MULTIVARIATE FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS: A NORMALIZATION APPROACH

Jeng-Min Chiou, Yu-Ting Chen and Ya-FangYang

Academia Sinica

Supplementary Material

S1 Simulation settings for the multivariate functional data

We describe additional details in setting the multivariate covariance function and the eigenfunctions along with the eigenvalues for the simulated multivariate functional data. To consider the correlations between the random functions, we set the underlying eigenfunctions $\{\boldsymbol{\phi}_r\}$ coupled with the corresponding eigenvalues $\{\lambda_r\}$ through spectral decomposition of the multivariate correlation function $\boldsymbol{C}$. We consider the following correlation functions.

- The Bessel correlation function of the first kind [Abramowitz and Stegun (1965)],

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{j=0}^{\infty} \frac{(-z^2/4)^j}{j!\,\Gamma(\nu + j + 1)},$$

  with order $\nu = 0$, where $z = |\,t\,|/\rho_1^o$.

- The Matérn correlation function [Minasny and McBratney (2005)],

$$F(z) = \frac{1}{2^{\nu-1}\Gamma(\nu)} z^\nu K_\nu(z),$$

  with order $\nu = 2.5$, where $z = 2|t|\sqrt{\nu}/\rho_2^o$, where

$$K_\nu(z) = \frac{\pi}{2}\frac{I_{-\nu}(z) - I_\nu(z)}{\sin(\nu\pi)}$$

  with the modified Bessel function

$$I_{\pm\nu}(z) = \left(\frac{z}{2}\right)^{\pm\nu} \sum_{j=0}^{\infty} \frac{(z^2/4)^j}{j!\,\Gamma(j \pm \nu + 1)}.$$

- The rational quadratic correlation function [Abrahamsen (1997)],

$$R(z) = \frac{1}{1+z^2}$$

with $z = t/\rho_3^o$.

Here, the constants ρ_1^o, ρ_2^o and ρ_3^o are scale parameters of the correlation functions. We take the $\{1, 2, \ldots, N - j + 1\}$th elements of $J_\nu(z)$, $F(z)$ and $R(z)$ as the elements of the jth row in the upper triangular matrix of C_{kk}, where $C_{kk} = C_{kk}^\top$ and N is the number of recording times and $j = 1, \ldots, N$. In the simulation study, we set $\boldsymbol{\rho}^o = (1, 2.2, 3)$ for Setting I and $\boldsymbol{\rho}^o = (1.1, 8, 8)$ for Setting II. We obtain the eigenfunctions $\{\boldsymbol{\phi}_r\}$ by the following steps.

1. Set the correlation function $C_{kk}(s,t)$, $k = 1, \ldots, 3$, based on the Bessel correlation function, the Matérn correlation function and the rational quadratic correlation function described above. Obtain $\{\vartheta_r\}$ and $\{\varphi_{kr}\}$ through the spectral decomposition of C_{kk} such that $C_{kk}(s,t) = \sum_{r=1}^{\infty} \vartheta_r \varphi_{kr}(s)\varphi_{kr}(t)$.
2. Construct the cross-correlation functions $C_{kl}(s,t)$, $k \neq l$, such that $C_{kl}(s,t) = \sum_{r=1}^{\infty} \bar{\vartheta}_r \varphi_{kr}^*(s)\varphi_{kr}^*(t)$, where $\varphi_{kr}^*(t) = \varphi_{kr}(t)/\sqrt{3}$ and $\bar{\vartheta}_r = (1/3)\sum_{k=1}^{3} \vartheta_{kr}$.
3. Based on the spectral decomposition of $\mathbf{C} = \{C_{kl}; 1 \leq k, l \leq 3\}$, obtain the eigenfunctions $\{\boldsymbol{\phi}_r\}$ with the corresponding eigenvalues $\{\lambda_r\}$ for $r = 1, \ldots, M$, where M is the number of positive eigenvalues.

