first term in the formula (13) for $(W_{m_1 m_2})_{i,\ell}$ vanishes. The invertible representation (29) of the PAR(1) process is

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_t(j) X_{t-j} = X_t - \phi X_{t-1}$$

and hence we have $\pi_t(0) = 1$, $\pi_t(1) = \pi(1) = -\phi$, and $\pi_t(j) = 0$ for all $j > 1$. Then

$$s_{i\ell} = (W_{00})_{i\ell} - \phi(W_{01})_{i,\ell-1} - \phi(W_{10})_{i-1,\ell} + \phi^2(W_{11})_{i-1,\ell-1}. \tag{71}$$

From Proposition 2.5 we have that, for $\ell \geq i$, the first term (with $m_1 = m_2 = 0$) is

$$\begin{aligned} (W_{00})_{i\ell} \;&=\; 2\sum_{n=-\infty}^{\infty} \gamma_i(n\nu + \ell - i)^2 = 2\sum_{n=-\infty}^{\infty} \phi^{2|n\nu+\ell-i|}\gamma(0)^2 \\ &=\; 2\phi^{2(\ell-i)}\gamma(0)^2 \sum_{n=0}^{\infty} \phi^{2n\nu} + 2\phi^{2(\nu+i-\ell)}\gamma(0)^2 \sum_{n=0}^{\infty} \phi^{2n\nu} \\ &=\; 2\gamma(0)^2 \left[\frac{\phi^{2(\ell-i)} + \phi^{2(\nu+i-\ell)}}{1-r^2}\right]. \end{aligned} \tag{72}$$

The remaining calculations are similar. If $\ell = i$, the second term (with $m_1 = 0$ and $m_2 = 1$) is

$$(W_{01})_{i,\ell-1} \;=\; 2\gamma(0)^2 \left[\frac{\phi + \phi^{2\nu-1}}{1-r^2}\right]. \tag{73}$$

If $\ell > i$ we have

$$(W_{01})_{i,\ell-1} \;=\; 2\gamma(0)^2 \left[\frac{\phi^{2\ell-2i-1}}{1-r^2} + \frac{\phi^{2\nu-2\ell+2i+1}}{1-r^2}\right].$$

If $\ell = i$ the second and third terms are the same, and we get

$$s_{i\ell} = 2\gamma(0)^2 \left[\frac{\phi^{2(\ell-i)} + \phi^{2(\nu+i-\ell)}}{1-r^2}\right] - 2\phi \times 2\gamma(0)^2 \left[\frac{\phi + \phi^{2\nu-1}}{1-r^2}\right] + \phi^2\gamma(0)^2 \left[\phi^2 + \frac{1+3r^2}{1-r^2}\right]. \quad (74)$$

If $\ell > i$ we obtain

$$\begin{aligned} s_{i\ell} \quad = \quad & 2\gamma(0)^2 \left[\frac{\phi^{2(\ell-i)} + \phi^{2(\nu+i-\ell)}}{1-r^2}\right] - 2\gamma(0)^2\phi \left[\frac{\phi^{2\ell-2i-1}}{1-r^2} + \frac{\phi^{2\nu-2\ell+2i+1}}{1-r^2}\right] \\ & -2\gamma(0)^2\phi \left[\frac{\phi^{2\ell-2i+1} + \phi^{2\nu-2\ell+2i-1}}{1-r^2}\right] + 2\gamma(0)^2\phi^2 \left[\frac{\phi^{2\ell-2i} + \phi^{2\nu-2\ell+2i}}{1-r^2}\right]. \end{aligned} \quad (75)$$

Since $\Sigma_{\boldsymbol{\sigma}^2}$ is symmetric, this completes the calculation.

5 High frequency climate data

A North American Regional Climate Change Assessment Program (NARCCAP) simulation used a global circulation climate model called CCSM from the USA National Center for Atmospheric Research (NCAR) in Boulder, Colorado, together with a detailed local climate model (CRCM) from the Canadian Centre for Climate Modelling and Analysis. This produced a complete data set (no missing values) of maximum daily surface air temperature (°K) on a 140×115 grid over much of North America, for a period of $N = 29$ years, from 2041 through 2069. Each year has 365 days. We consider a single (typical) site on this grid, located at -98.229 degrees west and 37.783 degrees north. This is about 60 miles west of Wichita, Kansas.

