Comput Stat (2015) 30:719–744
DOI 10.1007/s00180-015-0574-x

ORIGINAL PAPER

Functional coefficient seasonal time series models with an application of Hawaii tourism data

Xialu Liu[1] · **Zongwu Cai**[2,3] · **Rong Chen**[1]

Received: 13 February 2014 / Accepted: 18 March 2015 / Published online: 5 April 2015

Abstract In this article, motivated by an analysis of the monthly number of tourists visiting Hawaii, we propose a new class of nonparametric seasonal time series models under the framework of the functional coefficient model. The coefficients change over time and consist of the trend and seasonal components to characterize seasonality. A local linear approach is developed to estimate the nonparametric trend and seasonal effect functions. The consistency of the proposed estimators is obtained without specifying the error distribution and the asymptotic normality of the proposed estimators is established under the α-mixing conditions. A consistent estimator of the asymptotic variance is also provided. The proposed methodologies are illustrated by two simulated examples and the model is applied to characterizing the seasonality of the monthly number of tourists visiting Hawaii.

The authors are grateful to Professors D. Findley, Tom Fomby and R. H. Shumway for their helpful and insightful comments and suggestions. Cai's research was supported, in part, by the National Nature Science Foundation of China Grants #71131008 (Key Project) and #70871003. Chen's research was supported in part by the National Science Foundation Grants DMS 1209085, DMS 0905763 and DMS 0915139.

✉ Rong Chen
rongchen@stat.rutgers.edu

Xialu Liu
xialu.liu@rutgers.edu

Zongwu Cai
caiz@ku.edu

1 Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA

2 The Charles Osward Professor of Econometrics, Department of Economics, University of Kansas, Lawrence, KS 66045, USA

3 Xiamen University, Xiamen, China

Keywords Fixed effects · Functional-coefficient models · Local linear fitting · Nonparametric smoothing · Seasonal variation

1 Introduction

Seasonal time series are commonly observed in various applications, including economic and business data, meteorological data and environmental data as well as other fields. There is a vast literature on seasonal time series analysis, ranging from stochastic seasonality models such as the seasonal ARIMA models (Box et al. 1994; Shumway and Stoffer 2000; Pena et al. 2001), the deterministic seasonal models such as the linear or polynomial additive or multiplicative seasonal component models (Shumway 1988; Brockwell and Davis 1991; Franses 1996, 1998). The books by Hylleberg (1992), Franses (1996, 1998), and Ghysels and Osborn (2001) provide a comprehensive review on the traditional seasonal time series analysis methods. Most of these methods are linear (or polynomial) and parametric in nature. However, it has been documented that time series are often nonlinear (Tong 1990; Tjøstheim 1994; Hylleberg 1992; Franses 1996, 1998) and often there is not enough information to determine a suitable parametric form for the nonlinear structure. Härdle et al. (2004) discussed and reviewed many the popular statistical nonparametric and semiparametric methods. These methods have been widely utilized for nonlinear time series analysis (e.g. Härdle and Vieu 1992; Tjøstheim 1994; Härdle et al. 1997; Fan and Yao 2003). To avoid the *curse of dimensionality*, nonparametric time series models with special structure have been proposed and applied to real applications, including the nonlinear additive ARX models (Chen and Tsay 1993b), the functional coefficient AR models (Chen and Tsay 1993a; Xia and Li 1999a; Cai 2007; Cai et al. 2000, 2009), the single index coefficient AR models (Härdle et al. 1993; Xia and Li 1999b; Fan and Yao 2003; Fan et al. 2003; Lu et al. 2007) and others. There was no systematic research done on nonparametric approaches to seasonal time series models, until Burman and Shumway (1998) proposed a nonparametric/semiparametric approach to seasonal time series, which opened the door in this area.

In this paper, to characterize the seasonality of the monthly number of tourists visiting Hawaii, we propose a nonparametric seasonal time series model with a functional coefficient structure. Different from a linear autoregressive seasonal model with possible regression terms, the coefficients in the proposed model change over time and consist of the trend and seasonal components to characterize the seasonality. This class of models includes, as its special cases, the standard additive trend and seasonal component models as well as other seasonal time series models. We use the local linear approach to estimate the trend and seasonal effect functions nonparametrically and study the asymptotic properties of the proposed estimators. The detailed analysis of this data set is reported in Sect. 4.

The rest of the paper is organized as follows. In Sect. 2 we first introduce the model and its motivation, and then we present the estimation procedure, followed by the asymptotic properties of the proposed estimators. Section 3 presents a Monte Carlo simulation study to illustrate the finite sample performance of the proposed estimation procedures and in Sect. 4, the proposed model and its modeling procedures are apply to an analysis of the monthly number of tourists visiting Hawaii. Finally, the mathematical Proof of the Theorem is given in the "Appendix".

2 Functional coefficient seasonal time series models

2.1 The model

Denote a seasonal time series as

$$y_{t1}, \ldots, y_{td}, \quad t = 1, 2, \ldots, n, \tag{1}$$

where d is the number of seasons within a period and n is the number of periods. We assume that there exist p other time series $\{x_{ktj}\}$, $k = 1, \ldots, p$, and $j = 1, \ldots, d$ that are related to the time series y_{tj}, and indexed according to y_{tj}. Those time series can be the lagged series of y_{tj} (in an AR fashion), or some exogenous variables.

The proposed *functional-coefficient seasonal time series model* assumes the form as

$$y_{tj} = \sum_{k=1}^{p} [\alpha_k(t) + \beta_{kj}(t)]\, x_{ktj} + e_{tj}, \tag{2}$$

where $\{\alpha_k(\cdot)\}$ are the trend functions for the coefficients, and $\{\beta_{jk}(\cdot)\}$ are the seasonal effect functions in the coefficient functions, satisfying constraints for the identification,

$$\sum_{j=1}^{d} \beta_{kj}(t) = 0, \qquad \text{for each } 1 \le k \le p \text{ and all } t,$$

and the error term $\{e_{tj}\}$ is stationary and satisfies $E(e_{tj} \mid \mathbf{X}_{tj}) = 0$ with $\mathbf{X}_{tj} = (x_{1tj}, \ldots, x_{ptj})^T$.

Remark 1 There is another way to denote seasonal time series with only one subscript as

$$y_1, \ldots, y_m, \ldots, y_T, \quad m = 1, 2, \ldots, T. \tag{3}$$

Both (1) and (3) are used in this paper exchangeably, identified by the number of subscripts. Time series denoted by the two different indexed methods satisfies the formula as $y_m = y_{tj}$, where $m = d(t-1) + j$ for $1 \le t \le n$ and $1 \le j \le d$.

Model (2), where coefficients combine of nonlinear trend and seasonal effect changing over time, is a generalization of the functional-coefficient time series model, a popular nonlinear time series model in the time series literature (Chen and Tsay 1993a; Xia and Li 1999a; Cai et al. 2000; Cai and Tiwari 2000), and the varying-coefficient model (Hastie and Tibshirani 1993; Yang et al. 2006) for i.i.d. samples.

This model is also motivated by the standard additive time trend and seasonal component model as

$$y_{tj} = T_t + S_{tj} + e_{tj}; \tag{4}$$

see Cleveland et al. (1990) and Cai and Chen (2006) where T_t is the common trend same to different seasons within a period, and S_{tj} is the seasonal effect, satisfying $\sum_{j=1}^{d} S_{tj} = 0$. A standard parametric model assumes a parametric function for the common trend T_t, such as linear or polynomial functions. The seasonal effects are usually assumed to be the same for different periods; that is, $S_{tj} = S_j$ for $j = 1, \ldots, d$ and all t. Note that if $p = 1$ and $x_{1tj} = 1$ for all t and j, then model (2) becomes

$$y_{tj} = \alpha(t) + \beta_j(t) + e_{tj}, \tag{5}$$

where $\{\beta_j(t)\}$ satisfy the condition $\sum_{j=1}^{d} \beta_j(t) = 0$; see Cai and Chen (2006) for details. This is the exact same as (4). Here, we assume nonparametric forms for both trend and seasonal component. If we further assume that $\beta_j(t) = \gamma_j \beta(t)$, then we obtained the model proposed by Burman and Shumway (1998), where $\{\gamma_j\}$ are seasonal factors. Hence, the overall seasonal effect changes over periods in accordance with the modulating function $\beta(t)$. Implicitly, this model assumes that the seasonal effect curves have the same shape (up to a multiplicative constant) for all seasons.

The AR model with trend and seasonal component is also commonly used in modeling seasonal time series (e.g., Hylleberg 1992; Franses 1996, 1998; Ghysels and Osborn 2001),

$$y_{tj} = T_t + S_{tj} + \phi y_{tj-1} + e_{tj}. \tag{6}$$

Our model allows both AR terms and exogenous variables to enter the model in a linear fashion. The AR coefficients and the coefficients of the exogenous variables are commonly assumed to be constant over different periods. However, for seasonal time series models, it is difficult to justify that the relationships between y_t and its lag variables and exogenous variables are the same for different periods. Allowing different functions for different periods (hence seasonality) has an ability to enhance the model to adopt the nature of the underlying time series and to capture the seasonality better.

In addition, if $p = 1$ and x_t is the lag d variable of y_t, say $x_m = y_{m-d}$, or $x_{tj} = y_{(t-1)j}$, then this model assumes a pure seasonal AR model with d different series, each with seasonality of 1 as

$$y_{tj} = (\alpha(t) + \beta_j(t)) y_{(t-1)j} + e_{tj}, \quad j = 1, \ldots, d, \tag{7}$$

where the coefficients change over time, with $\alpha(t)$ being the common trend and $\beta_j(t)$ being the seasonal effect, special to each season j in the period. Both trend and seasonal effect functions are nonparametric. An extreme case is that $\beta_j(t) = 0$ and $\alpha(t) = \alpha$, a constant. In this case, Eq. (7) becomes

$$y_m - \alpha y_{m-d} = e_m,$$

which is a pure seasonal AR model. Therefore, with certain combinations of the variables x_m and the coefficient functions, the proposed model in (2) is flexible enough to cover many existing seasonal models.

2.2 Estimation procedure

For technical reasons, we change the time unit in the coefficient functions to $s_t = t/n$. Then, we can express (2) in a matrix notation,

$$\mathbf{Y}_t = \mathcal{X}_t\, \boldsymbol{\theta}(s_t) + \mathbf{e}_t,$$

where

$$\mathbf{Y}_t = \begin{pmatrix} y_{t1} \\ \vdots \\ y_{td} \end{pmatrix}, \quad \mathcal{X}_t = \begin{pmatrix} \mathbf{X}_{t1}^T & \mathbf{X}_{t1}^T & \mathbf{0} & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{X}_{t,d-1}^T & \mathbf{0} & \mathbf{0} & \mathbf{X}_{t,d-1}^T \\ \mathbf{X}_{td}^T & -\mathbf{X}_{td}^T & \dots & -\mathbf{X}_{td}^T \end{pmatrix}, \quad \mathbf{e}_t = \begin{pmatrix} e_{t1} \\ \vdots \\ e_{td} \end{pmatrix},$$

$$\text{and} \quad \boldsymbol{\theta}(s_t) = \begin{pmatrix} \boldsymbol{\alpha}(s_t) \\ \boldsymbol{\beta}_1(s_t) \\ \vdots \\ \boldsymbol{\beta}_{d-1}(s_t) \end{pmatrix},$$

with $\boldsymbol{\alpha}(s_t) = (\alpha_1(s_t)\dots\alpha_p(s_t))^T$ and $\boldsymbol{\beta}_j(s_t) = (\beta_{1j}(s_t)\dots\beta_{pj}(s_t))^T$. Again, the error term $\{\mathbf{e}_t\}$ is assumed to be stationary with $E(\mathbf{e}_t) = \mathbf{0}$ and $\mathrm{var}(\mathbf{e}_t) = \boldsymbol{\Sigma}_e$.