We can obtain $G_{kl}(s,t)$ simply by $G_{kl}(s,t) = \{v_k(s)v_l(t)\}^{1/2} C_{kl}(s,t)$. We generate the multivariate functional data $\widetilde{\mathbf{Y}}_{ij} = \left(\widetilde{\mathrm{Y}}_{1ij}, \ldots, \widetilde{\mathrm{Y}}_{pij}\right)^\top$, the jth observation of the ith subject observed at t_{ij}, by the truncated version of model (3.1),

$$\widetilde{\mathbf{Y}}_{ij} = \boldsymbol{\mu}(t_{ij}) + \sum_{r=1}^{L} \xi_{ri} \left\{(\boldsymbol{D}\boldsymbol{\phi}_r)(t_{ij})\right\} + \boldsymbol{\epsilon}_{ij}.$$

For Setting I, Figure S1.1 displays true covariance function $G_{kk}(s,t)$ of X_k (diagonal blocks), the cross-covariance functions $G_{kl}(s,t)$ of X_k and X_l (upper triangular part), and the cross-correlation functions $C_{kl}(s,t)$ of Z_k and Z_l (lower triangular part), for $1 \leq k \neq l \leq 3$. Figure S1.2 displays the first four eigenfunctions of $\mathbf{C}$ for $m\mathrm{FPC}_n$. Using the 90% as the threshold for the selection criterion of the percentage of variance explained, the target number of components are 3 for $m\mathrm{FPC}_n$ and $m\mathrm{FPC}_u$ as shown in Figure S1.3 (a)–(b), and are are 2, 3 and 2 for each X_1, X_2 and X_3, as shown in Figure S1.3 (c)–(e).

Similarly, Figure S1.4 displays the true covariance functions $G_{kk}(s,t)$ of X_k (diagonal blocks), the cross-covariance functions $G_{kl}(s,t)$ of X_k and X_l (upper triangular part), and the cross-correlation functions $C_{kl}(s,t)$ of Z_k and Z_l (lower triangular part), $1 \leq k \neq l \leq 3$, for Setting II. Figure S1.5 displays the first four eigenfunctions of $\mathbf{G}$ for $m\mathrm{FPC}_n$. Using the 90% as the threshold for the selection criterion of the percentage of variance explained, the target number of components are 3 for $m\mathrm{FPC}_n$ and $m\mathrm{FPC}_u$ as shown in Figure S1.6 (a)–(b), and are are 2, 1 and 1 for each X_1, X_2 and X_3, as shown in Figure S1.6 (c)–(e).

