Statistics and Its Interface Volume 10 (2017) 71–79

Semiparametric analysis for environmental time series

Lin Tang and Qin Shao*

Time series that contain a trend, a seasonal component and periodically correlated time series are commonly encountered in environmental sciences. A semiparametric three-step method is proposed to analyze such time series. The seasonal component and trend are estimated by means of B-splines, and the Yule-Walker estimates of the time series model coefficient are calculated via the residuals after removing the estimated seasonality and trend. The oracle efficiency of the proposed Yule-Walker type estimators is established. Simulation studies suggest that the performance of the estimators coincide with the theoretical results. The proposed method is applied to the monthly global temperature data provided by the National Space Science and Technology Center.

Keywords and phrases: Periodic autoregressive time series, Partially linear models, Yule-Walker estimators, B-splines, Oracle efficiency, Confidence band, Trend, Seasonality.

1. INTRODUCTION

Time series of environmental sciences often contain trend, seasonality, and periodically correlated random components due to the seasonal and periodic nature of dynamical systems. To emphasize such a cycling pattern, the time index t is often written as $iT+\nu$, where i is an integer, T is the period, and ν $(1 \leq \nu \leq T)$ is called the season. Reference [9] provided a comprehensive framework for modeling such time series. In particular, the periodically correlated random terms $\{x_t\}_{t=1}^n$ are often well described by the following periodic autoregressive model with order p (PAR(p))

$$x_{iT+\nu} - \sum_{k=1}^{p} \phi_k(\nu) x_{iT+\nu-k} = \sigma_\nu \epsilon_{iT+\nu}, \tag{1}$$

where $\{\epsilon_{iT+\nu}\}$ is white noise with $\mathrm{E}(\epsilon_{iT+\nu}) = 0$ and $\mathrm{Var}(\epsilon_{iT+\nu}) = 1$. A PAR($p$) becomes an autoregressive time series (AR(p)) when $T = 1$.

This paper is motivated by the analysis of the monthly global temperature data from January 1979 to December 2014 provided by the National Space Science and Technology Center. It seems from the time series plot in Figure 4 that the data set contains a trend and possibly seasonal means. A great deal of research has been done from the parametric approach when the trend $g(\cdot)$ follows a known analytic function with some unknown parameters and the error sequence $\{x_{iT+\nu}\}$ is time series, for example, [23, 4, 1, 12]. In particular, [1] took the seasonal-varying constant trend into account for periodic autoregressive error terms; [12] modeled the trend using a trigonometric function. Although parametric trend estimation works well when the shape of a trend can be approximated by a known analytical function, its major drawback is that the assumed model is usually subjective and is likely misspecified.

The temperature data in Figure 4, however, do not reveal that the trend follows the shape of any well-known analytic function. Moreover, given the seasonal nature of the observations, it is possible that the monthly means are not constant. The semiparametric approach we propose in the paper aims at a more flexible modeling alternative which does not require a trend function to have an explicit format, but only a certain degree of smoothness. We generalize the partially linear models proposed in [5] to the following periodically correlated error terms:

$$y_{iT+\nu} = \sum_{k=1}^{T-1} I_k(\nu)\beta_k + g(u_{iT+\nu}) + x_{iT+\nu}, \tag{2}$$

where $u_t = t/n$; $g(\cdot)$ represents the unknown smooth trend function defined in the interval $[0,1]$; $I_k(\nu)$ is the indicator function for season ν with $I_k(\nu) = 1$ if $k = \nu$ and $I_k(\nu) = 0$ otherwise; $\{\beta_k\}_{k=1}^{T-1}$ are the seasonal effects compared with the reference level $\nu = T$; $\{x_t\}_{t=1}^n$ is a PAR(p) time series defined by (1). Define an $n \times (T-1)$ matrix $\boldsymbol{D} = (\boldsymbol{D}_0, \boldsymbol{D}_0, \ldots, \boldsymbol{D}_0)'$, where a $T \times (T-1)$ matrix $\boldsymbol{D}_0 = (\boldsymbol{I}, \boldsymbol{0})'$ with $\boldsymbol{I}$ being the $(T-1)\times(T-1)$ identity matrix and $\boldsymbol{0}$ the $(T-1)$-dimensional zero vector. The partially linear model (2) can be rewritten in vector format as

$$\boldsymbol{y} = \boldsymbol{D}\boldsymbol{\beta} + \boldsymbol{g} + \boldsymbol{x},$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{T-1})'$, $\boldsymbol{g} = (g(u_1), g(u_2), \ldots, g(u_n))'$, and $\boldsymbol{x} = (x_1, \ldots, x_n)'$. Throughout the paper, we will use bold lower case letters for vectors and bold upper case letters for matrices.

*Corresponding author.

Partially linear models have received considerable attention due to their flexibility and wide applications. For example, see [8, 21, 7, 24]. In particular, [21] provided a detailed discussion about inference for partially linear models with independent and identically distributed error terms. Reference [17] took correlation in the error terms into consideration and applied partially linear models to analyzing time series using local linear smoothing for trends. Our proposed semiparametric three-step approach not only can be used to estimate all the components in model (2), but can be applied to making inference about them as well. In particular, we provide a theoretical justification for residual based Yule-Walker estimators for the time series model coefficients $\boldsymbol{\phi}$ by generalizing the results in [19] and [15] for the oracle efficiency of Yule-Walker estimators for autoregressive coefficients when time series observations contain a trend. It is straightforward to construct a 95% confidence interval for $\boldsymbol{\phi}$ and $\boldsymbol{\beta}$ based on Theorems 2.1–2.2, and a 95% confidence band for $g(\cdot)$ from the B-spline residuals using the method in [18] for AR(p). We will illustrate how to apply the procedure by analyzing the monthly global temperature data in Section 4.

The essence of the proposed three-step method is to replace the unobservable time series $\{x_t\}_{t=1}^n$ by the residuals of the B-spline estimates in partially linear models: in the first step, $\boldsymbol{\beta}$ are estimated based on B-splines; in the second step, $g\,(u)$ is estimated and the residuals are calculated by subtracting the estimates $\hat{g}\,(u)$ and $\hat{\boldsymbol{\beta}}$ from the observations $\{y_t\}_{t=1}^n$; in the third step, $\boldsymbol{\phi}$ is estimated from the residuals in the second step. One of the advantages of the procedure is that it is very easy to implement. Practitioners can utilize any software package that has built-in functions for AR and PAR. For example, the simulation studies and data analysis of this paper are accomplished using "`ar`" for autoregressive time series and "`pear`" ([13]) for periodic autoregressive time series. They are packages in R which is an open access environment for statistical computing and graphics developed by [16].

The paper is organized as follows. In Section 2, we will consider how to estimate the trend function $g\,(u)$ and seasonal effects $\boldsymbol{\beta}$ by B-spline smoothing, and how to calculate the Yule-Walker estimates for $\boldsymbol{\phi}$ from the residuals. In particular, we will provide the asymptotic properties of the estimators for the constant B-spline. In Section 3, we will illustrate implementation of the procedure and the performance of the estimators by simulation studies. In Section 4, we will apply the method to the monthly global temperature data which is the motivation for this research. In Section 5, we will summarize the paper and provide some concluding remarks. The proofs of the theoretical results in Section 2 are given in the Appendix.

2. CONSTRUCTION OF ESTIMATORS

In this section, we will discuss the details of how to estimate $g\,(u)$ and $\boldsymbol{\beta}$ by B-spines, and how to calculate $\hat{\boldsymbol{\phi}}$ from residuals $\{\hat{x}_t\}_{t=1}^n$.