For estimating $\alpha(\cdot)$ and $\{\beta_j(\cdot)\}$, a local linear method is employed, although a general local polynomial method is also applicable. Local linear (polynomial) methods have been widely used in nonparametric regression due to their attractive mathematical efficiency, bias reduction and adaptation of edge effects (see Fan and Gijbels 1996). We assume throughout that the trend functions $\{\alpha_k(\cdot)\}$ and the seasonal effect functions $\{\beta_{kj}(\cdot)\}$ have a continuous second derivative. Then, based on the local linear fitting scheme of Fan and Gijbels (1996), the locally weighted least squares is given by

$$\sum_{t=1}^{n} [\mathbf{Y}_t - \mathcal{X}_t\, \boldsymbol{\theta}_0 - (s_t - s)\, \mathcal{X}_t\, \boldsymbol{\theta}_1]^T\, [\mathbf{Y}_t - \mathcal{X}_t\, \boldsymbol{\theta}_0 - (s_t - s)\, \mathcal{X}_t\, \boldsymbol{\theta}_1]\;\; K_h(s_t - s), \quad (8)$$

where $K_h(u) = K(u/h)/h$, $K(\cdot)$ is a kernel function and h is the bandwidth satisfying $h \to 0$ and $n\,h \to \infty$ as $n \to \infty$. Let $\widehat{\boldsymbol{\theta}}_0$ and $\widehat{\boldsymbol{\theta}}_1$ be the minimizer of (8). Then,

$$\begin{pmatrix} \widehat{\boldsymbol{\theta}}_0 \\ \widehat{\boldsymbol{\theta}}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{G}_0 & \mathbf{G}_1 \\ \mathbf{G}_1 & \mathbf{G}_2 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{M}_0 \\ \mathbf{M}_1 \end{pmatrix}, \quad (9)$$

where

$$\mathbf{G}_k = \frac{1}{n} \sum_{t=1}^{n} \mathcal{X}_t^T\; \mathcal{X}_t\; (s_t - s)^k\; K_h(s_t - s) \;\text{ and }\; \mathbf{M}_k = \frac{1}{n} \sum_{t=1}^{n} \mathcal{X}_t^T\; \mathbf{Y}_t\; (s_t - s)^k\; K_h(s_t - s).$$

Therefore, the local linear estimates of $\boldsymbol{\theta}(s)$ and $\boldsymbol{\theta}'(s)$ (the first order derivative of $\boldsymbol{\theta}(s)$) are $\widehat{\boldsymbol{\theta}}(s) = \widehat{\boldsymbol{\theta}}_0$ and $\widehat{\boldsymbol{\theta}}'(s) = \widehat{\boldsymbol{\theta}}_1$, respectively.

Remark 2 Note that many other nonparametric smoothing methods can be used here. The locally weighted least square method is just one of the choices. There is a vast amount of literature in theory and empirical study on the comparison of different methods (see Fan and Gijbels 1996).

Remark 3 The restriction to the locally weighted least square method suggests that the normality is at least being considered as a baseline. However, when the non-normality is clearly present, a robust approach would be considered. Cai and Ould-Said (2001) considered this aspect in nonparametric regression estimation for time series.

Remark 4 The bandwidth selection is always one of the most important parts of any nonparametric procedure. There are several bandwidth selectors in the literature, including the leave-one-out cross validation of Härdle and Marron (1985), the generalized cross-validation of Wahba (1977), the plug-in method of Jones et al. (1996), and the empirical bias method of Ruppert (1997), among others. They all can be used here. A comparison of different procedures can be found in Jones et al. (1996). In this article we use a procedure proposed in Fan et al. (2003), which combines the generalized cross-validation and the empirical bias method.

Remark 5 In the above estimation procedure, one bandwidth is used for all functions. It is possible to use different bandwidths for different seasons by a more computational intensive two-step method (see Cai 2002). We can also incorporate the covariance structure of $\mathbf{e}_t$ in the estimation.

Remark 6 Since data are observed in time order as in Burman and Shumway (1998), we assume that $s_t = t/n$ for simplicity although the theoretical results developed later still hold for non-equally spaced design points.

2.3 Asymptotic properties

The estimation method described above can accommodate both fixed and random designs of the time index (at which the time series are observed). In this paper we focus on fixed-design since time series are commonly observed over a fixed time interval and we now use $\mathbf{e}_{nt}$ to denote $\mathbf{e}_t$. In such a case, we assume that, for each n, $\{\mathbf{e}_{n1}, \ldots, \mathbf{e}_{nn}\}$ have the same joint distribution as $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n\}$, where $\boldsymbol{\xi}_t$, $t = \ldots, -1, 0, 1, \ldots$, is a strictly stationary time series defined on a probability space $(\Omega, \mathcal{A}, P)$ and taking values on $\Re^d$. This type of assumption is commonly used in fixed-design regression for time series contexts. Detailed discussions on this respect can be found in Roussas (1989), Roussas et al. (1992), and Tran et al. (1996) for nonparametric regression estimation for dependent data.

Traditionally, the error component in a deterministic trend and seasonal component model (4) is assumed to follow certain linear time series models such as an ARMA process. Here we consider a more general structure—the α-mixing process, which includes many linear and nonlinear time series models as special cases (see Remark 7 later). The theoretical results are derived under this assumption. For reference convenience, the mixing coefficient is defined as

$$\alpha(t) = \sup\left\{|P(A \cap B) - P(A)\,P(B)| : \; A \in \mathcal{F}_{-\infty}^{0}, \; B \in \mathcal{F}_{t}^{\infty}\right\},$$

where $\mathcal{F}_a^b$ is the σ-algebra generated by $\{\boldsymbol{\xi}_t\}_{t=a}^{b}$. If $\alpha(t) \to 0$ as $t \to \infty$, the process is called strongly mixing or α-mixing.

Remark 7 Among various mixing conditions used in the literature, α-mixing is reasonably weak and is known to be fulfilled for many linear and nonlinear time series models under some regularity conditions. Gorodetskii (1977) and Withers (1981) derived the conditions under which a linear process is α-mixing. In fact, under very mild assumptions, linear autoregressive and more generally bilinear time series models are α-mixing with mixing coefficients decaying exponentially. Auestad and Tjøstheim (1990) provided illuminating discussions on the role of α-mixing (including geometric ergodicity) for model identification in nonlinear time series analysis. Chen and Tsay (1993a) showed that the functional autoregressive process is geometrically ergodic under certain conditions. Further, Masry and Tjøstheim (1995, 1997) and Lu (1998) demonstrated that under some mild conditions, both autoregressive conditional heteroscedastic processes and nonlinear additive autoregressive models with exogenous variables, particularly popular in finance and econometrics, are stationary and α-mixing. Roussas (1989) considered linear processes without satisfying the mixing condition. Potentially our results can be extended to such cases.

To establish the asymptotic properties of $\widehat{\boldsymbol{\theta}}(s)$, denote $\mathbf{G} = E\left(\mathcal{X}^T\,\mathcal{X}\right)$, where $\mathcal{X}$ has the same distribution as $\mathcal{X}_t$ for all t. Define $\mathbf{e}_t^* = \mathcal{X}_t^T\,\mathbf{e}_t$. Then, $\{\mathbf{e}_t^*\}$ is stationary by Assumption A2 in the "Appendix". Also, define $\mathbf{R}(k-l) = \mathrm{cov}(\mathbf{e}_k^*, \mathbf{e}_l^*)$ and $\boldsymbol{\Sigma}_0 = \sum_{k=-\infty}^{\infty}\mathbf{R}(k)$. Then, $\boldsymbol{\Sigma}_0$ exists by Assumption A2 in the "Appendix". Finally, define, for $k \geq 0$, $\mu_k = \int u^k K(u)\,d\,u$ and $\nu_k = \int u^k\,K^2(u)\,d\,u$. We have the following theorem with its sketch proof given in the "Appendix".

Theorem 1 *Assume that Assumptions A1–A3 in the "Appendix" hold, then,*

$$\sqrt{n\,h}\,\left\{\widehat{\boldsymbol{\theta}}(s) - \boldsymbol{\theta}(s) - \frac{h^2}{2}\,\mu_2\,\boldsymbol{\theta}''(s) + o_p(h^2)\right\} \;\longrightarrow\; N(\mathbf{0},\;\boldsymbol{\Sigma}_\theta),$$

where $\boldsymbol{\Sigma}_\theta = \nu_0\,\mathbf{G}^{-1}\,\boldsymbol{\Sigma}_0\,\mathbf{G}^{-1}$ *and* $\boldsymbol{\theta}''(s)$ *is the second order derivative of* $\boldsymbol{\theta}(s)$.

Remark 8 As a consequence of Theorem 1, $\widehat{\boldsymbol{\theta}}(s) - \boldsymbol{\theta}(s) = O_p\left(h^2 + (n\,h)^{-1/2}\right)$, so that $\widehat{\boldsymbol{\theta}}(t)$ is a consistent estimator of $\boldsymbol{\theta}(s)$. Also, the asymptotic variance of the estimator does not depend on the grid point t. More importantly, it shows that the asymptotic variance of the estimator depends not only on the covariance structure of the seasonal effects ($\mathbf{R}(0) = \mathrm{var}(\mathbf{e}_i^*)$) but also the autocorrelations over periods ($\sum_{k=1}^{\infty}\mathbf{R}(k)$). Further, it is interesting to note that in general, $\boldsymbol{\Sigma}_\theta$ might not be diagonal. This implies that $\widehat{\alpha}(s)$ and $\widehat{\beta}_j(s)$ $(1 \leq j \leq d-1)$ may be asymptotically correlated. Finally, the asymptotic mean square error (AMSE) of $\widehat{\boldsymbol{\theta}}(s)$ is given by

$$\text{AMSE} = \frac{h^4}{4}\,\mu_2^2\;||\boldsymbol{\theta}''(s)||_2^2 + \frac{\mathrm{tr}(\boldsymbol{\Sigma}_\theta)}{n\,h},$$

which gives the optimal bandwidth,

$$h_{opt} = n^{-1/5} \left\{ \text{tr}(\boldsymbol{\Sigma}_\theta)\, \mu_2^{-2} \, \left\| \boldsymbol{\theta}''(s) \right\|_2^{-2} \right\}^{-1/5},$$

by minimizing the AMSE. Hence, the optimal convergence rate of the AMSE for $\widehat{\boldsymbol{\theta}}(t)$ is of the order of $n^{-4/5}$, as one would have expected.