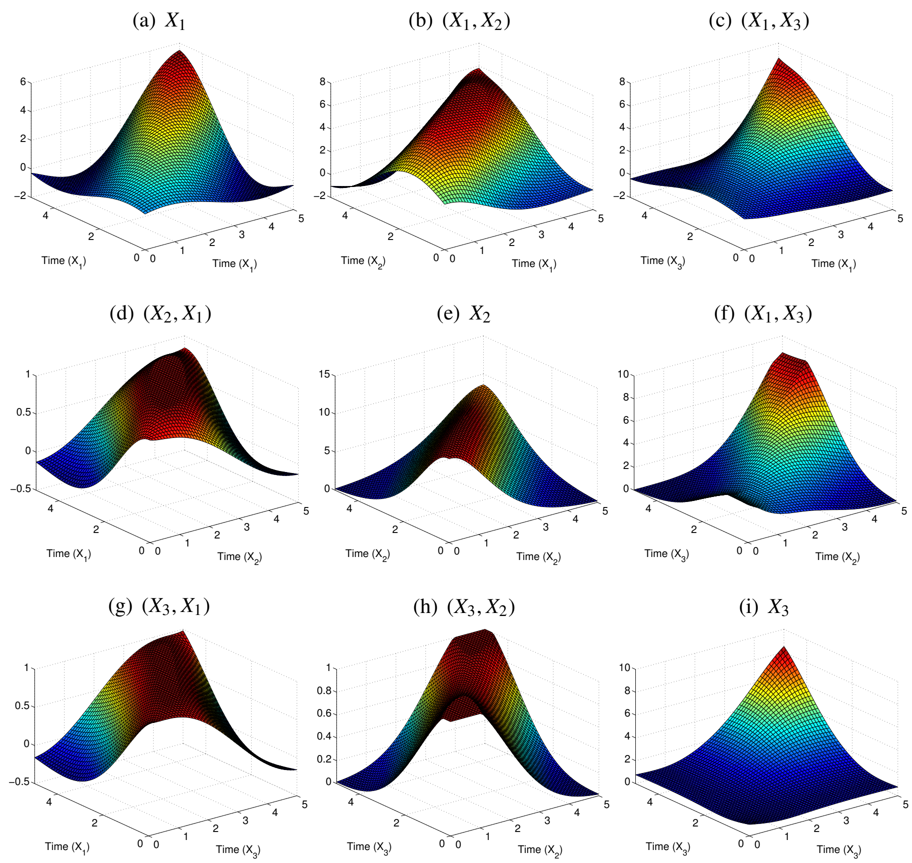

Figure S1.1: True covariance functions $G_{kk}(s,t)$ of X_k (diagonal blocks), the cross-covariance functions $G_{kl}(s,t)$ of X_k and X_j (upper triangular part), and the cross-correlation functions $C_{kl}(s,t)$ of Z_k and Z_l (lower triangular part), $1 \leq k \neq l \leq 3$, for simulation Setting I.

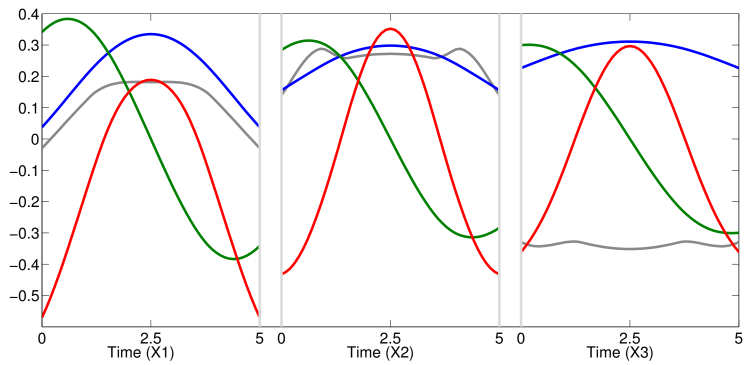

Figure S1.2: The first four true eigenfunctions $\{\phi_{kr}\}$ based on mFPC$_n$ for $r = 1$ (blue), $r = 2$ (green), $r = 3$ (red), and $r = 4$ (gray), in simulation Setting I.

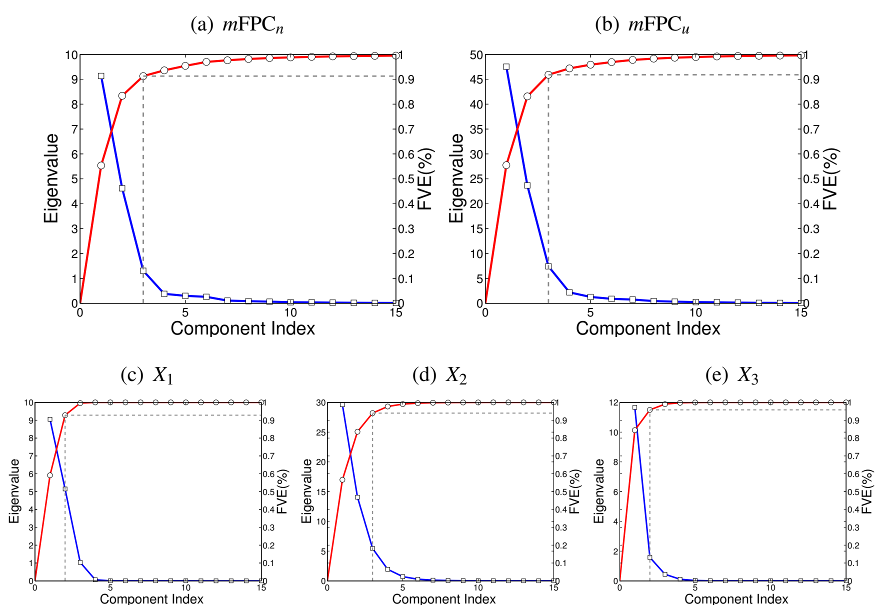

Figure S1.3: The first 15 true eigenvalues and the cumulative fraction of variance (FVE) of total variance explained, obtained by the spectral decomposition of **C** for mFPC$_n$ in (a) and **G** for mFPC$_u$ in (b) and uFPC in (c)–(e), respectively, in simulation Setting I.