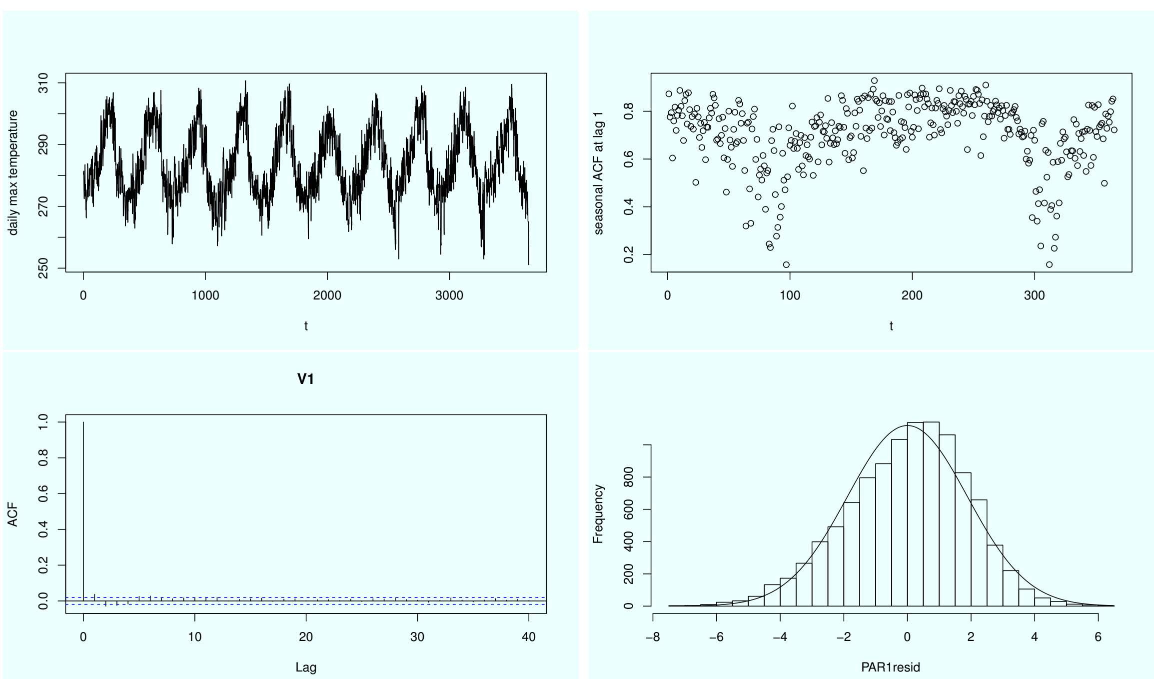

Figure 1: Upper left panel: The first ten years of daily maximum surface temperature data indicate strong seasonal variations. Upper right panel: The autocorrelation function also has strong seasonal variations. Lower left panel: The autocorrelation function of the model residuals are consistent with white noise. Lower right panel: The model residuals fit a Gaussian distribution.

The time series, $\tilde{X}_t$, records the maximum daily surface temperature on the t^{th} day, with $t = 1$ corresponding to January 1, 2041. The sample size is $n = 29 \times 365 = 10,585$. The first 10 years of data, shown in the upper left panel of Figure 1, exhibits significant seasonal variation. The right panel shows the seasonal autocorrelation function

$$\rho_t(m) = \frac{\gamma_t(m)}{\sqrt{\gamma_t(0)\gamma_{t+m}(0)}} \tag{76}$$

at lag $m = 1$. Since this plot also varies significantly with the time of year, neither differencing nor subtracting the seasonal mean will produce a second order stationary time series. Hence a periodically stationary time series model is indicated. We will attempt to fit this data using a PARMA$_\nu(p,q)$ with $\nu = 365$ days per year, $p = 1$ autoregressive terms, and $q = 0$ moving average terms in Equation (1). The calculations for this model were detailed at the end of Section 4. This PAR$_{365}(1)$ model can be written in the simplified form

$$X_t = \phi_t X_{t-1} + \varepsilon_t \tag{77}$$

where $X_t = \tilde{X}_t - \mu_t$, and $\{\varepsilon_t\}$ are independent with mean zero and standard deviation σ_t. The innovations algorithm was used to compute the estimates $\hat{\boldsymbol{\sigma}}^2$ and $\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\psi}}(1)$ of the model parameters. For $N = 29$ years of data, we found that $k = 5$ iterations of the innovations algorithm were sufficient, since the parameter estimates settled down around this value. The standardized model residuals (estimates of the noise series) were then computed using