Remark 9 In practice, it is desirable to have a quick and easy implementation to estimate the asymptotic variance of $\widehat{\boldsymbol{\theta}}(s)$ to construct a pointwise confidence interval. The explicit expression of the asymptotic variance in Theorem 1 provides two direct estimators. From Lemma A1 in the "Appendix", for any $0 < s < 1$, we have

$$\boldsymbol{\Sigma}_0 = \lim_{n\to\infty} \frac{h}{n} \, \text{var}\left(\sum_{t=1}^{n} \mathbf{e}_{nt}^* \, K_h(s_t - s) \right).$$

Hence a direct (naive) estimator of $\boldsymbol{\Sigma}_0$ is given by

$$\widehat{\boldsymbol{\Sigma}}_0 = \widehat{\mathbf{P}}_{n0}\, \widehat{\mathbf{P}}_{n0}^T,$$

where $\widehat{\mathbf{P}}_{n0} = (h\, n^{-1})^{1/2} \sum_{t=1}^{n} \mathcal{X}_t^T \left\{ \mathbf{Y}_t - \mathcal{X}_t\, \widehat{\boldsymbol{\theta}}(s_t) \right\} K_h(s_t - s)$. However, in the finite sample, $\widehat{\boldsymbol{\Sigma}}_\theta$ might depend on s. To overcome this shortcoming, an alternative way to construct a consistent estimation of $\boldsymbol{\Sigma}_0$ is to use the sample autocovariance type estimation methods to estimate $\{\mathbf{R}(k)\}$, such as the heteroscedasticity and autocorrelation consistent (HAC) methods; see Newey and West (1987) and Andrews (1991) for further discussions. But it seems that they might require more data points than the previous method.

3 Simulated examples

In this section, a Monte Carlo simulation study is conducted to examine the finite sample performance of the proposed procedures. Throughout this section, we use the Epanechnikov kernel, $K(u) = 0.75\,(1 - u^2)\, I(|u| \le 1)$ and the bandwidth selector mentioned in Remark 4. For simulated examples, the performance of the estimators is evaluated by the mean absolute deviation error (MADE):

$$\mathcal{E}_k = n_0^{-1} \sum_{j=1}^{n_0} \left| \widehat{\alpha}_k(v_j) - \alpha_k(v_j) \right| \quad \text{and} \quad \mathcal{E}_{kj} = n_0^{-1} \sum_{j=1}^{n_0} \left| \widehat{\beta}_{kj}(v_j) - \beta_{kj}(v_j) \right|$$

for $\alpha_k(\cdot)$ and $\beta_{kj}(\cdot)$, respectively, where $k = 1, \ldots, p$, $j = 1, \ldots, d$, and $\{v_j,\ j = 1, \ldots, n_0\}$ are the grid points from $(0,\ 1]$. When $p = 1$, the subscript k can be omitted. Simulation is repeated 500 times for each model with different sample sizes. For demonstration purposes, when showing results of a particular simulated series, we use the series with median total MADE value (sum of all MADE values) equals among the 500 MADE values. Such a sample is referred to as *a typical sample*.

Example 1 We begin with a simple additive trend and seasonal component model

$$y_{tj} = \alpha(s_t) + \beta_j(s_t) + e_{tj}, \quad t = 1, \dots, n, \quad j = 1, \dots, 4,$$

where $s_t = t/n$, $\alpha(x) = \exp(-0.7 + 3.5x)$, $\beta_1(x) = -3.1x^2 + 17.1x^4 - 28.1x^5 + 15.7x^6$, $\beta_2(x) = -0.5x^2 + 15.7x^6 - 15.2x^7$, $\beta_3(x) = -0.2 + 4.8x^2 - 7.7x^3$, and $\beta_4(x) = -\beta_1(x) - \beta_2(x) - \beta_3(x)$, for $0 < x \le 1$. Here, the error $\{e_m\}$ are generated from the following AR(1) model:

$$e_m = 0.9e_{m-1} + \varepsilon_m,$$

and ε_t is generated from $N(0,\ 0.1^2)$.

The sample sizes are $n = 50$, 100, and 300, respectively. Figure 1 gives the time plot of a typical sample with the sample size $n = 100$. Figure 2 shows the estimated $\alpha(\cdot)$ and $\{\beta_j(\cdot)\}$ (dashed lines) from the typical sample, together with their true values (solid lines), and it can be seen that estimated values are very close to the true

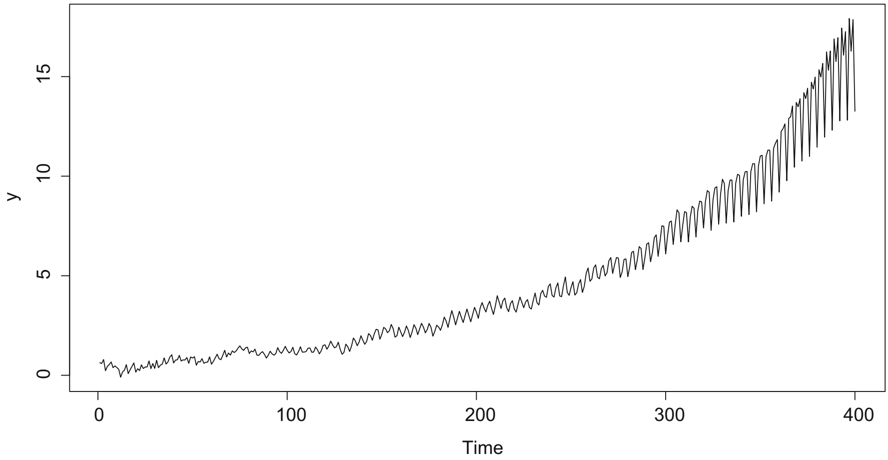

Fig. 1 Time series plot of a typical sample from Example 1 with $n = 100$

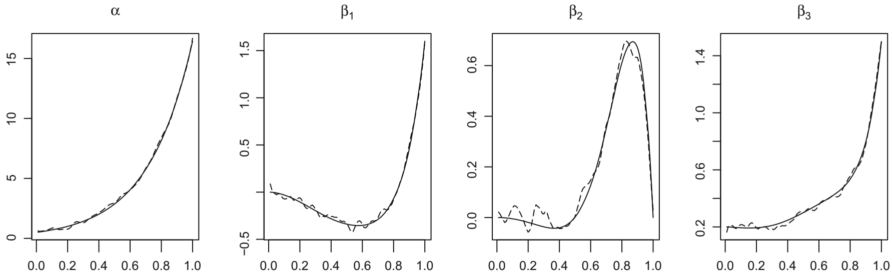

Fig. 2 Estimation results for a typical sample from Example 1 with $n = 100$. The local linear estimator (*dashed line*) of the trend function $\{\alpha(\cdot)\}$ and seasonal effect functions $\{\beta_j(\cdot)\}$ (*solid line*)

Table 1 The median and standard deviation of the 500 MADE values for Example 1

n	$\mathcal{E}$	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$
50	0.1511 (0.0309)	0.0409 (0.0064)	0.0301 (0.0049)	0.0407 (0.0067)	0.0301 (0.0048)
100	0.1324 (0.0241)	0.0263 (0.0037)	0.0212 (0.0032)	0.0262 (0.0035)	0.0210 (0.0031)
300	0.0931 (0.0137)	0.0131 (0.0018)	0.0125 (0.0018)	0.0130 (0.0018)	0.0121 (0.0018)

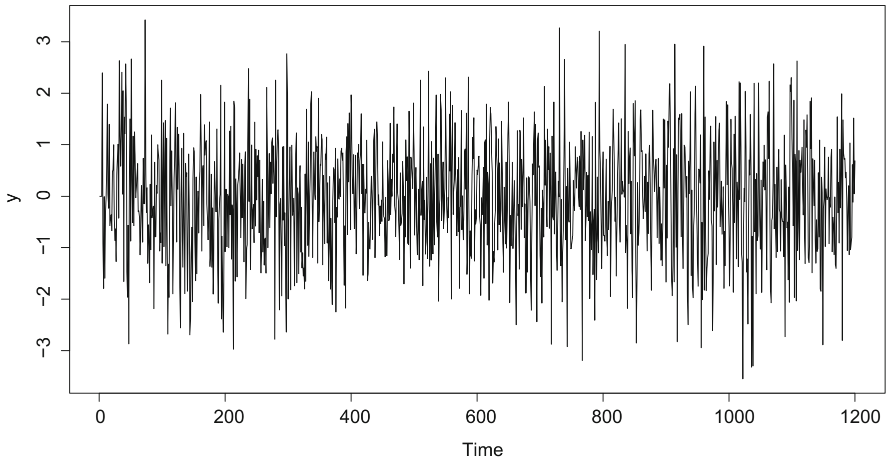

Fig. 3 Time plot of a typical sample from Example 2, with $n = 300$

values. The median and standard deviation (in parentheses) of the 500 MADE values are summarized in Table 1, which confirms that all the MADE values decrease as n increases, as dictated by the asymptotic theory. Clearly, the proposed modeling procedure performs fairly well.

Example 2 In this example, a seasonal AR model with functional coefficients is considered.

$$y_{tj} = (\alpha_1(s_t) + \beta_{1j}(s_t))y_{t,j-1} + (\alpha_2(s_t)) \\ +\beta_{2j}(s_t))y_{t-1,j} + e_{tj}, \quad t = 1, \ldots, n, \quad j = 1, \ldots, 4,$$

where $s_t = t/n$, $y_{t,0} = y_{t-1,4}$, $\alpha_1(x) = 0.5x^2 + 0.5x + 0.13$, $\beta_{11}(x) = -0.8x^2 + 0.5$, $\beta_{12}(x) = 0.2x^3 + 0.8x^2 - 0.4x$, $\beta_{13}(x) = 0.7x^4 - 0.1x^3 - 0.15x$, $\alpha_2(x) = 0.17\sin(2\pi x) - 0.2$, $\beta_{21}(x) = -0.5\cos(\pi x) + 0.1$, $\beta_{22}(x) = -0.5\sin(0.5\pi x) + 0.3$, $\beta_{23}(x) = -0.5\cos(0.5\pi x)$, and $\beta_{k4}(x) = -\beta_{k1}(x) - \beta_{k2}(x) - \beta_{k3}(x)$, $k = 1, 2$, for $0 < x \leq 1$. The errors, $\{e_{tj}\}$, are i.i.d. distributed as $N(0, 1)$. The seasonal AR coefficients at lag 1 are polynomial functions, and the seasonal AR coefficients at lag 4 are a combination of trigonometric functions plus some constants.