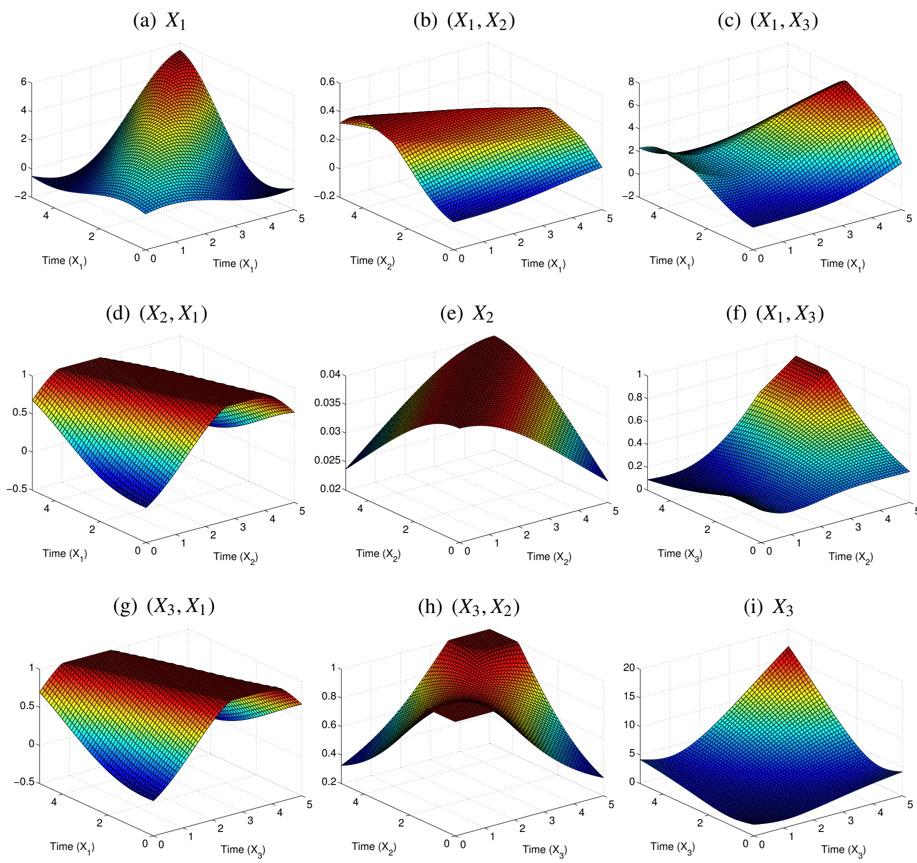

Figure S1.4: True covariance functions $G_{kk}(s,t)$ of X_k (diagonal blocks), the cross-covariance functions $G_{kl}(s,t)$ of X_k and X_j (upper triangular part), and the cross-correlation functions $C_{kl}(s,t)$ of Z_k and Z_l (lower triangular part), $1 \leq k \neq l \leq 3$, for simulation Setting II.