$$\hat{\epsilon}_t = \frac{\hat{X}_t - \hat{\phi}_t \hat{X}_{t-1}}{\hat{\sigma}_t} \tag{78}$$

where $\hat{X}_t = \tilde{X}_t - \hat{\mu}_t$ and $\hat{\mu}_t$ is given by (3). The autocorrelation function of the model residuals, shown in the lower left panel of Figure 1, indicates no serial correlation. Hence the PAR(1) model is sufficient to capture the dependence structure in this time series data. The histogram in the lower right panel of Figure 1 indicates that the residuals are reasonably well fit by a normal distribution. Hence we consider the PAR$_{365}(1)$ model with normal innovations to be an adequate representation of this time series data.

While the PAR$_{365}(1)$ model gives a reasonable fit to the data, the model has $3 \times 365 = 1095$ parameters $[(\mu_t, \sigma_t^2, \phi_t) : 0 \le t < 365]$. Next we will use the methods of this paper to obtain a significant reduction in the number of parameters, leading to a parsimonious time series model for this climate data. Table 1 shows the results of applying the tests from Section 4 to the climate data. The test statistics (66) were computed using the diagonal entries of the covariance matrices $\Sigma_{\boldsymbol{\mu}}$, $\Sigma_{\boldsymbol{\sigma}^2}$, and $\Sigma_{\boldsymbol{\phi}}$ The table shows all the harmonics of order $r \le 4$. The only significant harmonic of order $r > 4$ was $\hat{c}_{\phi 26} = 0.055$ with $z = 3.992$. Using only the statistically significant harmonics, we arrive at the Fourier-PARMA model

$$\begin{aligned}
\mu_t &= 284.132 - 13.304\cos\left(\frac{2\pi t}{365}\right) - 2.716\sin\left(\frac{2\pi t}{365}\right) + 3.312\sin\left(\frac{4\pi t}{365}\right) \\
\sigma_t^2 &= 10.742 + 1.522\sin\left(\frac{2\pi t}{365}\right) - 1.236\cos\left(\frac{4\pi t}{365}\right) + 1.437\sin\left(\frac{4\pi t}{365}\right) + 1.257\cos\left(\frac{6\pi t}{365}\right) \\
\phi_t &= 0.761 - 0.164\cos\left(\frac{2\pi t}{365}\right) - 0.054\sin\left(\frac{2\pi t}{365}\right) + 0.133\cos\left(\frac{4\pi t}{365}\right) + 0.063\sin\left(\frac{4\pi t}{365}\right) \\
&+ 0.078\cos\left(\frac{6\pi t}{365}\right) + 0.073\sin\left(\frac{6\pi t}{365}\right) + 0.055\cos\left(\frac{52\pi t}{365}\right).
\end{aligned}$$

Figure 2 shows the estimated parameter vectors $[(\hat{\mu}_t, \hat{\sigma}_t^2, \hat{\phi}_t) : 0 \le t < 365]$ together with the smoothed Fourier model. The bottom right panel in Figure 2 shows the autocorrelation function of the model residuals. The plot shows no evidence of serial dependence, and hence we conclude that

Table 1: First line: Low order harmonics of the discrete Fourier transform for the sample mean, variance, and ϕ-weights in the $\mathrm{PAR}_{365}(1)$ model for the time series of daily high surface temperature near Wichita, Kansas. Second line: Test statistic z in parentheses. The $*$ denotes statistically significant harmonics, with test statistic $|z| > 3.81$.

	DFT harmonics r				
	0	1	2	3	4
$\hat{c}_{\mu r}$	284.132*	-13.304*	0.734	-0.056	-0.731
	(1961.156)	(-65.061)	(3.609)	(-0.280)	(-3.678)
$\hat{s}_{\mu r}$		-2.716*	3.312*	-0.758	0.241
		(-13.284)	(16.293)	(-3.767)	(1.212)
	0	1	2	3	4
$\hat{c}_{\sigma^2 r}$	10.742*	1.146	-1.236*	1.257*	0.183
	(47.449)	(3.598)	(-3.859)	(3.926)	(0.571)
$\hat{s}_{\sigma^2 r}$		1.522*	1.437*	0.207	-0.558
		(4.752)	(4.487)	(0.646)	(-1.744)
	0	1	2	3	4
$\hat{c}_{\phi r}$	0.761*	-0.164*	0.133*	0.078*	0.014
	(78.309)	(-11.958)	(9.706)	(5.656)	(1.036)
$\hat{s}_{\phi r}$		-0.054*	0.063*	0.073*	-0.033
		(-3.909)	(4.556)	(5.340)	(-2.413)

*Fourier coefficients with test statistic $|z| \geq 3.81$

this Fourier-PARMA model provides a reasonable, parsimonious fit to the climate data, with only 16 parameters. In our view, the estimated parameter vectors are “fitting the noise” in this data, and the smoothed Fourier models uncover the basic structure.