The sample sizes used are $n = 300$, 500, and 1000, respectively. For a typical sample with the sample size $n = 300$, Figs. 3 and 4 give the time plots of $\{y_t\}$ and the

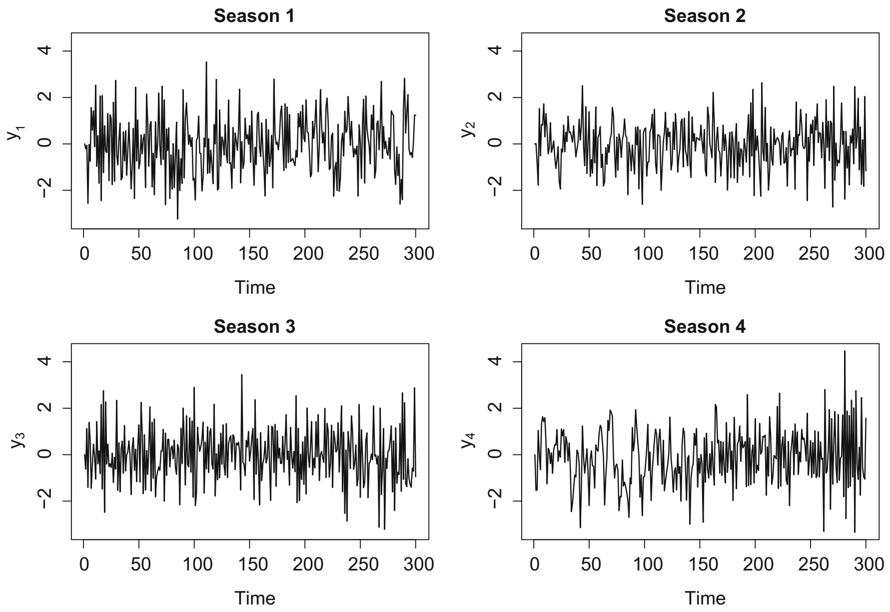

Fig. 4 Time plots of subseries y_{tj} for each season of a typical sample from Example 2 shown in Fig. 3

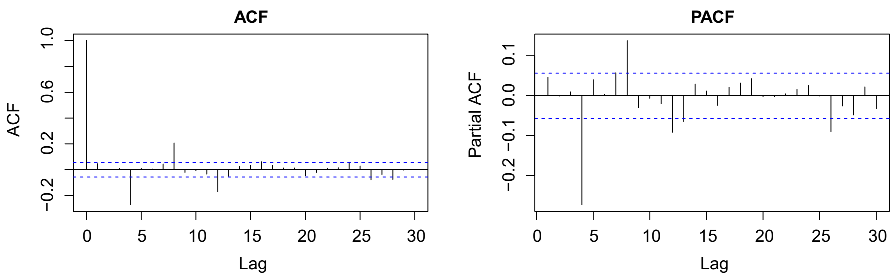

Fig. 5 ACF and PACF for a typical sample from Example 2 shown in Fig. 3

subseries $\{y_{tj}\}$ for each season. The seasonal pattern of the time series is not revealed here. However, the ACF and PACF of the time series (Fig. 5) demonstrate a clear indication of seasonality.

Figure 6 plots the estimated $\alpha_k(\cdot)$ and $\{\beta_{kj}(\cdot)\}$ (dashed lines) from a typical sample with $n = 300$, together with their true values (solid lines). It is seen that the estimation is reasonable, considering the small sample size. Note that the main function $\alpha_k(\cdot)$ has a much smaller scale than the rest of the functions. The median and standard deviation (in parentheses) of the 500 MADE values are summarized in Table 2.

4 An analysis of the Hawaiian tourism data

As a major international tourist site, Hawaii's economy relies heavily on tourism. For planning, marketing and pricing purposes, a deep understanding of the dynamics and

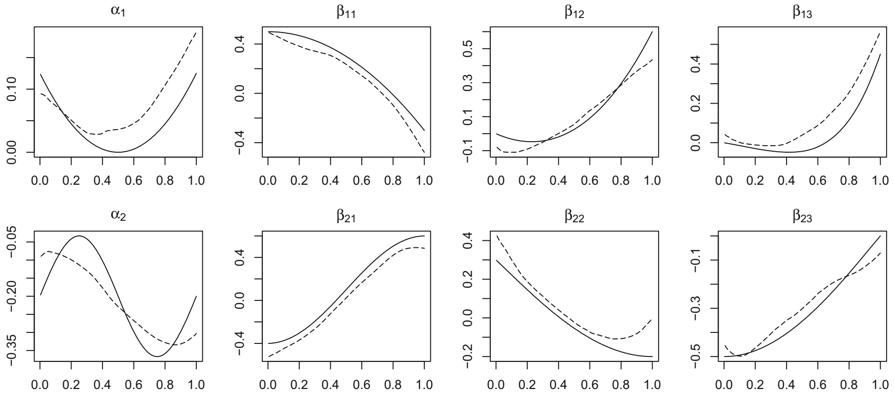

Fig. 6 Estimation results for a typical sample from Example 2 with $n = 300$. The local linear estimator (*dashed line*) of the trend function $\{\alpha_k(\cdot)\}$ and seasonal effect functions $\{\beta_{kj}(\cdot)\}$ (*solid line*)

Table 2 The median and standard deviation of the 500 MADE values for Example 2

n	$\mathcal{E}_1$	$\mathcal{E}_{11}$	$\mathcal{E}_{12}$	$\mathcal{E}_{13}$	$\mathcal{E}_{14}$
300	0.0345 (0.0146)	0.0577 (0.0247)	0.0607 (0.0273)	0.0614 (0.0270)	0.0636 (0.0263)
500	0.0275 (0.0108)	0.0447 (0.0177)	0.0487 (0.0207)	0.0517 (0.0193)	0.0508 (0.0206)
1000	0.0209 (0.0072)	0.0342 (0.0128)	0.0378 (0.0135)	0.0385 (0.0130)	0.0379 (0.0133)
n	$\mathcal{E}_2$	$\mathcal{E}_{21}$	$\mathcal{E}_{22}$	$\mathcal{E}_{23}$	$\mathcal{E}_{24}$
300	0.0534 (0.0131)	0.0581 (0.0256)	0.0590 (0.0245)	0.0540 (0.0251)	0.0593 (0.0233)
500	0.0455 (0.0120)	0.0474 (0.0187)	0.0466 (0.0200)	0.0414 (0.0174)	0.0475 (0.0177)
1000	0.0312 (0.0095)	0.0369 (0.0131)	0.0355 (0.0135)	0.0336 (0.0125)	0.0363 (0.0133)

a capability of accurate prediction of the number of tourists visiting Hawaii are very important to the tourist business and local economy in Hawaii. Due to weather, school schedule and other factors, the number of tourists often shows seasonality. Chen and Fomby (1999) used the stable seasonal pattern model to fit the monthly time series of number of tourists visiting Hawaii. Here we apply the proposed functional-coefficient seasonal time series model to analyze an updated version of Hawaiian tourism data (1970–2012), obtained from the Hawaii Visitors Bureau. Hence, $n = 43$, $d = 12$ and $T = 516$.

For expositional convenience, we re-scale the data by dividing 10^5. Figure 7a presents the monthly observations from January 1970 through December 2012 with the yearly averages (thick line). It demonstrates that the number of tourists visiting Hawaii experienced two growing stages. In the first stage, it increased rapidly from 1970 to 1990. In the second stage, the number of tourists still rose steadily from 1991 to 2012 although there were three down turns, which happened in the early 1990s (the economy recession), September 2001 (the 9/11 tragedy), and 2007–2010 (after the financial crisis), respectively. Figure 7b plots the monthly subseries $\{y_{tj}\}$ for each

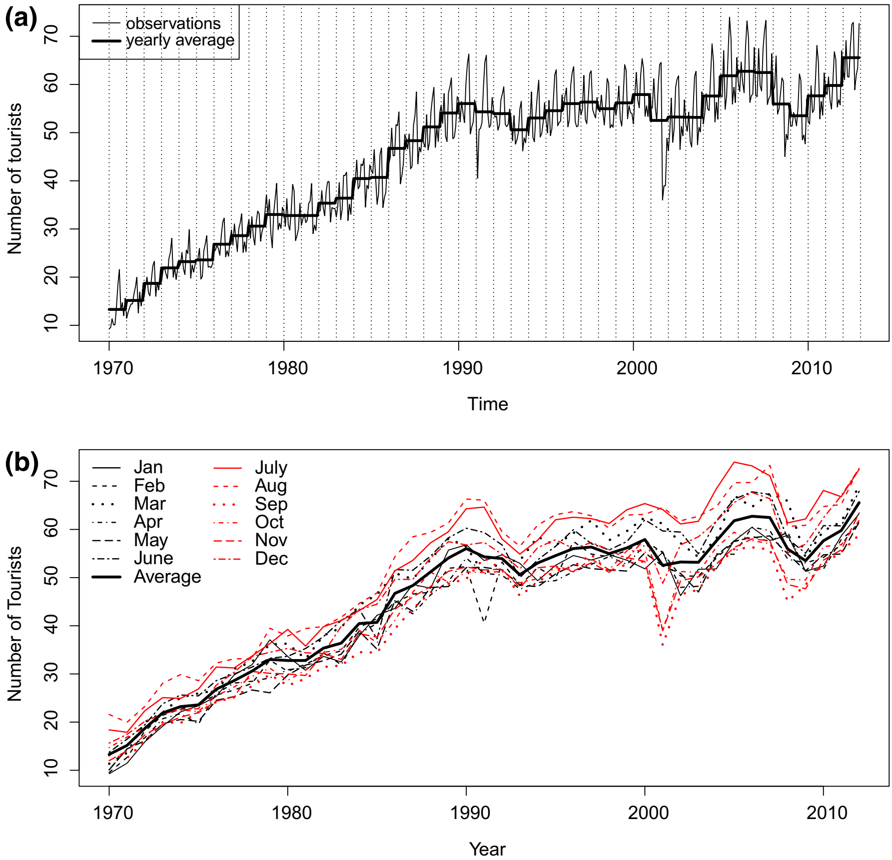

Fig. 7 Hawaiian tourism data from 1970 to 2012. **a** Time series plot of number of visitors (*solid line*) with yearly average (*thick line*), **b** time series plot of number of visitors for each month with yearly average (*thick line*)

month over the years. To see more clearly the seasonality, Fig. 8 gives the boxplot of deviations from the yearly average for each month. It shows that the heaviest travelled months in Hawaii are March, December and the summer.

We first use the nonparametric seasonal model

$$y_{tj} = \alpha(s_t) + \beta_j(s_t) + e_{tj}, \quad t = 1, \ldots, 43, \quad j = 1, \ldots, 12, \tag{10}$$

to fit the series, with the constraint $\sum_{j=1}^{12} \beta_j(s) = 0$ for all $s \in (0, 1]$. Figure 9a plots the estimated trend function (solid line) plus/minus twice estimated pointwise standard errors (dashed lines) with the bias ignored. The yearly average (thick line) is also included. We can see that the 95 % confidence interval covers most of the observed yearly averages except these in 1990–1992, in 2001 and around 2008 due to the economy recessions and the terrorist attack. Such sudden changes may cause additional bias in the estimation.

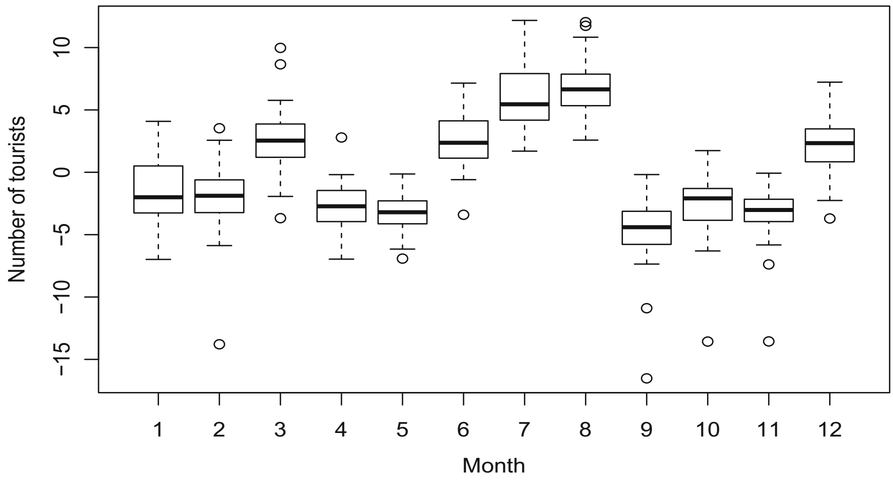

Fig. 8 Hawaiian tourism data from 1970 to 2012. *Boxplot* of deviations from the yearly average for each month

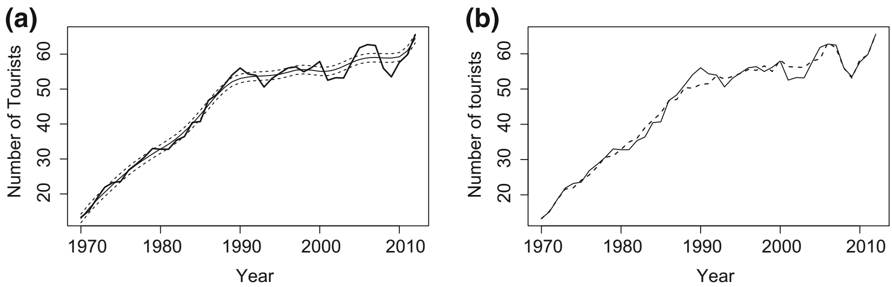

Fig. 9 Hawaiian tourism data from 1970 to 2012. **a** Estimated trend function (*solid line*) plus/minus twice estimated standard errors (*dashed lines*) with bias ignored and the yearly average (*thick line*) for model (10), **b** estimated trend function (*dashed line*) with the yearly average for model (11)

Figure 10 shows the estimated seasonal effect functions, and it can be seen that the seasonal effect functions of March, December and the summer months are all positive, and for the rest of them are negative. Also, the range of the seasonal effect functions increases over time, as the yearly average. Such dynamics are expected. In addition, economy downturn in 1990 has the largest negative impact on February and March; the 9/11 tragedy decreases the tourists severely on September, October, November and December in 2001; and financial crisis after 2007 does not make some very sharp turning points for seasonal functions, because its influence lasts for a few years (yearly average reduces greatly in 2008–2010). It is also interesting to see that December is becoming more and more popular to visit Hawaii in the recent years.