Consider a sequence of equally spaced points $\{(-m+1)h, (-m+2)h, \ldots, -h, 0, h, 2h, \ldots, Nh, 1\}$, where m is a positive integer. We will only provide the results for $m=1$ for simplicity. The interval $[0,1]$ is divided into $N+1$ subintervals of equal length $h = (N+1)^{-1}$ as $J_j = [jh, (j+1)h), j = 0, 1, 2, \ldots, N-1$ and $J_N = [Nh, 1]$. Let $G_N^{(m-2)} = G_N^{(m-2)}[0,1]$ denote the space of functions that are polynomial functions of degree $m-1$ on each J_j and have continuous $(m-2)$-th derivatives. Then the B-spline basis of $G_N^{(m-2)}$ is $\{b_{j,m}(u), j = -m+1, \ldots, N\}$. We will discuss two cases: $G_N^{(-1)}$ is constant on each J_j, where $m=1$; $G_N^{(0)}$ is linear on each J_j and continuous on $[0,1]$, where $m=2$. For case $G_N^{(-1)}$, the B-spline basis is $\{b_{j,1}(u)\}_{j=0}^N$, where $b_{j,1}(u)$ is defined as follows:

$$b_{j,1}(u) = \begin{cases} 1, & u \in J_j, \\ 0, & \text{otherwise.} \end{cases}$$

For case $G_N^{(0)}$, the B-spline basis is $\{b_{j,2}(u)\}_{j=0}^N$, where $b_{j,2}(u)$ is defined as follows:

$$b_{j,2}(u) = K\left(\frac{u-(j+1)h}{h}\right),$$

where $K(x) = (1-|x|)_+$ with $(x)_+ = \max(x,0)$. Thus we can write down $b_{j,2}(u)$ specifically as follows:

$$b_{j,2}(u) = \begin{cases} \frac{u}{h} - j, & u \in J_j, \\ j+2-\frac{u}{h}, & u \in J_{j+1}, \\ 0, & \text{otherwise.} \end{cases}$$

For observations $\{y_t\}_{t=1}^n$, we define a vector $\boldsymbol{b}_j = (b_{j,m}(u_1), \ldots, b_{j,m}(u_n))'$ with $u_i = i/n$ and $n \times (N+m)$ matrix

$$\boldsymbol{B} = (\boldsymbol{b}_{-m+1}, \ldots, \boldsymbol{b}_N),$$

n_j is the number of observations in J_j $(j = 0, \cdots, N)$, and $j_i = \sum_{k=0}^{j-1} n_k + i$. Thus x_{j_i} is the i-th observation in the j-th interval. Therefore, $\sum_{j=0}^N n_j = n$. The spline smoother is

$$\boldsymbol{P_B} = \boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'.$$

Although we anticipate the estimation procedure and theorems in this section to hold for $m \geq 2$, we only provide here the results for $m=1$ for simplicity.

The seasonal effects $\boldsymbol{\beta}$ and the trend $g(\cdot)$ in the partially linear model (2) are respectively estimated by

$$\hat{\boldsymbol{\beta}} = \{\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{D}\}^{-1}\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{y}, \tag{3}$$

and

$$\hat{\boldsymbol{g}} = \boldsymbol{P_B}(\boldsymbol{y}-\boldsymbol{D}\hat{\boldsymbol{\beta}}). \tag{4}$$

The residual sequence $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_n)'$ is calculated by subtracting $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{g}}$ from the observations $\{y_t\}_{t=1}^n$:

(5) $$\hat{\boldsymbol{x}} = \boldsymbol{y} - \hat{\boldsymbol{g}} - \boldsymbol{D}\hat{\boldsymbol{\beta}}.$$

We will replace $\{x_t\}_{t=1}^n$ by $\{\hat{x}_t\}_{t=1}^n$ in the calculation of the Yule-Walker estimate $\hat{\boldsymbol{\phi}}$.

The autocovariance function of a PAR(p) time series $\{x_t\}_{t=-\infty}^{\infty}$ with mean zero is defined by $\gamma_\nu(k) = \mathrm{E}(x_{iT+\nu}x_{iT+\nu-k})$, which is determined by both the lag k and the season ν. Without loss of generality, we only consider the case where $n_T = n/T$ is an integer; that is, there are n_T cycles of observations. When $\{x_t\}_{t=1}^n$ were observable, the sample autocovariance at lag k would be computed as follows:

$$\tilde{\gamma}_\nu(k) = \frac{1}{n_T}\sum_{i=i_0}^{n_T-1} x_{iT+\nu}x_{iT+\nu-k}, \quad k \geq 0$$

where i_0 is the smallest integer such that $i_0T + \nu - k \geq 1$.

Let $\tilde{\boldsymbol{\Gamma}}_\nu$ be a $p \times p$ symmetric matrix, the (i, j)-th entry of which is $\tilde{\gamma}_{\nu-i}(j-i+1)$, $1 \leq i \leq j \leq p$, $\tilde{\boldsymbol{\Gamma}}$ be a $pT \times pT$ diagonal partitioned matrix with $\mathrm{diag}\left(\tilde{\boldsymbol{\Gamma}}\right) = (\tilde{\boldsymbol{\Gamma}}_1, \cdots, \tilde{\boldsymbol{\Gamma}}_T)$, and a pT dimensional vector $\tilde{\boldsymbol{\gamma}} = (\tilde{\boldsymbol{\gamma}}_1', \tilde{\boldsymbol{\gamma}}_2', \ldots, \tilde{\boldsymbol{\gamma}}_T')'$ with $\tilde{\boldsymbol{\gamma}}_\nu = (\tilde{\gamma}_\nu(1), \tilde{\gamma}_\nu(2), \cdots, \tilde{\gamma}_\nu(p))'$. Then the Yule-Walker estimators from $\{x_t\}_{t=1}^n$ are defined by

(6) $$\tilde{\boldsymbol{\phi}} = \tilde{\boldsymbol{\Gamma}}^{-1}\tilde{\boldsymbol{\gamma}} \text{ and } \tilde{\sigma}_\nu^2 = \tilde{\gamma}_\nu(0) - \sum_{k=1}^p \tilde{\phi}_k(\nu)\,\tilde{\gamma}_{\nu-k}(-k).$$

We use $\hat{\boldsymbol{\Gamma}}$ and $\hat{\boldsymbol{\gamma}}$ to denote $\tilde{\boldsymbol{\Gamma}}$ and $\tilde{\boldsymbol{\gamma}}$ when $\{x_t\}_{t=1}^n$ is replaced by the residuals $\{\hat{x}_t\}_{t=1}^n$ in (5). The proposed Yule-Walker estimators are

(7) $$\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\Gamma}}^{-1}\hat{\boldsymbol{\gamma}} \text{ and } \hat{\sigma}_\nu^2 = \hat{\gamma}_\nu(0) - \sum_{k=1}^p \hat{\phi}_k(\nu)\,\hat{\gamma}_{\nu-k}(-k).$$

Formulas (6) and (7) are very similar except that the time series $\{x_t\}_{t=1}^n$ is replaced by the residuals $\{\hat{x}_t\}_{t=1}^n$. Hereafter, we will use hat and tilde to represent the formulas based on $\{\hat{x}_t\}$ and $\{x_t\}$, respectively, unless otherwise indicated.

These estimators are not only computationally simple, but they have desirable asymptotic properties under very general conditions. For example, Theorem 2.2 below indicates that although these two estimators are different, $\hat{\boldsymbol{\phi}}$ is oracally or asymptotically equivalent to $\tilde{\boldsymbol{\phi}}$. We summarize these conditions and results below.