Next we consider a reduced model, with a smaller number of parameters. We discard the terms $\hat{c}_{\sigma^2 2} = -1.236$ (-3.859), $\hat{c}_{\sigma^2 3} = 1.257$ (3.926), $\hat{c}_{\phi 1} = -0.054$ (-3.909) and $\hat{c}_{\phi 26} = 0.055$ (3.992) whose test statistic is barely over the cutoff of 3.81. The resulting plots for the mean and variance appear unchanged. Figure 3 shows the resulting model of the ϕ-weights, and the autocorrelation plot of the model residuals. The ϕ plot is smoother, since we zeroed out the high frequency term. The autocorrelation plot is consistent with white noise, and hence we conclude that the reduced model is adequate. Given the complexity of the NARCCAP model, we consider this modeling exercise a success. Using a parsimonious Fourier-PARMA model with only 12 parameters, we are able to capture all the essential features of this time series.

Remark 5.1 *For a purely autoregressive PAR(p) model, it should also be possible to develop Fourier-PARMA modeling using Yule-Walker estimates instead of the innovations algorithm. The parameter estimates will be different, and the asymptotic variance-covariance matrices will change, so the entire theory would have to be developed from first principles. One possible advantage of this approach would be that the nuisance parameter k in the innovations algorithm need not be selected. On the other hand, the innovations algorithm applies to a general PARMA(p,q) model.*

An alternative approach to the Fourier-PARMA model developed in this paper considers the z-scores as a stationary time series, e.g., see French et al. (2019). Compute the seasonal mean $\hat{\mu}_t$ using (3), the seasonal variance $\hat{\gamma}_t(0)$ using (27), and then consider the time series

$$Z_t = \frac{\tilde{X}_t - \hat{\mu}_t}{\hat{\gamma}_t(0)} \tag{79}$$

of z-scores. The autocorrelation function of this time series, shown in the left panel in Figure 4, suggests long range dependence, which can be captured using an ARFIMA (fractionally integrated