To model more accurately the negative impacts of tourism around 1991–1992 and 2008–2010, partially due to the economy recessions in the U.S. around these two periods, we incorporate some economic indices as exogenous variables. Since U.S. and Japan are the major regions that contribute about 85 % of the tourists to visit

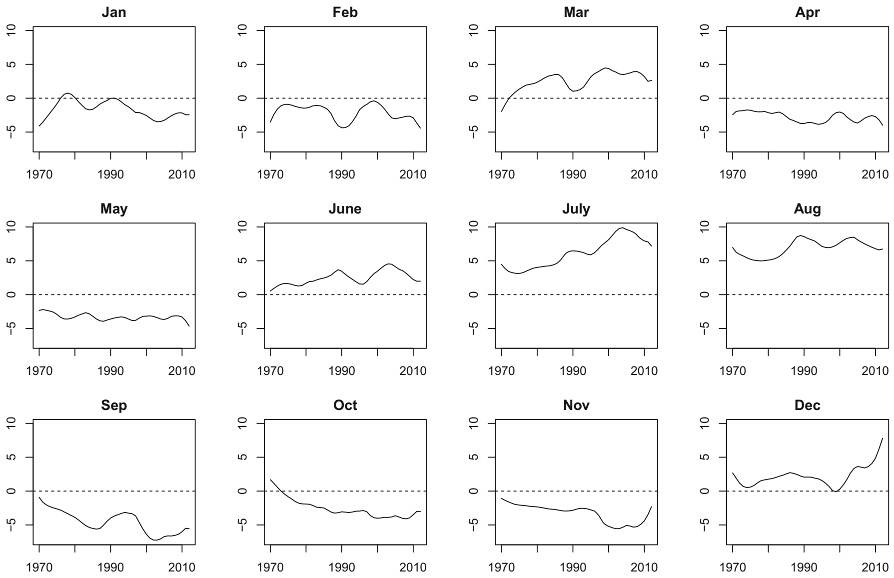

Fig. 10 Hawaiian tourism data from 1970 to 2012. Estimated seasonal functions (*solid line*) with the zero line (*dashed line*) for model (10)

Hawaii, we add the growth rate of annual personal disposable income (PDI) of both countries to the model, as in Chen and Fomby (1999). They are denoted by x_{1t} and x_{2t} for U.S. and Japan, respectively.

Specifically, we consider the following seasonal functional-coefficient model

$$y_{tj} = [\alpha_0(s_t) + \beta_{0j}(s_t)] + [\alpha_1(s_t) + \beta_{1j}(s_t)]x_{1t} + [\alpha_2(s_t) + \beta_{2j}(s_t)]x_{2t} + e_{tj}, \quad (11)$$

$t = 1, \ldots, 43,\ j = 1, \ldots, 12$, subject to the constraints

$$\sum_{j=1}^{12} \beta_{kj}(s) = 0 \quad \text{for each } k = 0, 1, 2 \text{ and all } s \in (0, 1].$$

Comparing to model (10), the two extra terms in model (11) try to make adjustments using the economic variables. In Fig. 9b, the dash line shows the estimated overall trend function $\widehat{\alpha}_0(s_t) + \widehat{\alpha}_1(s_t)\, x_{1t} + \widehat{\alpha}_2(s_t)\, x_{2t}$ against t, calculated with the observed values of x_{1t} and x_{2t}. The solid line shows observed yearly average. It is roughly the same as that using the simpler model (10) before 2005, but the adjustment to the overall annual trend improves the estimation for years after 2005 significantly. The estimated seasonal functions $\widehat{\beta}_{0j}(s_t) + \widehat{\beta}_{1j}(s_t)\, x_{1t} + \widehat{\beta}_{2j}(s_t)\, x_{2t}$, plotted against time, again calculated with the observed values of x_{1t} and x_{2t}, are depicted in Fig. 11. The basic shapes of the seasonal functions remain similar as those shown in Fig. 10, but the extra terms using economic indices make the seasonal functions less smooth and reflect the significant influence of the financial crisis.

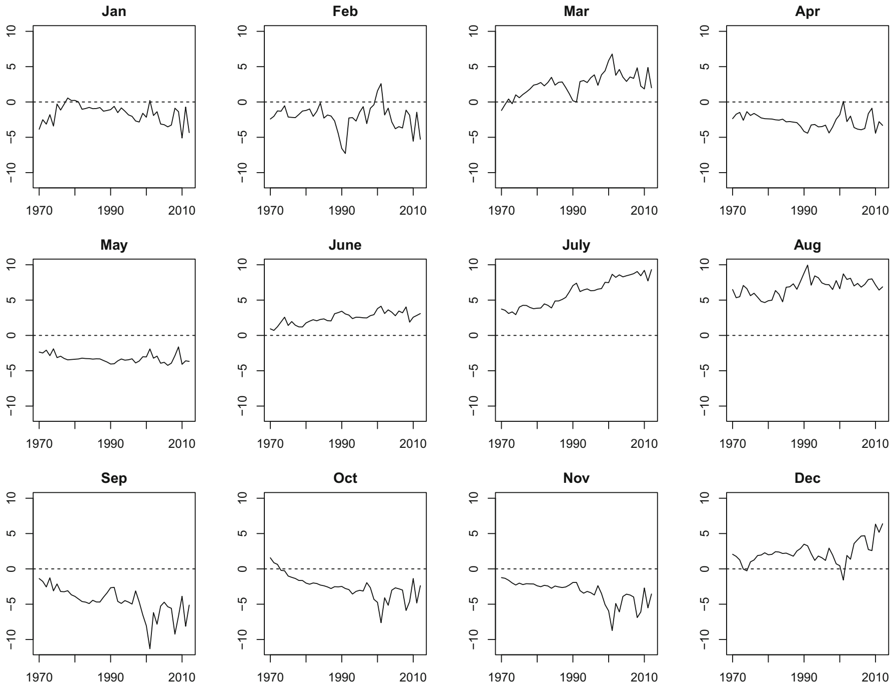

Fig. 11 Hawaiian tourism data from 1970 to 2012. Estimated seasonal functions (*solid line*) with the zero line (*dashed line*) for model (11)

Figure 12 shows the estimated seasonal trend, β_{0i} for $i = 1, \ldots, 12$. Comparing with Fig. 10, we can see that for most months, the pattern of the trend remains similar. However, β_{01} and β_{08} estimated for model (11) have different trend from these estimated for model (10). They increased in 1970s, and decreased after 2000 for model (11), while they basically remained at the same level for model (10). It partially indicates that the increases of tourism in January and August are due to the growth of PDI from 2000 to 2012 based on model (11).

Figures 13 and 14 display the estimated seasonal income effects of U.S. and Japan, respectively. For most months, the income effect of U.S. was weakened from 1970 to 1985, and then was strengthen until subprime mortgage crisis. For Japan, the income effect for the first half of the year was rather weak. However, PDI had a very strong impact on the number of tourists for the second half of the year, especially after 1995, and it had been increasing over the whole period.

We select two sets of estimated functions which have the larger variations among 12 months, and whose estimated seasonal trend functions are different from these for model (10), plotted in Fig. 15. The estimation results give us some detailed explanations of the Hawaiian tourism data. Figure 15a presents α_1, the overall income effect of U.S. over time. It is seen that the income effect decreased in the period of 1970–1988, then gradually increased until the financial crisis. However, the additional income effect of U.S. in the month of January β_{11} decreased in the period of 1970–1988, and then

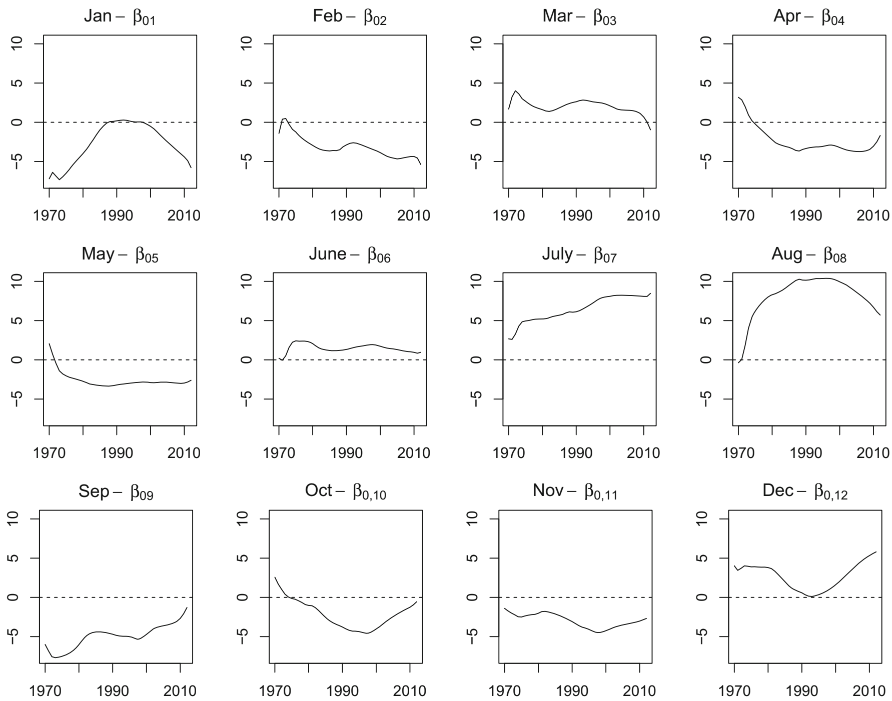

Fig. 12 Hawaiian tourism data from 1970 to 2012 for model (11). Estimated seasonal trend β_{0i} for $i = 1, \ldots, 12$

increased, shown in Fig. 15b. The overall income effect on U.S. in January over time is plotted in Fig. 15c. The overall income effect for Japan α_2 increases consistently over time, shown in Fig. 15d. The income effect of the month of August β_{28} decreased after 1995, and then rose very sharply after 2007 (Fig. 15e). Overall, the income growth becomes a more and more deciding factor on the number of Japanese tourists visiting Hawaii in August.