1. The trend function $g(\cdot) \in C^{(m)}[0, 1]$, $m = 1$; that is, the trend function has m continuous derivatives. In addition, the first derivative of $g(\cdot)$ is finite on the interval $[0, 1]$; that is, $g'(u) < \infty$ for every $u \in [0, 1]$.
2. The subinterval length $h \sim n^{-1/(2m+1)}$; that is, the number of interior knots $N \sim n^{1/(2m+1)}$.
3. The time series $\{x_{iT+\nu}, 1 \leq \nu \leq T\}_{i=0}^{n_T-1}$ is causal; that is, for each fixed ν $(1 \leq \nu \leq T)$, there exists a sequence of constants $\{\psi_j(\nu)\}_{j=0}^{\infty}$ such that $\sum_{j=0}^{\infty}|\psi_j(\nu)| < \infty$ and
$$x_{iT+\nu} = \sum_{j=0}^{\infty}\psi_j(\nu)\epsilon_{iT+\nu-j}.$$
4. $\mathrm{E}(\epsilon_t^4) < \infty$.

These conditions are very typical either for periodically stationary time series or for B-spline estimators. For example, Assumption 3 ensures $\sum_{|k|=0}^{\infty}|\gamma_\nu(k)| < \infty$ for each season ν, which is necessary for the asymptotic normality of the Yule-Walker estimator of $\boldsymbol{\phi}$ from $\{x_t\}_{t=1}^n$.

Theorem 2.1. *Under Assumptions 1–4, as $n_T \to \infty$,*

(8) $$\sqrt{n_T}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \stackrel{D}{\Longrightarrow} N(\boldsymbol{0}, \boldsymbol{V}),$$

where the (i, j)-th entry of the $(T-1) \times (T-1)$ matrix $\boldsymbol{V}$ is

$$(\boldsymbol{V})_{i,j} = \sum_{k=-\infty}^{\infty}\{\gamma_i(kT+i-j) - \gamma_i(kT+T-j) - \gamma_i(kT+i-T) + \gamma_i(kT)\}.$$

Theorem 2.2. *Under Assumptions 1–4, the Yule-Walker estimator $\hat{\boldsymbol{\phi}}$ defined in (7) and $\tilde{\boldsymbol{\phi}}$ defined in (6) satisfy*

$$\hat{\boldsymbol{\phi}} - \tilde{\boldsymbol{\phi}} = o_p(n^{-1/2}),$$

and

$$\hat{\sigma}_\nu^2 - \tilde{\sigma}_\nu^2 = o_p(1).$$

Under Assumptions 3–4, [14] showed that $\sqrt{n_T}(\tilde{\boldsymbol{\phi}} - \boldsymbol{\phi}) \stackrel{D}{\Longrightarrow} N(\boldsymbol{0}, \boldsymbol{\Sigma}^{-1})$ and $\tilde{\sigma}_\nu^2 \stackrel{P}{\longrightarrow} \sigma_\nu^2$ as $n_T \to \infty$, where $\boldsymbol{\Sigma}$ is the diagonal partitioned matrix with $\mathrm{diag}(\boldsymbol{\Sigma}) = (\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_T)$ and the (i, j)-th entry of $\boldsymbol{\Sigma}_\nu$ is $(\boldsymbol{\Sigma}_\nu)_{i,j} = \gamma_{\nu-i}(i-j)/\sigma_\nu^2$. Theorem 2.2 implies that $\hat{\boldsymbol{\phi}}$ and $\tilde{\boldsymbol{\phi}}$ are equivalent in terms of efficiency and both are $\sqrt{n}$-consistent estimators, which is given in the corollary below:

Corollary 2.1. *Under Assumptions 1–4, $\sqrt{n_T}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}) \stackrel{D}{\Longrightarrow} N(\boldsymbol{0}, \boldsymbol{\Sigma}^{-1})$ and $\hat{\sigma}_\nu^2 \stackrel{P}{\longrightarrow} \sigma_\nu^2$ as $n_T \to \infty$.*

We omit the proof of Corollary 2.1, as it is obvious according to Slusky's Theorem, while we postpone the detailed proofs of Theorems 2.1–2.2 to the Appendix.

3. IMPLEMENTATION AND SIMULATION STUDIES

In this section, we will illustrate the performance of the proposed procedure for the linear B-spline $m = 2$ using PAR(1) time series and the following trend function in [22]

(9) $$g(u) = \sin(2\pi u), u \in [0, 1].$$

We report the results for three PAR(1) models we adopt from [22] in Table 1. The sample sizes are n =

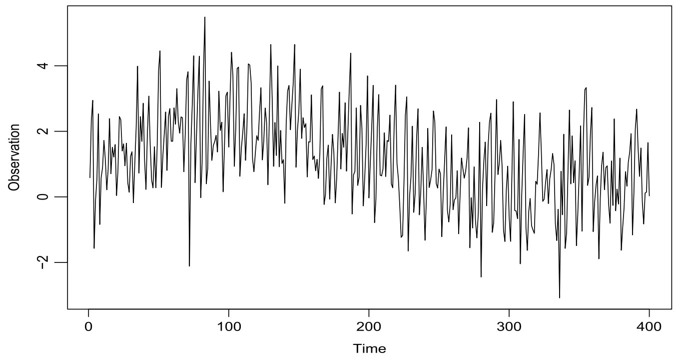

Figure 1. Generated Time Series Observations.

$200, 400, 800, 1600$ or $n_T = 50, 100, 200, 400$ with $T = 4$ so that the performance of the proposed estimators are illustrated for relatively small and large sample sizes. We simulate 100 sample paths of the time series from model (2) with

$$\boldsymbol{\beta} = (0.8, 1.5, 2.3)', \tag{10}$$

and the variances of periodic white noise $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2)' = (0.5,\ 0.7,\ 0.85,\ 1)'$. We omit the subscripts of $\boldsymbol{\phi}$, as there is only one coefficient for each season of PAR(1). Figure 1 is one sample path of the time series observations for Model 1 with $\boldsymbol{\phi} = (0.3, 0.6, 0.4, 0.2)'$ in Table 1 simulated by the following process:

1. generate PAR(1) time series $\{x_t\}_{t=1}^n$ based on the model coefficients and variances of periodic white noise;
2. obtain the observations $\{y_t\}_{t=1}^n$ based on model (2) with the trend function and the seasonal effects $\boldsymbol{\beta}$ respectively defined by (9) and (10);
3. estimate the seasonal effects $\boldsymbol{\beta}$ from (3) and the trend $g(\cdot)$ from (4) using the simulated observations $\{y_t\}_{t=1}^n$;
4. calculate the residual sequence $\hat{\boldsymbol{x}}$ from (5), and estimate the PAR model coefficients $\boldsymbol{\phi}$ from (7).

The number of knots is $N = [n^{1/5}]$ which is the same as in [22]. The sample means and sample standard deviations of $\hat{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{\beta}}$ are summarized in Table 1. It is worth pointing out that when the absolute values of the true coefficients $\boldsymbol{\phi}$ are close to or larger than one (e.g. Model 2), $\hat{\boldsymbol{\phi}}$ tends to be smaller than the true value and unstable with relatively large standard deviations, and the estimates of $\boldsymbol{\beta}$ perform well for all of these models. Overall, the bias and variability of both $\hat{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{\beta}}$ decrease when the sample size becomes larger.

We also calculate $\tilde{\boldsymbol{\phi}}$ based on time series $\{x_t\}_{t=1}^n$ without trend and seasonality and ratios $\{\tilde{\phi}(\nu)/\hat{\phi}(\nu)\}_{\nu=1}^4$. We include two of the boxplots of the ratios for illustration purpose in Figures 2–3. According to the boxplots, we conclude that these boxes become narrower and narrower and closer and closer to 1, as the sample size increases from 200 to 1600 for each season, which is in accordance with Theorem 2.2.