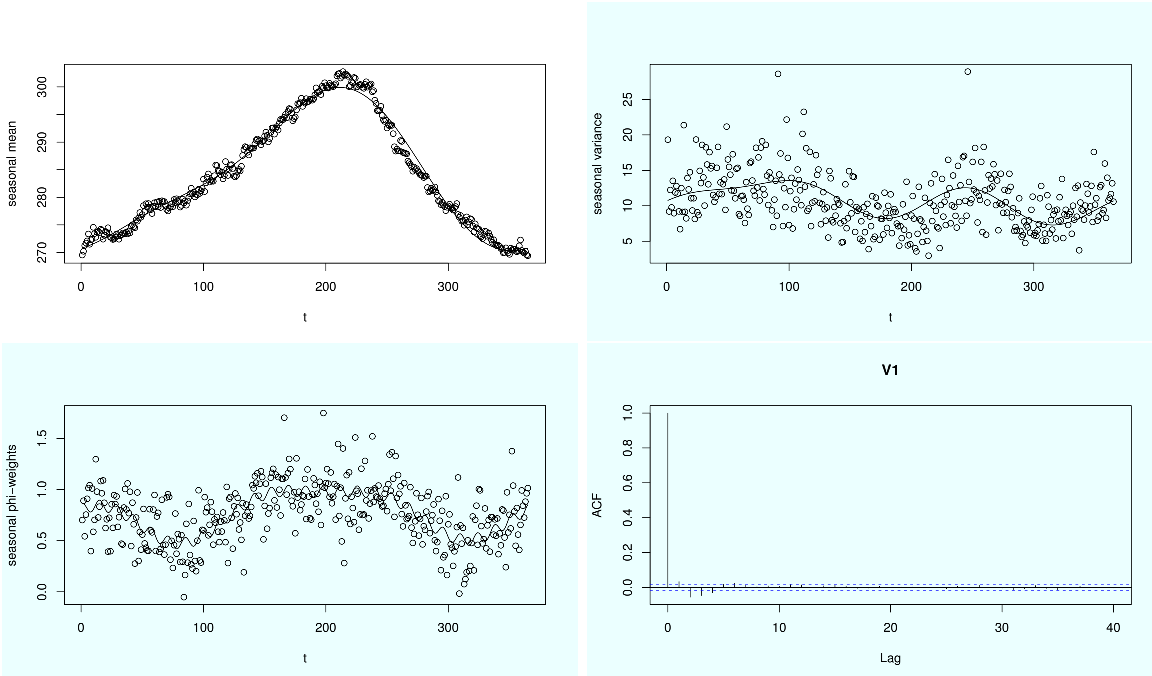

Figure 2: Parameters for the $\mathrm{PAR}_{365}(1)$ model (77) of daily maximum surface temperature near Wichita, Kansas: Upper left, the seasonal mean μ_t; upper right, the seasonal variance σ_t^2 of the noise term ε_t; lower left, the seasonal autoregressive parameters ϕ_t; and lower right, the autocorrelation function of the model residuals using the discrete Fourier transform model of the parameters.

ARMA) model. Using the `arfima` package in `R`, a good fit is obtained using a fractionally integrated white noise with fractional differencing parameter $d = 0.495$, indicating a strong long range dependence. The spectral density of the z-scores is shown in the right panel of Figure 4. It appears to follow a straight line with slope $-2d$ on a log-log plot, consistent with the long range dependent ARFIMA$(0, d, 0)$ model. A different model using *tempered* fractional differencing is discussed in Example 4.5 of Sabzikar, McLeod, and Meerschaert (2019). In our experience, the autocorrelation function of periodically stationary PARMA time series can often mimic long range dependence. Therefore, when a time series with strong seasonal variations shows the hallmarks of long range dependence, it may also be worth while to consider an alternative PARMA model.

Remark 5.2 *Another possible source of nonstationarity in this NARCCAP time series is a warming trend. We tested for this by regressing the data against the time variable. The slope term 0.000130 is statistically significant (*$p = 0.000$*) and represents a mean temperature change of 1.4* $^\circ$*K over 29 years. Using the regression equation temp* $= 283.442 + 0.000130t$*, we repeated the Fourier-PARMA modeling using* $y = 283.442 +$ *residual. The resulting model has the same nonzero harmonics, and the coefficients are close. For example, the coefficients for the mean are 283.442, -13.304, -2.701, and 3.320.*

6 Data Availability Statement

The data used in Section 4 are freely available at `http://www.narccap.ucar.edu/` from the North American Regional Climate Change Assessment Program (NARCCAP). The data can also be downloaded in compressed `R` format from `https://data.mendeley.com/datasets/jz553h7ytw/1`, using array indices `c(68, 34)`, or from the authors.

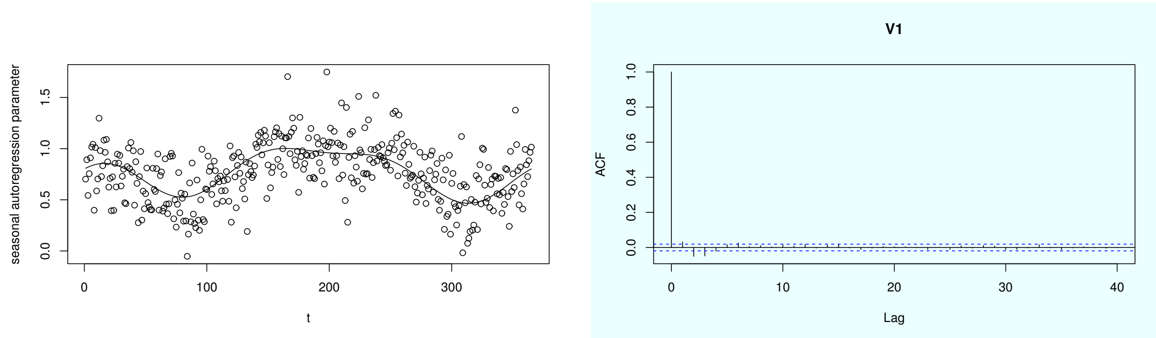

Figure 3: Left panel: reduced model for the seasonal autoregressive parameter ϕ_t. Right panel: autocorrelation function of the model residuals.