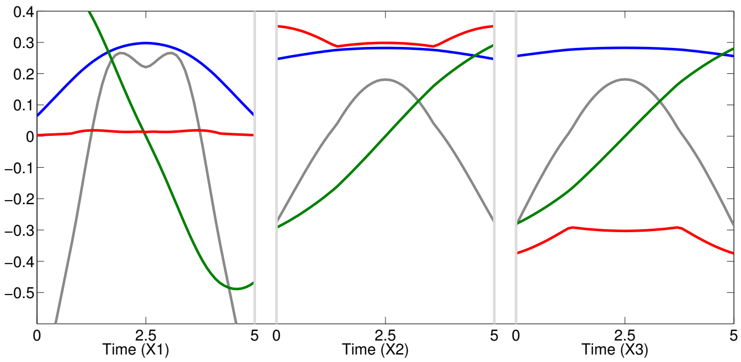

Figure S1.5: The first four true eigenfunctions $\{\phi_{kr}\}$ based on $m\text{FPC}_n$ for $r = 1$ (blue), $r = 2$ (green), $r = 3$ (red), and $r = 4$ (gray), in simulation Setting II.

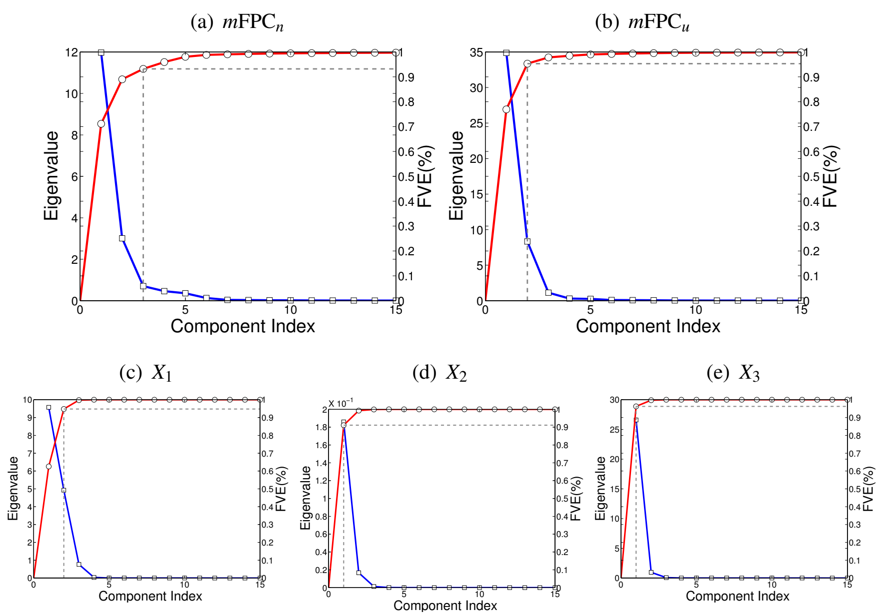

Figure S1.6: True first 15 eigenvalues and the cumulative fraction of variance (FVE) of total variance explained, obtained by the spectral decomposition of $\boldsymbol{C}$ for $m\text{FPC}_n$ and $\boldsymbol{G}$ for $m\text{FPC}_u$ and uFPC, respectively, in simulation Setting II.

References

Abramowitz , M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables* Technical Report **917**, Norwegian Computing Center.

Abramowitz , P. (1997). *A Review of Gaussian Random Fields and Correlation Functions, Norwegian Computing Center*, Dover Publications.

Minasny, B. and McBratney, A. B. (2005) The Matérn function as a general model for soil variograms. *Geoderma* **128** 192-207.

S2 Additional simulation

In the simulation study, we generate the synthetic curves according to the truncated version of (3.1) up to $L = 15$ components. To make the simulated data closer to the real scenario of traffic flow data, the unknown quantities are set as the model estimates of our traffic flow analysis obtained Section 4, including the unknown mean function $\boldsymbol{\mu}(t)$, variance function $\boldsymbol{v}(t)$ and eigenfunction $\boldsymbol{\phi}_{ri}(t)$, where $r = 1, \ldots, 15$. The multivariate FPC scores $\{\xi_{ri}\}$ are generated from $N(0, \lambda_r)$ for each r and the measurement errors $\{\epsilon_{ki}\}$ are generated from $N(0, \sigma_k^2)$ for $k = 1, \ldots, p$, where λ_r and σ_k^2 are also taken from the estimates of the traffic flow analysis. The recording times are equally spaced on [0.25, 24] with 96 time points, mimicking the 15-minute time intervals within a 24-hour period in the traffic flow analysis. We generate $n = 100$ and $n = 500$ multivariate random trajectories for each simulated data set with 200 simulation replicates.