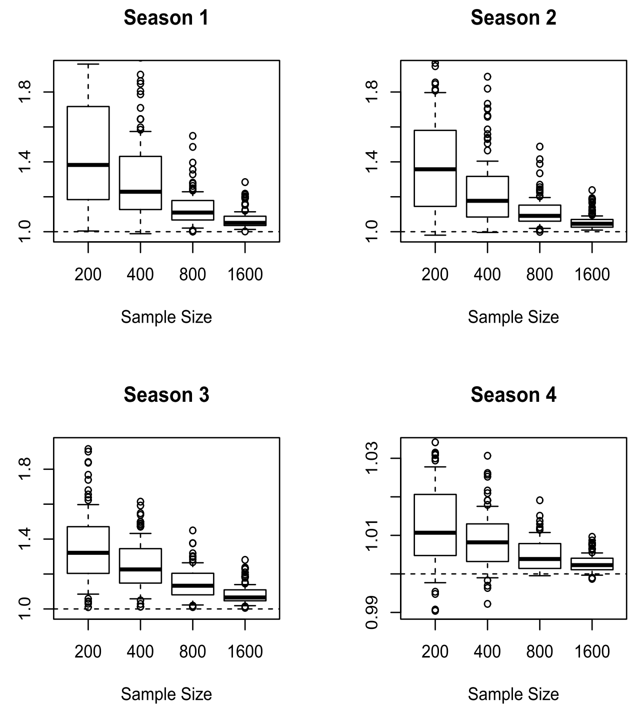

Figure 2. Boxplot of $\{\tilde{\phi}(\nu)/\hat{\phi}(\nu)\}_{\nu=1}^4$ for Model 2.

4. APPLICATION

In this section, we will apply our proposed method to the monthly global temperature data from January 1979 to December 2014. There are a total of $n_T = 36$ cycles for the $n = 432$ monthly temperatures with the period $T = 12$. Figure 4 is the scatter plot of the observations. The nonstationarity of the data is shown by a nonconstant trend of which the pattern is not clear in Figure 4 and a decreasing sample autocorrelation function in Figure 5. Thus the ARMA model class which works for the stationary time series is not appropriate for the data. Reference [3] recommended autoregressive integrated moving-average (ARIMA) models to handle time series with such a slowly decreasing autocorrelation function if the research interest is solely on forecasting. However, one of the interests here is whether there is a significant nonconstant trend. Therefore, we do not pursue analyzing the data by ARIMA.

We start with the full model (2), which includes a deterministic trend, seasonal components and a random error term. The seasonal effects and trend are respectively estimated by (3) and (4) with the linear B-spline ($m = 2$). The number of knots is $N = 4$ which is obtained by $N = [n^{1/5}]$, the same formula as in the simulation studies. We explore several AR models with different orders using the residuals $\{\hat{x}_t\}_{t=1}^{432}$, and an AR(2) model appears to be the most appropriate one with the smallest Akaike information criterion (AIC). The estimates with the standard errors for the full model (2) are $\hat{\phi}_1 = 0.565 \pm 0.047$, $\hat{\phi}_2 = 0.234 \pm 0.047$ for the AR(2) model coefficients, and $\hat{\beta}_1 = -0.002 \pm 0.124$, $\hat{\beta}_2 =$

Table 1. Estimate of Parameter ± Standard Deviation

Model 1: $(\phi_1, \phi_2, \phi_3, \phi_4) = (0.3, 0.6, 0.4, 0.2)$		
Sample Size	$\hat{\boldsymbol{\phi}}$	$\hat{\boldsymbol{\beta}}$
200	$(0.282 \pm 0.086,\ 0.569 \pm 0.163,\ 0.361 \pm 0.166,\ 0.169 \pm 0.162)$	$(0.776 \pm 0.148,\ 1.458 \pm 0.175,\ 2.272 \pm 0.181)$
400	$(0.285 \pm 0.079,\ 0.584 \pm 0.093,\ 0.379 \pm 0.092,\ 0.180 \pm 0.090)$	$(0.790 \pm 0.104,\ 1.475 \pm 0.123,\ 2.283 \pm 0.135)$
800	$(0.293 \pm 0.045,\ 0.594 \pm 0.082,\ 0.390 \pm 0.078,\ 0.184 \pm 0.075)$	$(0.796 \pm 0.073,\ 1.495 \pm 0.087,\ 2.291 \pm 0.088)$
1600	$(0.295 \pm 0.035,\ 0.597 \pm 0.056,\ 0.400 \pm 0.052,\ 0.198 \pm 0.047)$	$(0.797 \pm 0.051,\ 1.498 \pm 0.064,\ 2.300 \pm 0.058)$
Model 2: $(\phi_1, \phi_2, \phi_3, \phi_4) = (0.2, -2, -1.5, 0.9)$		
Sample Size	$\hat{\boldsymbol{\phi}}$	$\hat{\boldsymbol{\beta}}$
200	$(0.121 \pm 0.044,\ -1.360 \pm 0.382,\ -1.110 \pm 0.154,\ 0.878 \pm 0.048)$	$(0.826 \pm 0.546,\ 1.554 \pm 1.269,\ 2.276 \pm 0.150)$
400	$(0.153 \pm 0.027,\ -1.677 \pm 0.221,\ -1.218 \pm 0.127,\ 0.889 \pm 0.033)$	$(0.784 \pm 0.390,\ 1.468 \pm 0.934,\ 2.318 \pm 0.093)$
800	$(0.176 \pm 0.018,\ -1.818 \pm 0.123,\ -1.331 \pm 0.092,\ 0.897 \pm 0.023)$	$(0.790 \pm 0.273,\ 1.482 \pm 0.645,\ 2.305 \pm 0.076)$
1600	$(0.186 \pm 0.012,\ -1.900 \pm 0.085,\ -1.389 \pm 0.066,\ 0.899 \pm 0.015)$	$(0.793 \pm 0.205,\ 1.489 \pm 0.473,\ 2.298 \pm 0.053)$
Model 3: $(\phi_1, \phi_2, \phi_3, \phi_4) = (-0.1, -0.2, -0.4, -0.6)$		
Sample Size	$\hat{\boldsymbol{\phi}}$	$\hat{\boldsymbol{\beta}}$
200	$(-0.122 \pm 0.101,\ -0.218 \pm 0.181,\ -0.424 \pm 0.157,\ -0.626 \pm 0.133)$	$(0.750 \pm 0.188,\ 1.480 \pm 0.173,\ 2.230 \pm 0.253)$
400	$(-0.106 \pm 0.056,\ -0.210 \pm 0.115,\ -0.411 \pm 0.110,\ -0.612 \pm 0.100)$	$(0.783 \pm 0.138,\ 1.485 \pm 0.116,\ 2.285 \pm 0.185)$
800	$(-0.106 \pm 0.039,\ -0.207 \pm 0.079,\ -0.392 \pm 0.068,\ -0.608 \pm 0.073)$	$(0.809 \pm 0.100,\ 1.512 \pm 0.100,\ 2.306 \pm 0.126)$
1600	$(-0.102 \pm 0.029,\ -0.198 \pm 0.061,\ -0.400 \pm 0.053,\ -0.596 \pm 0.053)$	$(0.801 \pm 0.076,\ 1.503 \pm 0.062,\ 2.293 \pm 0.094)$