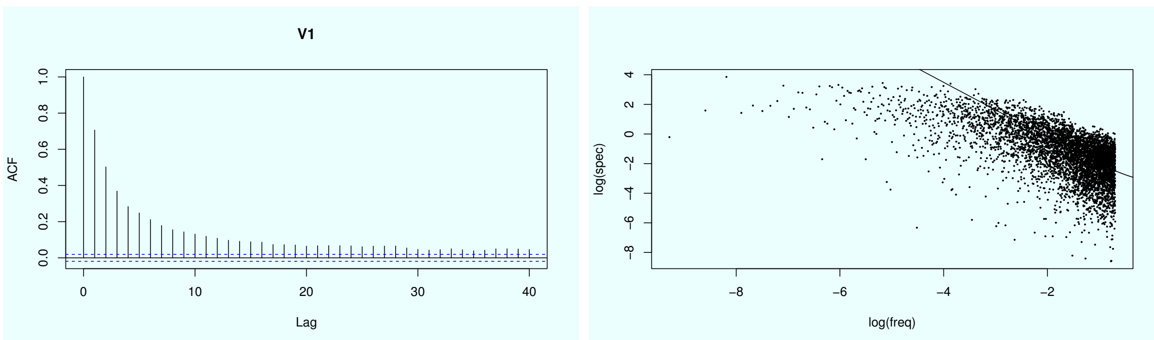

Figure 4: Left panel: The autocorrelation function of the z-scores suggests long range dependence. Right panel: The corresponding power spectrum follows a straight line on a log-log plot, providing further evidence of long range dependence.

References

Adams, G.J and G.C. Goodwin (1995), Parameter estimation for periodically ARMA models. *Journal of Time Series Analysis*, *16*(2), 127–145.

Anderson, P.L. and A.V. Vecchia (1993), Asymptotic results for periodic autoregressive moving -average processes. *Journal of Time Series Analysis, 14*, 1–18.

Anderson, P.L. and M.M. Meerschaert (1997), Periodic moving averages of random variables with regularly varying tails. *Annals of Statistics, 25*, 771–785.

Anderson, P.L. and M.M. Meerschaert (1998), Modeling river flows with heavy tails. *Water Resources Research* , *34*(9), 2271–2280.

Anderson, P.L., M.M. Meerschaert and A.V. Veccia (1999), Innovations algorithm for periodically stationary time series. *Stochastic Processes and their Applications, 83*, 149–169.

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

Anderson, P.L., Y.G. Tesfaye and M.M. Meerschaert (2007), Fourier-PARMA models and their application to river flows. *Journal of Hydrologic Engineering*, *12*(7), 462–472.

Anderson, P.L., L. Kavalieris and M.M. Meerschaert (2008), Innovations algorithm asymptotics for periodically stationary time series with heavy tails. *Journal of Multivariate Analysis, 99*, 94–116.

Basawa, I. V., R.B. Lund and Q. Shao (2004), First-order seasonal autoregressive processes with periodically varying parameters. *Statist. Probab. Lett., 67*(4), 299–306.