Figure S2.2 illustrates the boxplots of cASE (N_c=1) as defined in (4.1) for the methods

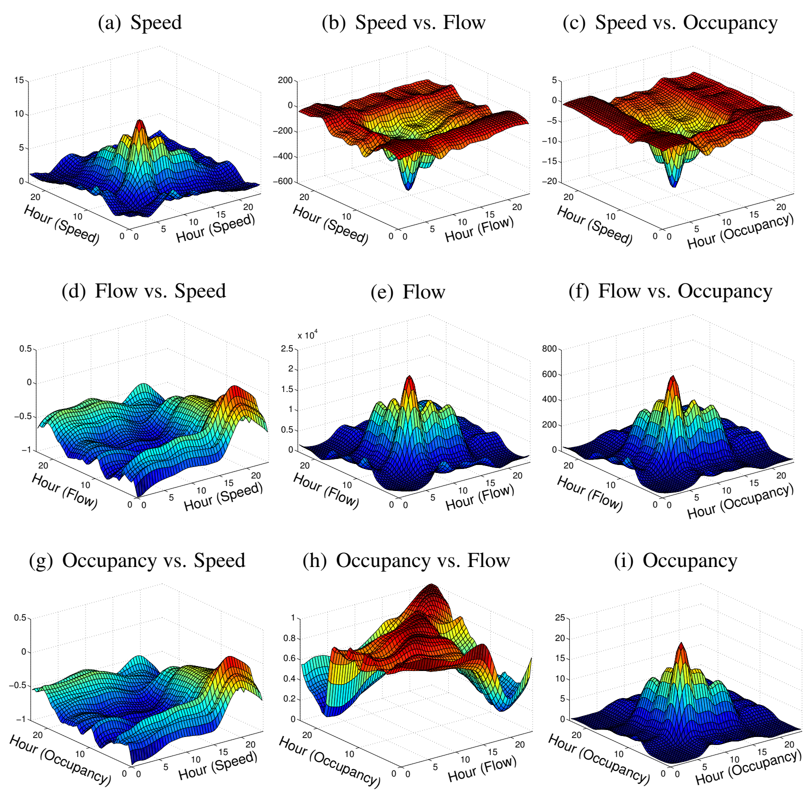

Figure S2.1: The estimated covariance functions $\hat{G}_{kk}(s, t)$ (diagonal blocks) and the estimated cross-covariance functions $\hat{G}_{kl}(s, t)$ (upper triangular part) for X_j, and the estimated cross-covariance functions $\hat{C}_{kl}(s, t)$ (lower triangular part) for Z_j, $1 \leq j \leq 3$ and $1 \leq k \neq l \leq 3$.

mFPC$_n$, mFPC$_u$ and uFPC, and Table S2.1 lists the cASE (N_c=1) ratios of mFPC$_n$ to mFPC$_u$ (denoting the ratio by R1), and those of mFPC$_n$ to uFPC (R2), in terms of cASEs (N_c=1). The results indicate significant reductions in cASE measures from mFPC$_u$ to mFPC$_n$ for the three variables and for the WLS and CE methods. Furthermore, while the cASE measures of mFPC$_n$ in X_1 are slightly larger than uFPC, mFPC$_n$ has significantly smaller cASEs than uFPC in X_2 and X_3. Overall, the proposed mFPC$_n$ perform relatively well in the simulation study. Furthermore, the boxplots of cASEs (N_c=1) in Figure S2.2 also indicate that in mFPC$_n$ the WLS approach performs slightly better than those using CE in this simulation study.

Figure S2.3 displays the boxplots for the number of components and fraction of total vari-

Table S2.1: Relative performance in terms of cASE (N_c=1) ratios of mFPC$_n$ to mFPC$_u$ (R1) and of mFPC$_n$ to uFPC (R2) based on 200 simulation replicates.