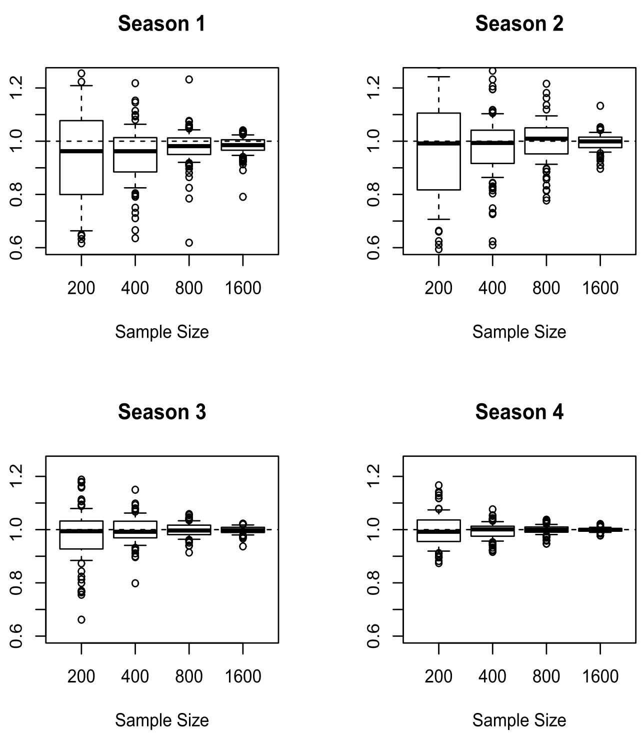

Figure 3. Boxplot of $\{\tilde{\phi}(\nu)/\hat{\phi}(\nu)\}_{\nu=1}^{4}$ for Model 3.

-0.012 ± 0.139, $\hat{\beta}_3 = -0.014 \pm 0.156$, $\hat{\beta}_4 = -0.005 \pm 0.165$, $\hat{\beta}_5 = -0.001 \pm 0.170$, $\hat{\beta}_6 = 0.011 \pm 0.171$, $\hat{\beta}_7 = -0.007 \pm 0.169$, $\hat{\beta}_8 = 0.0049 \pm 0.163$, $\hat{\beta}_9 = 0.019 \pm 0.154$, $\hat{\beta}_{10} = 0.011 \pm 0.136$, $\hat{\beta}_{11} = 0.002 \pm 0.120$ for the seasonal components.

Since none of these seasonal effects are significantly different from zero according to Theorem 2.1, we reduce the model to

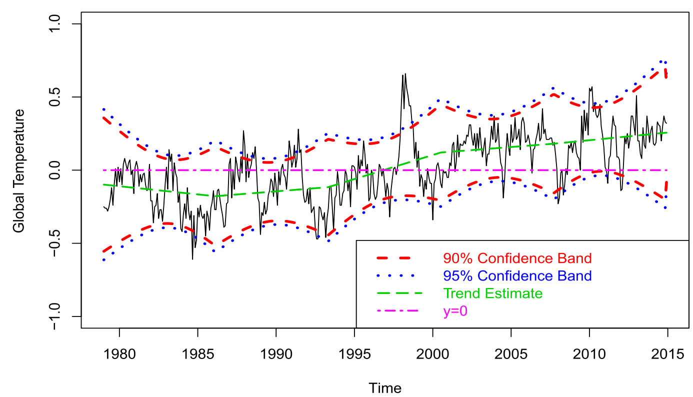

Figure 4. Monthly Global Temperature Data, January 1979–December 2014.

$$y_{iT+\nu} = g(u_{iT+\nu}) + x_{iT+\nu}. \tag{11}$$

The observations $\{y_t\}_{t=1}^{n}$ contain a trend in the reduced model. We use the two-step method proposed by [19] and extended by [22] to analyze $\{y_t\}_{t=1}^{n}$. In particular, we estimate the trend of the reduced model from $\hat{\boldsymbol{g}} = \boldsymbol{P_B y}$, and then calculate the residuals from $\hat{\boldsymbol{x}} = \boldsymbol{y} - \hat{\boldsymbol{g}}$. An AR(2) appears to be the best again for the residual sequence according to AIC. Model adequacy checking is conducted for the reduced model (11) with AR(2) residual autocorrelations. The residual sample autocorrelation function at lags 0–20 in Figure 6 shows no significant serial correlation. Therefore, the reduced model (11) with AR(2) is the most appropriate model for the monthly global temperature data. The Yule-Walker estimates are $\hat{\phi}_1 = 0.565 \pm 0.047$, $\hat{\phi}_2 = 0.233 \pm 0.047$. It is worth mentioning that the AR(2) model is not only adequate for the data, but also parsimonious compared with

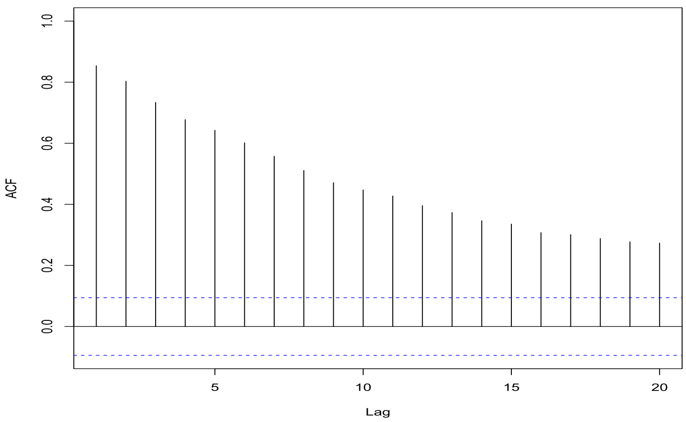

Figure 5. Sample Autocorrelations for Monthly Global Temperature Data.

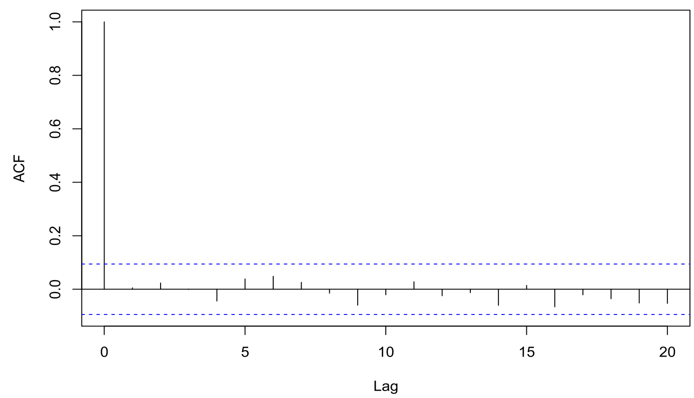

Figure 6. AR(2) Residual Autocorrelations for Monthly Global Temperature Data.

a PAR, as even a PAR(1) model with $T = 12$ has 12 time series model parameters.

Moreover, we obtain 90% and 95% confidence bands using the method of [18] based on the reduced model (11). These confidence bands which have at least those significance levels are conservative. In Figure 4, the dashed curve is the linear B-spline estimate of the trend, the dashed and dotted bands respectively have the significance levels 90% and 95%, and the dash-dot line is $y = 0$. It is interesting to note that the upward trend statistically is not very significant in that part of $y = 0$ is very close to the 95% confidence band or overlapped with the 90% confidence band. However, since there are $n_T = 36$ cycles in the data set, further research is needed when more observations are available.