Berk, K. (1974), Consistent autoregressive spectral estimates. *Ann. Statist.*, 2, 489–502.
Bhansali, R. (1978), Linear prediction by autoregressive model fitting in the time domain. *Ann. Statist.*, 6, 224–231.
Box, G.E., and G.M. Jenkins (1976), *Time Series Analysis: Forcasting and Control*, 2nd edition, San Francisco: Holden-Day.
Brockwell, P.J. and R.A. Davis (1988), Simple consistent estimation of the coefficients of a linear filter. *Stoch. Proc. Appl.*, 28, 47–59.
Brockwell, P.J. and R.A. Davis (1991), *Time Series: Theory and Methods*, 2nd ed. New York: Springer-Verlag.
Dudek, A.E., H. Hurd and W. Wojtowicz, (2016), Periodic autoregressive moving average methods based on Fourier representation of periodic coefficients. *WIREs Comput Stat*, 8, 130-149.
Franses, P.H. and Paap, R. (2011) , Random-coefficient periodic autoregressions. *Statistica Neerlandica*, 65, 101-115.
Feller, W. (1971), *An introduction to Probability Theory and Its Applications*. Second edition, John Wiley & Sons, New York.
French, J., P. Kokoszka, S. Stoev and L. Hall, (2019), Quantifying the risk of heat waves using extreme value theory and spatio-temporal functional data. *Computational Statistics and Data Analysis 131*, 176–193.
Jones, R. H. and W. M. Brelsford (1967), Times series with periodic structure. *Biometrika, 54*, 403–408.
Lund, R. B. and I. V. Basawa (1999), Modeling and inference for periodically correlated time series. *Asymptotics, Nonparameterics, and Time sereis*(ed. S. Ghosh), 37–62. Marcel Dekker, New York.
Lund, R. B. and I. V. Basawa (2000), Recursive prediction and likelihood evaluation for periodic ARMA models. *Journal of Time Series Analysis*, *20*(1), 75–93.
Meerschaert, M.M. and H.P. Scheffler (2001), *Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice*, Wiley Interscience, New York.
Mei, J., Shao, Q., Liu, R. (2017), Efficient inference for parameters of unobservable periodic autoregressive time series. *Communications in Statistics - Theory and Methods*, 46, Issue 15.
Pagano, M. (1978), On periodic and multiple autoregressions. *Annals of Statistics*, *6*(6), 1310–1317.
Sabzikar, F., A.I. McLeod, and M.M.Meerschaert, (2019), Parameter estimation for ARTFIMA time series. *J. Statist. Plann. Inf. 200*, 129–145.
Salas, J.D, J.W. Delleur, V. Yevjevich and W.L. Lane (1980), Applied modeling of hydrologic time series. *Water Resource Publications*, Littleton, Colorado.
Salas, J.D, J. T.B Obeysekera and R.A. Smith (1981), Identification of streamflow stochastic models. *ASCE Journal of the Hydraulics Division*, *107*(7), 853–866.
Salas, J.D, D.C. Boes and R.A. Smith (1982), Estimation of ARMA models with seasonal parameters. *Water Resources Research*, *18*, 1006–1010.
Salas, J.D, G.Q. Tabios III and P. Bartolini (1985), Approaches to multivariate modeling of water resources time series. *Water Resources Bulletin*, *21*, 683–708.
Samorodnitsky, G., and Taqqu, M. (1994), *Stable non-Gaussian Random Processes*, Chapman & Hall, New York.
Shao, Q. and R.B. Lund (2004), Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *Journal of Time Series Analysis*, *25*(3), 359–372.
Tesfaye, Y.G., M.M. Meerschaert and P.L. Anderson (2005), Identification of PARMA models and their application to the modeling of river flows. *Water Resour. Res.*,*42*(1), W01419.
Tesfaye, Y.G., P.L. Anderson, and M.M. Meerschaert (2011), Asymptotic results for Fourier-PARMA time series. *Journal of Time Series Analysis, 32*, 157-174.
Tiao, G.C. and M.R. Grupe (1980), Hidden periodic autoregressive moving average models in

time series data. *Biometrika 67*, 365-373.
Tjøstheim, D. and J. Paulsen (1982), Empirical identification of multiple time series. *Journal of Time Series Analysis, 3*, 265–282.
Thompstone, R.M., K.W. Hipel and A.I. McLeod (1985), Forecasting quarter-monthly riverflow. *Water Resources Bulletin*, *25*(5),731–741.
Troutman, B.M. (1979), Some results in periodic autoregression. *Biometrika, 6*, 219–228.
Ula, T.A. (1990), Periodic covariance stationarity of multivariate periodic autoregressive moving average processess. *Water Resources Research*, *26*(5), 855–861.
Ula, T.A. (1993), Forcasting of multivariate periodic autoregressive moving average processes. *Journal of Time Series Analysis, 14*, 645–657.
Ula, T.A. and A.A Smadi (1997), Periodic stationarity conditions for periodic autoregressive moving average processess as eigenvalue problems. *Water Resources Research*, *33*(8), 1929–1934.
Ula, T.A. and A.A Smadi (2003), Identification of periodic moving average models. *Communications in Statistics: Theory and Methods, 32*(12), 2465–2475.
Vecchia , A.V. (1985a), Periodic autoregressive-moving average (PARMA) modelling with applications to water resources. *Water Resources Bulletin 21*, 721–730.
Vecchia, A.V. (1985b), Maximum likelihood estimation for periodic moving average models. *Technometrics*, *27*(4), 375–384.
Vecchia , A.V. and R. Ballerini (1991), Testing for periodic autocorrelations in seasonal time series data. *Biometrika*, *78*(1), 53–63.