We compare model (11) with the seasonal ARIMA model by out-of-sample rolling forecasting. Specifically, for $m_0 = T_0, \ldots, T - \ell$, we use data observed at time m_0, $\{y_j, j = 1, \ldots, m_0\}$ to predict number of tourists visiting Hawaii at time $m_0 + \ell$, $\{y_{m_0+\ell}\}$, where the forecast horizon is ℓ months. Here we set $T_0 = 408$, and let ℓ take values from 1 to 48. For computational convenience, when the forecast origin is m_0, where $m_0 = 12(t_0 - 1) + j_0$, $1 \leq j_0 \leq 12$, data is separated into $(t_0 - 1)$ periods, not by the calendar year, but by the following rule: the months $j_0 + 1, \ldots, 12$ in the year h, and the months $1, \ldots, j_0$ in the year $h + 1$ are defined as the h-th period, for each $h = 1, \ldots, t_0 - 1$. In other words, data is separated into periods as $\{y_{j_0+1}, \ldots, y_{j_0+12}\}, \ldots, \{y_{m_0-11}, \ldots, y_{m_0}\}$, when the forecast origin is m_0.

Figure 16 shows the time series plot and sample autocorrelations of residuals $\{\widehat{e}_m\}$ by model (11). There is no significant seasonality but serial dependence in the data after extracting seasonal trend and income effects. Hence, we specify an AR(1) model for the residuals

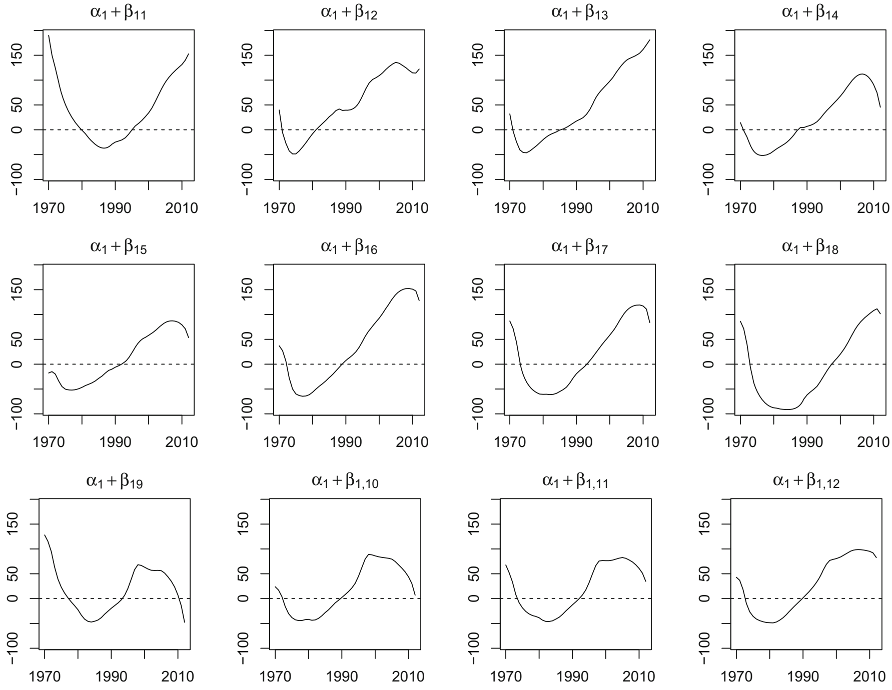

Fig. 13 Hawaiian tourism data from 1970 to 2012 for model (11). Estimated seasonal income effect of U.S. for each month, estimated $\alpha_1 + \beta_{1i}$ for $i = 1, \ldots, 12$

$$\widehat{e}_m = \phi \widehat{e}_{m-1} + \eta_m,$$

where $\{\eta_m\}$ is a white noise process.

When the forecast origin is m_0, ϕ can be estimated by least squares, i.e., $\widehat{\phi} = \arg\min \sum_{m=2}^{m_0} (\widehat{e}_m - \phi \widehat{e}_{m-1})^2$, and $y_{m_0+\ell}$ can be predicted as

$$\widehat{y}_{m_0+\ell} = [\widehat{\alpha}_0(s_t) + \widehat{\beta}_{0j}(s_t)] + [\widehat{\alpha}_1(s_t) + \widehat{\beta}_{1j}(s_t)]x_{1t} + [\widehat{\alpha}_2(s_t) + \widehat{\beta}_{2j}(s_t)]x_{2t} + \widehat{\phi}^{\ell}\, \widehat{e}_{m_0}, \tag{12}$$

where the $(m_0 + \ell)$-th month is the j-th month in the t-th period, i.e. $m_0 + \ell - j_0 = 12(t-1) + j$, $1 \le j \le 12$, $\{\widehat{\alpha}_k(s_t), \widehat{\beta}_{kj}(s_t), k = 0, 1, 2\}$ are estimates of trend and seasonal components in j-th month and t-th period based on data observed at time m_0 with Eq. (9), and $\widehat{e}_{m_0}$ is the residual at time m_0.

For the seasonal ARIMA model, we select the following model based on AIC to fit the data

$$(1 - \phi_1 B)(1 - \phi_{12} B^{12})(1 - B^{12}) y_m = a_m, \tag{13}$$

where B is the back-shift operator, ϕ_1 and ϕ_{12} are the AR coefficient and the seasonal AR coefficient respectively, and $\{a_m\}$ is a white noise process.

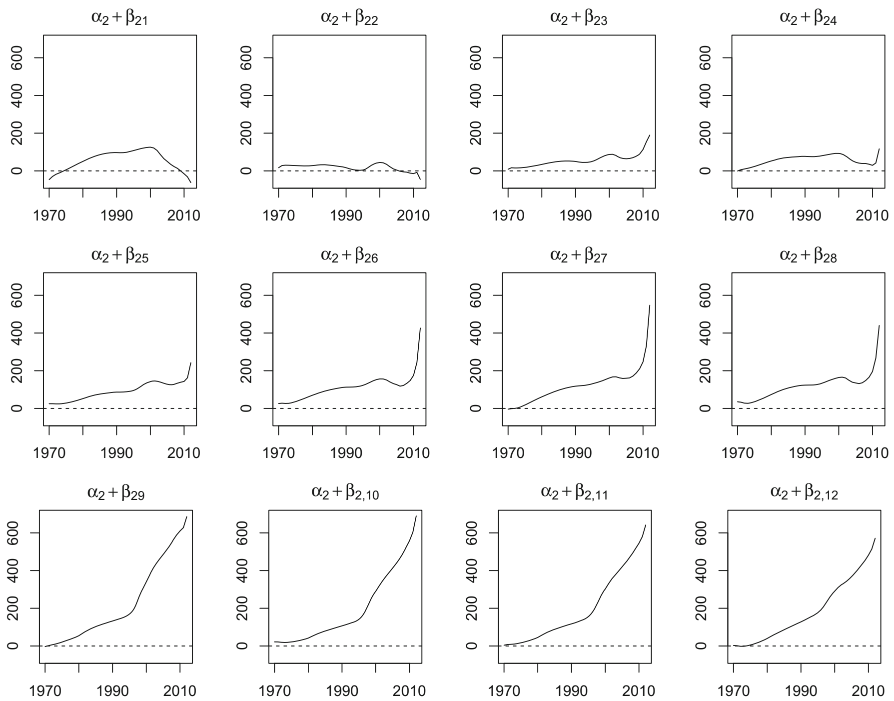

Fig. 14 Hawaiian tourism data from 1970 to 2012 for model (11). Estimated seasonal income effect of Japan for each month, estimated $\alpha_2 + \beta_{2i}$ for $i = 1, \ldots, 12$

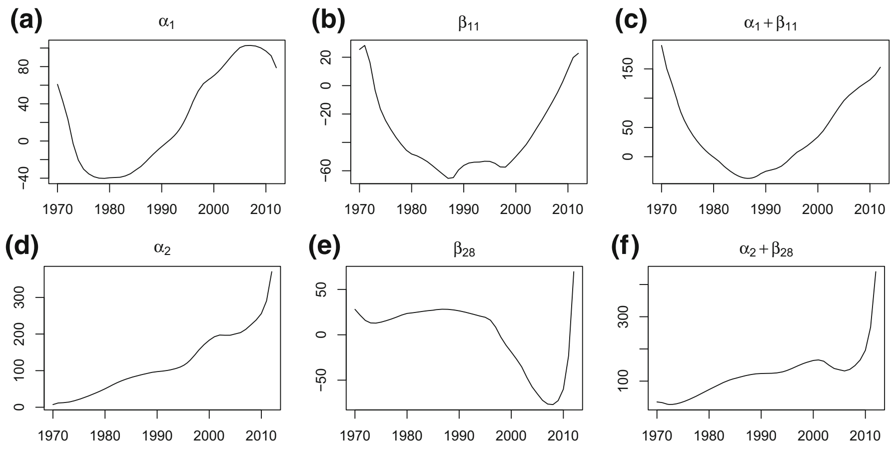

Fig. 15 Hawaiian tourism data from 1970 to 2012 for model (11). *Top panel* shows the seasonal income effect of U.S. in January: **a** estimated α_1, **b** estimated β_{11}, **c** estimated $\alpha_1 + \beta_{11}$. *Bottom panel* shows the seasonal income effect of Japan in August: **d** estimated α_2, **e** estimated β_{28}, **f** estimated $\alpha_2 + \beta_{28}$

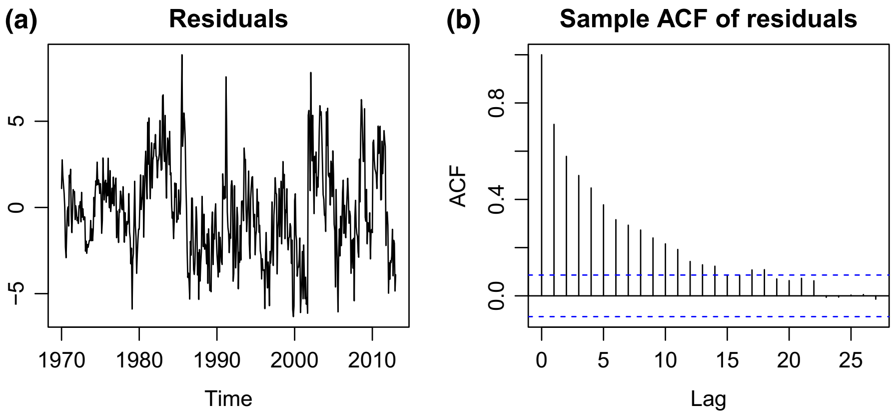

Fig. 16 Hawaiian tourism data from 1970 to 2012. **a** Time series plot of residuals for model (11), **b** sample auto-correlations of residuals for model (11)

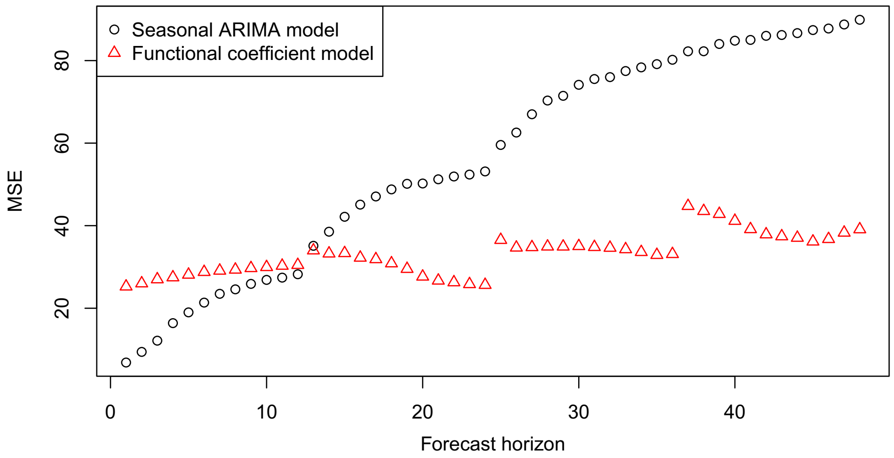

Fig. 17 Hawaiian tourism data from 1970 to 2012. The mean squared forecasting error for Model (11) and the seasonal ARIMA model against different forecast horizon

Figure 17 plots the out-of-sample mean squared prediction error against different forecast horizon for two models. Although the seasonal ARIMA model predicts the number of tourists less than 1-year ahead better than our model, it suffers severely from the increase of forecast horizon. Predictions by our model are more stable, and outperform when forecast horizon is longer. The functional-coefficient seasonal model characterizes the long-term trend of the series, and describes the dynamic relationship between growth of PDI and the number of tourists visiting Hawaii.