5. CONCLUDING REMARKS

In this paper, we proposed a semiparametric three-step method for analyzing periodic time series with trend and seasonality. The proposed procedure is not only computationally accessible but theoretically well-justified for the constant B-spline. We anticipate that it can be generalized to higher order B-splines. We expect that the oracle efficiency showed in this paper can be extended to partially linear models of which trends are estimated by other nonparametric methods, such as smoothing splines (e.g. [11, 25]), under general conditions. Interested readers can refer to [6] for a comprehensive review about nonparametric estimation methods for time series. References [26, 27] compared how faster B-splines is than kernel estimation for functions. The simulation studies shows that our method works well for the linear B-spline estimation. Replacing unobservable error terms by residuals has been widely applied in time series analysis, see for example [2, 20]. However, an interesting and critical question is whether the estimators are oracally efficient; in other words, whether the analysis from the residuals $\{\hat{x}_t\}_{t=1}^n$ is asymptotically equivalent to the one based on $\{x_t\}_{t=1}^n$. For the partially linear model (2), we established the oracle efficiency of the model parameters, which is a desirable property and ensures that replacement is appropriate.

ACKNOWLEDGMENTS

The authors would like to thank the referees for the critical remarks and constructive suggestions which improved the manuscript considerably.

APPENDIX

In this section, we will show Theorems 2.1–2.2 through three lemmas. Hereafter, $\|\cdot\|$ refers to the Euclidean norm, $\|g - \tilde{g}\|_\infty$ stands for the supremum norm, and $U(\cdot)$ and $u(\cdot)$ denote the uniform boundedness of a matrix and a scalar, respectively. Without loss of generality, we always assume that the number of observations in one of $N + 1$ intervals is an multiple of the period T and the order of PAR is less than the period (i.e. $p < T$). Although the assumption is not necessarily true, it will greatly simplify the notation without changing the asymptotic properties of a statistic under consideration.

We decompose $\hat{\boldsymbol{g}}$ in (4) into three terms:

$$\hat{\boldsymbol{g}} = \tilde{\boldsymbol{g}} + \tilde{\boldsymbol{x}} + \boldsymbol{P_B}\left(\boldsymbol{D\beta} - \boldsymbol{D}\hat{\boldsymbol{\beta}}\right),$$

where $\tilde{\boldsymbol{g}} = \boldsymbol{P_B g}$ and $\tilde{\boldsymbol{x}} = \boldsymbol{P_B x}$ are the projections of $\boldsymbol{g}$ and $\boldsymbol{x}$, respectively. Furthermore, we decompose the residuals $\hat{\boldsymbol{x}}$ in (5) into three components as follows:

$$\text{(A.1)} \qquad \hat{\boldsymbol{x}} = (\boldsymbol{g} - \tilde{\boldsymbol{g}}) + (\boldsymbol{x} - \tilde{\boldsymbol{x}}) + (\boldsymbol{I} - \boldsymbol{P_B})\,\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right).$$

We will consider the asymptotics of the three terms in (A.1).

Under Assumptions 1–2, according to Theorem 5.1 of [10],

$$\text{(A.2)} \qquad \|g - \tilde{g}\|_\infty = \sup_{u\in[0,1]} |g(u) - \tilde{g}(u)| = O(h^m).$$

According to [19],

$$\text{(A.3)} \qquad \frac{1}{n}(\boldsymbol{g}-\tilde{\boldsymbol{g}})'(\boldsymbol{g}-\tilde{\boldsymbol{g}}) = O(n^{-2m/(2m+1)}),$$

$$\text{(A.4)} \qquad \frac{1}{n}(\boldsymbol{g}-\tilde{\boldsymbol{g}})'\boldsymbol{x} = O_p(n^{-(4m+1)/(4m+2)}),$$

$$\text{(A.5)} \qquad \frac{1}{n}(\boldsymbol{g}-\tilde{\boldsymbol{g}})'\tilde{\boldsymbol{x}} = O_p(n^{-2m/(2m+1)}).$$

Lemma A.1. *Under Assumptions 1–4,* $\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{D}/n_T = \boldsymbol{A}_0+o(1)$, *where the* (i,j)*-th entry of* $(T-1)\times(T-1)$ *matrix* $\boldsymbol{A}_0$ *is defined as* $1-1/T$ *if* $i=j$ *and* $-1/T$ *otherwise.*

Proof. In this case where $m=1$, simple algebra yields

$$\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B}) = (\boldsymbol{A}_0,\cdots,\boldsymbol{A}_0) + u\left\{(nh)^{-1}\right\},$$

and thus

$$\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{D}/n_T = \boldsymbol{A}_0 + o(1).$$

The proof is complete.

Lemma A.2. *Under Assumptions 1–4,* $\|E(\hat{\boldsymbol{\beta}})-\boldsymbol{\beta}\| = O(n^{-1})$.

Proof. Notice that

$$\text{(A.6)} \qquad \mathrm{E}(\hat{\boldsymbol{\beta}})-\boldsymbol{\beta} = \{\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{D}\}^{-1}\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{g}.$$

According to Lemma A.1,

$$\begin{aligned}\mathrm{E}(\hat{\boldsymbol{\beta}})-\boldsymbol{\beta} &= \left\{\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{D}\right\}^{-1}\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{g} \\ &= \frac{1}{n_T}(\boldsymbol{A},\cdots,\boldsymbol{A})\boldsymbol{g} + o(n_T^{-1}),\end{aligned}$$

where the $(T-1)\times T$ matrix $\boldsymbol{A}=(\boldsymbol{I},-\boldsymbol{1})$ with $\boldsymbol{I}$ being the identity matrix and $-\boldsymbol{1}=(-1,-1,\cdots,-1)'$. According to Assumption 1, we have

$$\|\mathrm{E}(\hat{\boldsymbol{\beta}})-\boldsymbol{\beta}\| = O(n^{-1}).$$

The proof is complete.

Proof of Theorem 2.1. Let $\boldsymbol{x}_i = (x_{iT+1},\cdots,x_{iT+T})'$. Then $\{\boldsymbol{x}_i\}$ is a multivariate stationary time series with $\mathrm{E}(\boldsymbol{x}_i)=\boldsymbol{0}$ and $\mathrm{Cov}(\boldsymbol{x}_i,\boldsymbol{x}_{i-k})=\boldsymbol{W}_k$ $(k=0,\pm1,\cdots)$, where the $T\times T$ matrix $\boldsymbol{W}_k = (\gamma_i(kT+i-j))_{i,j=1,\cdots,T}$. For any T dimension vector $\boldsymbol{a}$, define $u_i=\boldsymbol{a}'\boldsymbol{x}_i$. Then $\{u_i\}$ is a stationary time series with $\mathrm{E}(u_i)=0$ and $\gamma_u(k) = \mathrm{Cov}(u_i,u_{i-k}) = \boldsymbol{a}'\boldsymbol{W}_k\boldsymbol{a}$. Since $\{x_{iT+\nu}, 1\le\nu\le T\}_{i=0}^{n_T-1}$ is causal, $\{u_i\}$ satisfies Theorem 7.1.2 of [2], which implies

$$\sqrt{n_T}\left(\frac{1}{n_T}\sum_{i=0}^{n_T-1}\boldsymbol{a}'\boldsymbol{x}_i\right) \xrightarrow{D} N(0,\boldsymbol{a}'\boldsymbol{W}\boldsymbol{a}),$$

where

$$\boldsymbol{W} = \sum_{k=-\infty}^{\infty}\boldsymbol{W}_k = \left(\sum_{k=-\infty}^{\infty}\gamma_i(kT+i-j)\right)_{i,j=1,\cdots,T}.$$

Since $\boldsymbol{a}$ is arbitrary, from the Cramér-Wold device,

$$\sqrt{n_T}\left(\frac{1}{n_T}\sum_{i=0}^{n_T-1}\boldsymbol{x}_i\right) \xrightarrow{D} N(0,\boldsymbol{W}).$$

Define a $(T-1)\times T$ matrix $\boldsymbol{A}_1 = (\boldsymbol{A}_0, \boldsymbol{1}/T)$. Simple algebra yields the (i,j)-th entry of $\boldsymbol{A}_0^{-1}$ is 2 if $i=j$ and 1 otherwise, and $\boldsymbol{A}_0^{-1}\boldsymbol{A}_1=\boldsymbol{A}$ defined in Lemma A.2. Notice that

$$\begin{aligned}&\sqrt{n_T}\left\{\hat{\boldsymbol{\beta}}-\mathrm{E}\left(\hat{\boldsymbol{\beta}}\right)\right\} \\ =\;& \sqrt{n_T}\left\{\frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{D}\right\}^{-1}\frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I}-\boldsymbol{P_B})\boldsymbol{x} \\ =\;& \frac{1}{\sqrt{n_T}}\boldsymbol{A}_0^{-1}\sum_{i=0}^{n_T-1}\boldsymbol{A}_1\boldsymbol{x}_i + o_p\left(n_T^{-1/2}\right) \\ =\;& \sqrt{n_T}\left(\frac{1}{n_T}\boldsymbol{A}\sum_{i=0}^{n_T-1}\boldsymbol{x}_i\right) + o_p\left(n_T^{-1/2}\right).\end{aligned}$$

Hence,

$$\text{(A.7)} \qquad \sqrt{n_T}\left\{\hat{\boldsymbol{\beta}}-\mathrm{E}\left(\hat{\boldsymbol{\beta}}\right)\right\} \xrightarrow{D} N(\boldsymbol{0},\boldsymbol{V}),$$

where the $(T-1)\times(T-1)$ matrix $\boldsymbol{V}=\boldsymbol{A}\boldsymbol{W}\boldsymbol{A}'$. In particular, the (i,j)-th entry of $\boldsymbol{V}$ is

$$\begin{aligned}(\boldsymbol{V})_{i,j} = \sum_{k=-\infty}^{\infty}&\{\gamma_i(kT+i-j)-\gamma_i(kT+T-j) \\ &-\gamma_i(kT+i-T)+\gamma_i(kT)\}.\end{aligned}$$

According to Lemma A.2, $\mathrm{E}(\hat{\boldsymbol{\beta}})-\boldsymbol{\beta} = o_p(n^{-1/2})$. Slusky's Theorem together with (A.7) implies (8). The proof is complete.

Notice that the difference of the Yule-Walker estimators $\hat{\boldsymbol{\phi}}$ and $\tilde{\boldsymbol{\phi}}$ is

$$\hat{\boldsymbol{\phi}}-\tilde{\boldsymbol{\phi}} = \hat{\boldsymbol{\Gamma}}^{-1}(\hat{\boldsymbol{\gamma}}-\tilde{\boldsymbol{\gamma}}) + \hat{\boldsymbol{\Gamma}}^{-1}(\tilde{\boldsymbol{\Gamma}}-\hat{\boldsymbol{\Gamma}})\tilde{\boldsymbol{\Gamma}}^{-1}\tilde{\boldsymbol{\gamma}}.$$

The proof of Theorem 2.2 is complete if we can show that $\hat{\boldsymbol{\phi}}-\tilde{\boldsymbol{\phi}} = o_p(n^{-1/2})$ or equivalently

$$\text{(A.8)} \qquad \hat{\gamma}_\nu(k)-\tilde{\gamma}_\nu(k) = o_p(n^{-1/2}).$$

We will show (A.8) by Lemma A.3 below. Define the following vectors

$$\begin{aligned}\boldsymbol{x}_\nu &= (x_\nu, x_{T+\nu},\ldots,x_{(n_T-1)T+\nu})', \\ \boldsymbol{g}_\nu &= (g(u_\nu), g(u_{T+\nu}),\ldots,g(u_{(n_T-1)T+\nu}))'.\end{aligned}$$

This notation sometimes is applied to $\hat{\boldsymbol{x}}_\nu$, $\tilde{\boldsymbol{x}}_\nu$, $\tilde{\boldsymbol{g}}_\nu$ and so on. Notice that the vector $\boldsymbol{x}_\nu$ includes almost all but a finite number of the observations, and the vectors $\boldsymbol{x}_{iT+\nu}$ and $(x_\nu, x_{T+\nu},\ldots,x_{(n_T-1)T+\nu})'$ are equivalent in the sense of asymptotics for any fixed i. We will show (A.8) by the following lemma.

Lemma A.3. *Under Assumptions 1–4, for any fixed* ν $\hat{\gamma}_\nu(k) - \tilde{\gamma}_\nu(k) = O_p(h^2)$ *and* $\hat{\gamma}_\nu(k) \xrightarrow{P} \gamma_\nu(k)$.

Proof. Without loss of generality, we assume that $\nu - k \geq 0$. According to (A.1),

$$\hat{\gamma}_\nu(k) - \tilde{\gamma}_\nu(k) = \frac{1}{n_T}\hat{\boldsymbol{x}}'_\nu \hat{\boldsymbol{x}}_{\nu-k} - \frac{1}{n_T}\boldsymbol{x}'_\nu \boldsymbol{x}_{\nu-k}$$

$$\begin{aligned}
= \;& \frac{1}{n_T}(\boldsymbol{g}_\nu - \tilde{\boldsymbol{g}}_\nu)' \boldsymbol{x}_{\nu-k} - \frac{1}{n_T}(\boldsymbol{g}_\nu - \tilde{\boldsymbol{g}}_\nu)' \tilde{\boldsymbol{x}}_{\nu-k} \\
& + \frac{1}{n_T}\tilde{\boldsymbol{x}}'_\nu \tilde{\boldsymbol{x}}_{\nu-k} - \frac{1}{n_T}\boldsymbol{x}'_{\nu-k}\tilde{\boldsymbol{x}}_\nu - \frac{1}{n_T}\tilde{\boldsymbol{x}}'_{\nu-k}\boldsymbol{x}_\nu \\
& + \frac{1}{n_T}(\boldsymbol{g}_\nu - \tilde{\boldsymbol{g}}_\nu)'(\boldsymbol{g}_{\nu-k} - \tilde{\boldsymbol{g}}_{\nu-k}) \\
& + \frac{1}{n_T}(\boldsymbol{g}_{\nu-k} - \tilde{\boldsymbol{g}}_{\nu-k})' \boldsymbol{x}_\nu - \frac{1}{n_T}(\boldsymbol{g}_{\nu-k} - \tilde{\boldsymbol{g}}_{\nu-k})' \tilde{\boldsymbol{x}}_\nu \\
& + \frac{1}{n_T}(\boldsymbol{g}_\nu - \tilde{\boldsymbol{g}}_\nu)' \left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}_{\nu-k} \\
& + \frac{1}{n_T}(\boldsymbol{g}_{\nu-k} - \tilde{\boldsymbol{g}}_{\nu-k})' \left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}_{\nu} \\
& + \frac{1}{n_T}(\boldsymbol{x}_\nu - \tilde{\boldsymbol{x}}_\nu)' \left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}_{\nu-k} \\
& + \frac{1}{n_T}(\boldsymbol{x}_{\nu-k} - \tilde{\boldsymbol{x}}_{\nu-k})' \left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}_{\nu} \\
& + \frac{1}{n_T}\boldsymbol{S}'_\nu \boldsymbol{S}_{\nu-k},
\end{aligned} \tag{A.9}$$

where $\boldsymbol{S}_\nu = \left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}_\nu$. We will consider the orders of the terms above. First

$$\begin{aligned}
& \frac{1}{n_T}\left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\right\}'(\boldsymbol{x} - \tilde{\boldsymbol{x}}) \\
= \;& \frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{x} \\
= \;& \frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{x} = \frac{1}{n_T}\sum_{i=0}^{n_T-1}\boldsymbol{A}_1\boldsymbol{x}_i + u_p(n_T^{-1}).
\end{aligned}$$