5 Concluding remarks

In this article, we propose a nonparametric seasonal time series model with functional coefficients. By allowing the coefficients to change over time, it describes the time-varying impact the trend and possible exogenous variables exert on the process. The seasonal components in the model help to characterize the periodic behaviors of the time series data. The paper focuses on the nonparametric approach, with its flexibility and minimum subjective assumptions. It should be pointed out that the results from the nonparametric approach can be used as a first step for building more parsimonious models which may lead to more accurate and stable estimation and better performance. The proposed method is implemented to analyze the Hawaii tourism data, and results show that our model provides easier interpretation and better long term prediction than a linear seasonal ARIMA models.

Appendix: Mathematical proofs

We first list all the assumptions needed for the asymptotic theory in Sect. 2.2 although some of them might not be the weakest possible. Note that the same notations in Sect. 2 are used here. Throughout this appendix, we denote by C a generic constant, which may take different values at different appearances.

Assumption A A1. The kernel $K(u)$ is symmetric and satisfies the Lipschitz condition and $u\, K(u)$ is bounded.

A2. For each n, $\{(\mathcal{X}_t,\ \mathbf{e}_{nt})\}_{t=1}^{n}$ have the same joint distribution as $\{(\mathcal{X}_t,\ \boldsymbol{\xi}_t)\}_{t=1}^{n}$, where the time series $\{(\mathcal{X}_t,\ \boldsymbol{\xi}_t)\}$ is strictly stationary α-mixing. Assume that there exists $\delta > 0$ such that $E|\boldsymbol{\xi}_t|^{2(1+\delta)} < \infty$, $E|\mathcal{X}_t|^{4(1+\delta)} < \infty$, and the mixing coefficient $\alpha(n) = O\left(n^{-(2+\delta)(1+\delta)/\delta}\right)$.

A3. $n\, h^{1+4/\delta} \to \infty$.

Remark 10 Let $r_{jm}(k)$ denote the (j, m)-th element of $\mathbf{R}(k)$. By the Davydov's inequality (see, e.g., Corollary A.2 in Hall and Heyde 1980), Assumption A2 implies that $|r_{jm}(k)| \le C\, \alpha^{\delta/(2+\delta)}(k)$ so that $\sum_{k=-\infty}^{\infty} |r_{jm}(k)| < \infty$.

Lemma A1 *Under the assumptions of Theorem 1, we have*

$$Var(\mathbf{P}_{n0}) = \nu_0\, \boldsymbol{\Sigma}_0 + o(1) \quad \textit{and} \quad \mathbf{P}_{n1} = o_p(1),$$

where, for $k = 0, 1$,

$$\mathbf{P}_{nk} = h^{1/2}\, n^{-1/2} \sum_{t=1}^{n} (s_t - s)^k\, \mathbf{e}^*_{nt}\, K_h(s_t - s)$$

with $\mathbf{e}^*_{nt} = \mathcal{X}_t^T\, \mathbf{e}_{nt}$.

Proof By the stationarity of $\{\boldsymbol{\xi}_j\}$ and $\mathcal{X}_t$,

$$\begin{aligned}\text{var}(\mathbf{P}_{n0}) &= n^{-1}\, h \sum_{1 \le k,\, l \le n} \mathbf{R}(k - l)\, K_h(s_k - s)\, K_h(s_l - s) \\ &= n^{-1}\, h\, \mathbf{R}(0) \sum_{k=1}^{n} K_h^2(s_k - s)\end{aligned}$$

$$+2\,n^{-1}\,h \sum_{1\le l<k\le n} \mathbf{R}(k-l)\,K_h(s_k-s)\,K_h(s_l-s)$$
$$\equiv \mathbf{I}_1+\mathbf{I}_2.$$

Clearly, by the Riemann sum approximation of an integral,

$$\mathbf{I}_1 \approx \mathbf{R}(0)\,h\int_0^1 K_h^2(u-s)du \approx \nu_0\,\mathbf{R}(0).$$

Since $n\,h\to\infty$, there exists $c_n\to\infty$ such that $c_n/(n\,h)\to 0$. Let $S_1=\{(k,l):1\le k-l\le c_n;\ 1\le l<k\le n\}$ and $S_2=\{(k,l):\ 1\le l<k\le n\}\backslash S_1$. Then, $\mathbf{I}_2$ is split into two terms as $\sum_{S_1}(\cdots)$, denoted by $\mathbf{I}_{21}$, and $\sum_{S_2}(\cdots)$, denoted by $\mathbf{I}_{22}$. Since $K(\cdot)$ is bounded, then, $K_h(\cdot)\le C/h$ and $n^{-1}\sum_{k=1}^n K_h(t_k-t)\le C$. In conjunction with the Davydov's inequality (see, e.g., Corollary A.2 in Hall and Heyde 1980), we have, for the (j,m)-th element of $\mathbf{I}_{22}$,

$$\begin{aligned}|\mathbf{I}_{22(jm)}| &\le 2n^{-1}h\sum_{S_2}|r_{jm}(k-l)|\,K_h(s_k-s)\,K_h(s_l-s)\\ &\le C\,n^{-1}\,h\sum_{S_2}\alpha^{\delta/(2+\delta)}(k-l)K_h(s_k-s)\,K_h(s_l-s)\\ &\le C\,n^{-1}\sum_{k=1}^n K_h(s_k-s)\sum_{k_1>c_n}\alpha^{\delta/(2+\delta)}(k_1)\\ &\le C\sum_{k_1>c_n}\alpha^{\delta/(2+\delta)}(k_1)\\ &\le C\,c_n^{-\delta}\to 0\end{aligned}$$

by Assumption A2 and the fact that $c_n\to\infty$. For any $(k,l)\in S_1$, by Assumption A1

$$|K_h(s_k-s)-K_h(s_l-s)|\le C\,h^{-1}\,(s_k-s_l)/h\le C\,c_n/(n\,h^2),$$

which implies that

$$\begin{aligned}|\mathbf{I}_{212(jm)}| &\equiv \left|2\,n^{-1}\,h\sum_{l=1}^{n-1}\sum_{1\le k-l\le c_n} r_{jm}(k-l)\,\{K_h(s_k-s)-K_h(s_l-s)\}\,K_h(s_l-s)\right|\\ &\le C\,c_n\,n^{-2}\,h^{-1}\sum_{l=1}^{n-1}\sum_{1\le k-l\le c_n}|r_{jm}(k-l)|\,K_h(s_l-s)\\ &\le C\,c_n\,n^{-2}\,h^{-1}\sum_{l=1}^{n-1}K_h(s_l-s)\sum_{k\ge 1}|r_{jm}(k)|\\ &\le C\,c_n/(n\,h)\to 0\end{aligned}$$

by Remark 10 and the fact that $c_n/(n\,h) \to 0$. Therefore,

$$\begin{aligned}\mathbf{I}_{21(jm)} &= 2\,n^{-1}\,h \sum_{l=1}^{n-1} \sum_{1\le k-l\le c_n} r_{jm}(k-l)\,K_h(s_k-s)\,K_h(s_l-s)\\ &= 2\,n^{-1}\,h \sum_{l=1}^{n-1} K_h^2(s_l-s) \sum_{1\le k-l\le c_n} r_{jm}(k-l) + \mathbf{I}_{212(jm)}\\ &\to 2\,\nu_0 \sum_{k=1}^{\infty} r_{jm}(k).\end{aligned}$$

Thus,

$$\mathrm{var}(\mathbf{P}_{n0}) \to \nu_0 \left(\mathbf{R}(0) + 2\sum_{k=1}^{\infty} \mathbf{R}(k) \right) = \nu_0\,\boldsymbol{\Sigma}_0.$$

On the other hand, by Assumption A1, we have

$$\begin{aligned}\mathrm{var}(\mathbf{P}_{n1}) &= n^{-1}\,h \sum_{1\le k,\,l\le n} \mathbf{R}(k-l)\,(s_k-s)(s_l-s)K_h(s_k-s)\,K_h(s_l-s)\\ &\le C\,n^{-1}\,h \sum_{1\le k,\,l\le n} |\mathbf{R}(k-l)|\\ &\le C\,h \sum_{k=-\infty}^{\infty} |\mathbf{R}(k)| \to \mathbf{0}.\end{aligned}$$

This proves the lemma. □

Proof of Theorem 1 Similar to the proof used in Lemma A1, we have

$$h^{-k}\,\mathbf{G}_k(s) = \mu_k\,\mathbf{G} + o_p(1), \tag{14}$$

where $\mathbf{G}_k$ for $k \ge 0$ is defined in (9), so that

$$\begin{pmatrix} \mathbf{G}_0 & \mathbf{G}_1/h \\ \mathbf{G}_1/h & \mathbf{G}_2/h^2 \end{pmatrix} = \mathrm{diag}\{\mathbf{G},\,\mu_2\,\mathbf{G}\} + o_p(1).$$

We re-write $\mathbf{M}_k$ as

$$\mathbf{M}_k = \mathbf{M}_k^* + (nh)^{-1/2}\mathbf{P}_{nk},$$

where $\mathbf{M}_k$ is defined in (9), $\mathbf{P}_{nk}$ is defined in Lemma A1, and $\mathbf{M}_k^* = n^{-1}\sum_{t=1}^{n}(s_t - s)^k\,\mathcal{X}_t^T\boldsymbol{\theta}(s_t)\,K_h(s_t - s)$. By a Taylor expansion, for any $k \ge 0$ and s_t in a neighborhood of s,

$$\mathbf{M}_k^* = \mathbf{G}_k\,\boldsymbol{\theta}(s) + \mathbf{G}_{k+1}\,\boldsymbol{\theta}'(s) + \frac{1}{2}\,\mathbf{G}_{k+2}\,\boldsymbol{\theta}''(s) + o_p(h^2),$$

so that by (9),

$$\begin{pmatrix}\widehat{\boldsymbol{\theta}}_0\\ \widehat{\boldsymbol{\theta}}_1'\end{pmatrix} - \begin{pmatrix}\boldsymbol{\theta}_0\\ \boldsymbol{\theta}_1'\end{pmatrix} = \frac{1}{2}\begin{pmatrix}\mathbf{G}_0 & \mathbf{G}_1\\ \mathbf{G}_1 & \mathbf{G}_2\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{G}_2\\ \mathbf{G}_3\end{pmatrix}\boldsymbol{\theta}''(s) + o_p(h^2) + (nh)^{-1}\begin{pmatrix}\mathbf{G}_0 & \mathbf{G}_1\\ \mathbf{G}_1 & \mathbf{G}_2\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{P}_{n0}\\ \mathbf{P}_{n1}\end{pmatrix},$$

which implies that

$$\sqrt{n\,h}\,\left\{\widehat{\boldsymbol{\theta}}(s) - \boldsymbol{\theta}(s) - \frac{h^2}{2}\,\mu_2\,\boldsymbol{\theta}''(s) + o(h^2)\right\} = \mathbf{G}^{-1}\,\mathbf{P}_{n0} + o_p(1). \tag{15}$$

Therefore, it follows from (15) that the term $\frac{1}{2}\,h^2\,\mu_2\,\boldsymbol{\theta}''(t)$ on the right hand side of (15) serves as the asymptotic bias, and that to establish the asymptotic normality of $\widehat{\boldsymbol{\theta}}(s)$, one only needs to establish the asymptotic normality for $\mathbf{P}_{n0}$. To this end, the Cramér-Wold device is used. For any unit vector $\mathbf{d} \in \Re^d$, let $Z_{n,t} = n^{-1/2}\,h^{1/2}\,\mathbf{d}^T\,\mathbf{e}_{nt}\,K_h(s_t - s)$. Then, $\mathbf{d}^T\,\mathbf{P}_{n0} = \sum_{t=1}^n Z_{n,t}$ and by Lemma A1,

$$\text{var}\left(\mathbf{d}^T\,\mathbf{P}_{n0}\right) = \nu_0\,\mathbf{d}^T\,\boldsymbol{\Sigma}_0\,\mathbf{d}\,\{1 + o(1)\} \equiv \theta_d^2\,\{1 + o(1)\}. \tag{16}$$

Now, the Doob's small-block and large-block technique is used. Namely, partition $\{1, \ldots, n\}$ into $2\,q_n + 1$ subsets with large-block of size $r_n = \lfloor (n\,h)^{1/2} \rfloor$ and small-block of size $s_n = \lfloor (n\,h)^{1/2}/\log n \rfloor$, where $q_n = \left\lfloor \frac{n}{r_n+s_n} \right\rfloor$. Then, $q_n\,\alpha(s_n) \le C\;n^{-(\tau-1)/2}\,h^{-(\tau+1)/2}\,\log^{\tau} n$, where $\tau = (2+\delta)(1+\delta)/\delta$, and $q_n\,\alpha(s_n) \to 0$ by Assumption A3. Let $r_j^* = j\,(r_n + s_n)$ and define the random variables, for $0 \le j \le q_n - 1$,

$$\eta_j = \sum_{t=r_j^*+1}^{r_j^*+r_n} Z_{n,t}, \quad \zeta_j = \sum_{t=r_j^*+r_n+1}^{r_{j+1}^*} Z_{n,t}, \quad \text{and} \quad \mathbf{Q}_{n,3} = \sum_{t=r_{q_n}^*+1}^{n} Z_{n,t}.$$

Then, $\mathbf{d}^T\,\mathbf{P}_{n0} = \mathbf{Q}_{n,1} + \mathbf{Q}_{n,2} + \mathbf{Q}_{n,3}$, where $\mathbf{Q}_{n,1} = \sum_{j=0}^{q_n-1}\eta_j$ and $\mathbf{Q}_{n,2} = \sum_{j=0}^{q_n-1}\zeta_j$. Next we prove the following four facts: (i) as $n \to \infty$,

$$E(\mathbf{Q}_{n,2})^2 \to 0, \quad E(\mathbf{Q}_{n,3})^2 \to 0, \tag{17}$$

(ii) as $n \to \infty$ and $\theta_d^2(t)$ defined as in (16), we have

$$\sum_{j=0}^{q_n-1} E\left(\eta_j^2\right) \to \theta_d^2, \tag{18}$$

(iii) for any s and $n \to \infty$,

$$\left| E\left[\exp(i\, s\, \mathbf{Q}_{n,1})\right] - \prod_{j=0}^{q_n-1} E\left[\exp(i\, s\, \eta_j)\right] \right| \to 0, \tag{19}$$

and (iv) for every $\varepsilon > 0$,

$$\sum_{j=0}^{q_n-1} E\left[\eta_j^2 I\left\{|\eta_j| \geq \varepsilon\, \theta_d\right\}\right] \to 0. \tag{20}$$

(17) implies that $\mathbf{Q}_{n,2}$ and $\mathbf{Q}_{n,3}$ are asymptotically negligible in probability. (19) shows that the summands $\{\eta_j\}$ in $\mathbf{Q}_{n,1}$ are asymptotically independent, and (18) and (20) are the standard Lindeberg-Feller conditions for asymptotic normality of $\mathbf{Q}_{n,1}$ for the independent setup. The rest proof is to establish (17)–(20) and it can be done by following the almost same lines as those used in the Proof of Theorem 2 in Cai et al. (2000) with some modifications. This completes the Proof of Theorem 1. □

References

Andrews DWK (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59:817–858

Auestad B, Tjøstheim D (1990) Identification of nonlinear time series: first order characterization and order determination. Biometrika 77:669–687

Box GEP, Jenkins GM, Reinsel GC (1994) Time series analysis, forecasting and control, 3rd edn. Prentice-Hall, Englewood Cliffs, NJ

Brockwell PJ, Davis RA (1991) Time series theory and methods. Springer, New York

Burman P, Shumway RH (1998) Semiparametric modeling of seasonal time series. J Time Ser Anal 19:127–145

Cai Z (2002) Two-step likelihood estimation procedure for varying-coefficient models. J Multivar Anal 82:189–209

Cai Z (2007) Trending time varying coefficient time series models with serially correlated error. J Econ 136:163–188

Cai Z, Chen R (2006) Flexible seasonal time series models. In: Fomby T, Terrell D (eds) Advances in econometrics volume honoring Engle and Granger, Econometric analysis of financial and economic time series 20:63–87

Cai Z, Fan J, Yao Q (2000) Functional-coefficient regression models for nonlinear time series. J Am Stat Assoc 95:941–956

Cai Z, Li Q, Park JY (2009) Functional-coefficient models for nonstationary time series data. J Econom 148:101–113

Cai Z, Ould-Said E (2001) Local robust regression estimation for time series. Manuscript

Cai Z, Tiwari RC (2000) Application of a local linear autoregressive model to BOD time series. Environmetrics 11:341–350

Chen R, Fomby T (1999) Forecasting with stable seasonal pattern models with an application of Hawaiian tourist data. J Bus Econ Stat 17:497–504

Chen R, Tsay RS (1993a) Functional-coefficient autoregressive models. J Am Stat Assoc 88:298–308

Chen R, Tsay RS (1993b) Nonlinear additive ARX models. J Am Stat Assoc 88:955–967

Cleveland RB, Cleveland WS, McRae JE, Terpenning I (1990) STL: a seasonal-trend decomposition procedure based on loess. J Off Stat 6:3–73

Fan J, Gijbels I (1996) Local polynomial modelling and its applications. Chapman and Hall, London

Fan J, Yao Q (2003) Nonlinear time series: nonparametric and parametric methods. Springer, New York

Fan J, Yao Q, Cai Z (2003) Adaptive varying-coefficient linear models. J R Stat Soc Ser B 65:57–80
Franses PH (1996) Periodicity and stochastic trends in economic time series. Cambridge University Press, New York
Franses PH (1998) Time series models for business and economic forecasting. Cambridge University Press, New York
Gorodetskii VV (1977) On the strong mixing property for linear sequences. Theory Probab Appl 22:411–413
Ghysels E, Osborn DR (2001) The econometric analysis of seasonal time series. Cambridge University Press, New York
Härdle W, Hall P, Ichimura H (1993) Optimal smoothing of single index models. Ann Stat 21:157–178
Härdle W, Lütkepohl H, Chen R (1997) A review of nonparametric time series analysis. Int Stat Rev 65:45–72
Härdle W, Marron S (1985) Optimal bandwidth selection in nonparametric regression function estimation. Ann Stat 13:1465–1481
Härdle W, Müller M, Sperlich S, Werwatz A (2004) Nonparametric and semiparametric models. Springer, Heidelberg
Härdle W, Vieu P (1992) Kernel regression smoothing of time series. J Time Ser 13:209–232
Hastie TJ, Tibshirani R (1993) Varying-coefficient models (with discussion). J R Stat Soc Ser B 55:757–796
Hylleberg S (1992) The historical perspective. In: Hylleberg S (ed) Modelling seasonality. Oxford University Press, Oxford
Jones MC, Marron JS, Sheather SJ (1996) A brief survey of bandwidth selection for density estimation. J Am Stat Assoc 91:401–407
Lu Z (1998) On the ergodicity of non-linear autoregressive model with an autoregressive conditional heteroscedastic term. Stat Sin 8:1205–1217
Lu Z, Tjøstheim D, Yao Q (2007) Adaptive varying-coefficient linear models for stochastic processes: asymptotic theory. Stat Sin 17:177–198
Masry E, Tjøstheim D (1995) Nonparametric estimation and identification of nonlinear ARCH time series: strong convergence and asymptotic normality. Econom Theory 11:258–289
Masry E, Tjøstheim D (1997) Additive nonlinear ARX time series and projection estimates. Econom Theory 13:214–252
Newey WK, West KD (1987) A simple, positive-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55:703–708
Pena D, Tiao GC, Tsay RS (2001) A course in time series analysis. Wiley, New York
Roussas GG (1989) Consistent regression estimation with fixed design points under dependence conditions. Stat Probab Lett 8:41–50
Roussas GG, Tran LT, Ioannides DA (1992) Fixed design regression for time series: asymptotic normality. J Multivar Anal 40:262–291
Ruppert D (1997) Empirical-bias bandwidths for local polynomial nonparametric regression and density estimation. J Am Stat Assoc 92:1049–1062
Shumway RH (1988) Applied statistical time series analysis. Prentice-Hall, Englewood Cliffs, NJ
Shumway RH, Stoffer DS (2000) Time series analysis and its applications. Springer, New York
Tjøstheim D (1994) Non-linear time series: a selective review. Scand J Stat 21:97–130
Tong H (1990) Nonlinear time series: a dynamic system approach. Clarendon Press, Oxford
Tran L, Roussas G, Yakowitz S, Van BT (1996) Fixed-design regression for linear time series. Ann Stat 24:975–991
Wahba G (1977) A survey of some smoothing problems and the method of generalized cross-validation for solving them. In: Krisnaiah PR (ed) Appl Stat. North Holland, Amsterdam, pp 507–523
Withers CS (1981) Conditions for linear processes to be strong mixing. Zeitschrift fur Wahrs- cheinlichkeitstheorie verwandte Gebiete 57:477–480
Xia Y, Li WK (1999a) On the estimation and testing of functional-coefficient linear models. Stat Sin 9:735–757
Xia Y, Li WK (1999b) On single-index coefficient regression models. J Am Stat Assoc 94:1275–1285
Yang L, Park B, Xue L, Härdle W (2006) Estimation and testing for varying coefficients in additive models with marginal integration. J Am Stat Assoc 101:1212–1227