Following a similar discussion to the derivation of (5), we obtain $\sum_{i=0}^{n_T-1}\boldsymbol{A}_1\boldsymbol{x}_i/n_T = O_p(n^{-1/2})$. Thus

$$\frac{1}{n_T}\left\|\left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\right\}'(\boldsymbol{x} - \tilde{\boldsymbol{x}})\right\| = O_p(n^{-1/2}). \tag{A.10}$$

Next, we note $\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B}) = u(1)$, and thus the (i,j)-th entry of $\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})(\boldsymbol{g} - \tilde{\boldsymbol{g}})$ is $O(n_T)\|g - \tilde{g}\|_\infty$. Therefore,

$$\begin{aligned}
& \frac{1}{n_T}\left\|\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})(\boldsymbol{g}_\nu - \tilde{\boldsymbol{g}}_\nu)\right\| \\
\leq \;& \frac{1}{n_T}O(n_T)\|g - \tilde{g}\|_\infty = O(n^{-m/(2m+1)}).
\end{aligned} \tag{A.11}$$

According to Theorem 2.1, $\left\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right\| = O_p(n^{-1/2})$, which, together with (A.10) and (A.11), implies

$$\frac{1}{n_T}(\boldsymbol{g}_\nu - \tilde{\boldsymbol{g}}_\nu)'\left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}_{\nu-k} = O_p(n^{-5/6}),$$

$$\frac{1}{n_T}(\boldsymbol{x}_\nu - \tilde{\boldsymbol{x}}_\nu)'\left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}_{\nu-k} = O_p(n^{-1}).$$

Finally,

$$\begin{aligned}
& \left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\}'\left\{(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right\} \\
= \;& \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)'\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right) \\
= \;& \boldsymbol{g}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left\{\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\right\}^{-1}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{g} \\
& + \boldsymbol{x}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\left\{\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\right\}^{-1}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{x} \\
= \;& \left\|\left\{\frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\right\}^{-1}\left\{\frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{g}\right\}\right\|^2 \\
& + \left\|\left\{\frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{D}\right\}^{-1}\left\{\frac{1}{n_T}\boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{P_B})\boldsymbol{x}\right\}\right\|^2 \\
= \;& O(1) + O_p\left(n_T^{-1}\right).
\end{aligned}$$

Thus

$$\frac{1}{n_T}\boldsymbol{S}'_\nu \boldsymbol{S}_{\nu-k} = O_p(1). \tag{A.12}$$

According to (A.3), (A.4), (A.5), (A.10), (A.11), (A.12), and Theorem 2.1, the order of the dominant terms in (A.9) is $O_p(n^{-2/3})$, which implies $\hat{\gamma}_\nu(k) - \tilde{\gamma}_\nu(k) = O_p(h^2) = O_p(n^{-2/3})$ and $\hat{\gamma}_\nu(k) \xrightarrow{P} \gamma_\nu(k)$ for any fixed $1 \leq \nu \leq T$. The proof is complete.

Received 1 May 2015

REFERENCES

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

[2] Brockwell, P. J. and Davis, R. A. (1991) *The Analysis of Time Series: Theory and Methods*, 2nd ed, Springer, New York. MR1093459

[3] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994) *Time Series Analysis. Forecasting and Control*, 3rd ed, Prentice-Hall, Englewood Cliffs, NJ. MR1312604

[4] Davis, R. A. and Dunsmuir, W. T. (1997) Least absolute deviation estimation for regression with ARMA errors. *Journal of Theoretical Probability* 10, 481–497. MR1455154

[5] Engle, R., Granger, C., Rice, J., and Weiss, A. (1986) Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association* 81, 310–320.

[6] Fan, J. and Yao, Q. W. (2003) *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer, New York. MR1964455

[7] Gao, J. T. (1995) Asymptotic theory for partly linear models. *Communications in Statistics: Theory Methods* 24, 1985–2009. MR1345231

[8] Heckman, H. (1986) Spline smoothing in partly linear model. *Journal of Royal Statistical Society Series B* 48, 244–248. MR0868002

[9] Hipel, W. K. and McLeod, I. A. (1994) *Time Series Modelling of Water Resources and Environmental Systems*, Elsevier.

[10] Huang, J. Z. (2003) Local asymptotics for polynomial spline regression. *Annals of Statistics* 31, 1600–1635. MR2012827

[11] Kohn, R., Ansley, C. F., and Wong, C. M. (1992) Nonparametric spline regression with autoregressive moving average errors. *Biometrika* 79, 335–346. MR1185135

[12] Lund, R. B., Shao, Q., and Basawa, I. V. (2006) Parsimonious periodic time series modeling. *Australian & New Zealand Journal of Statistics* 48, 33–47. MR2234777

[13] McLeod, A. I. and Balcilar, M. (2011) pear: Package for Periodic Autoregression Analysis, R package. http://cran.r-project.org/web/packages/pear/.

[14] Pagano, M. (1978) On periodic and multiple autoregressions. *The Annals of Statistics* 6, 1310–1317. MR0523765

[15] Qiu, D, Shao, Q., and Yang, L. (2013) Efficient inference for autoregressive coefficient in the presence of trend. *Journal of Multivariate Analysis* 114, 40–53. MR2993872

[16] R Core Team (2014) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[17] Shao, Q. (2009) Seasonality analysis of time series in partial linear models. *Journal of Nonparametric Statistics* 21, 827–837. MR2572585

[18] Shao, Q. and Yang, L. (2012) Polynomial spline confidence bands for time series trend. *Journal of Statistical Planning and Inference* 142, 1678–1689. MR2903380

[19] Shao, Q. and Yang, L. (2011) Autoregressive coefficient estimation in nonparametric analysis. *Journal of Time Series Analysis* 32, 587–597. MR2846558

[20] Shumway, R. H. and Stoffer, D. S. (2011) *Time Series Analysis and Its Applications with R Examples*, 3rd ed, Springer. MR2721825

[21] Speckman, P. (1988) Kernel smoothing in partial linear models. *Journal of the Royal Statistical Society. Series B (Methodological)* 50, 413–436. MR0970977

[22] Tang, L. and Shao, Q. (2014) Efficient estimation for periodic autoregressive coefficients via residuals. *Journal of Time Series Analysis* 35, 378–389. MR3226451

[23] Tsay, R. S. (1984) Regression models with time series errors. *Journal of the American Statistical Association* 79, 118–124. MR0742860

[24] You, J., and Chen, G. (2007) Semiparametric generalized least squares estimation in partially linear regression models with correlated errors. *Journal of Statistical Planning and Inference* 137, 117–132. MR2292845

[25] Wang, Y. (1998) Smoothing spline models with correlated random errors. *Journal of the American Statistical Association* 93, 341–348.

[26] Wang, L. and Yang, L. (2007) Spline-backfitted kernel smoothing of nonlinear additive autoregression model. *Annals of Statistics* 35, 2474–2503. MR2382655

[27] Xue, L. and Yang, L. (2006) Additive coefficient modelling via polynomial spline. *Statistica Sinica* 16, 1423–1446. MR2327498

Lin Tang
Department of Mathematics and Statistics
The University of Toledo
Toledo, OH 43606
USA
E-mail address: lin.tang@rockets.utoledo.edu

Qin Shao
Department of Mathematics and Statistics
The University of Toledo
Toledo, OH 43606
USA
E-mail address: qin.shao@utoledo.edu