	WLS		CE	
Variable	R1	R2	R1	R2
(n=100)				
X_1	0.588	1.020	0.739	1.137
X_2	0.754	0.762	0.825	0.874
X_3	0.702	0.763	0.868	0.875
(n=500)				
X_1	0.635	1.108	0.737	1.154
X_2	0.773	0.783	0.832	0.881
X_3	0.723	0.797	0.879	0.899

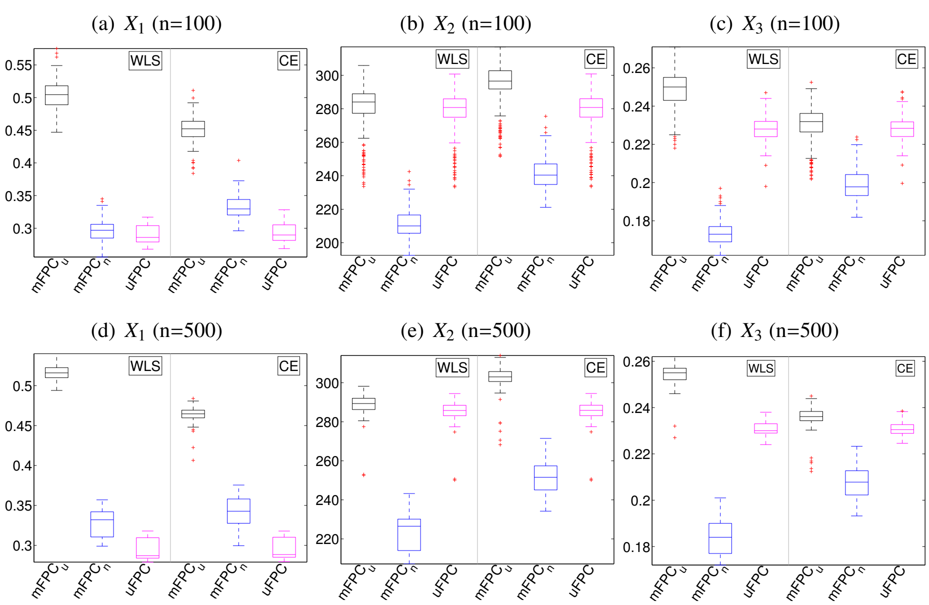

Figure S2.2: Boxplots of ASE based on 200 simulation replicates for comparisons among mFPC$_u$, mFPC$_n$ and uFPC.

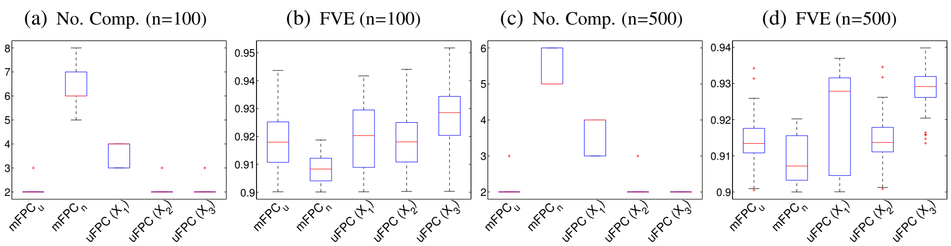

Figure S2.3: Boxplots for the number of components and fraction of total variance explained (FVE) based on 200 simulation replicates for comparisons among $m\mathrm{FPC}_u$, $m\mathrm{FPC}_n$ and uFPC .

ance explained (FVE) based on 200 simulation replicates, with sample curves $n = 100$ and $n = 500$, respectively. Under the criterion of achieving 90% of total variance explained, we see that $m\mathrm{FPC}_n$ selects 5 to 8 components for $n = 100$ and 5 to 6 components for $n = 500$ with the FVE about 90.7%, while $m\mathrm{FPC}_u$ selects 2 components only with the FVE interquartile ranges from 91.1% to 92.5% for $n = 100$ and from 91.1% to 91.8% for $n = 500$. For uFPC, the median number of components for X_1 is 4, while the selected number is 2 for X_2 and X_3, and all the there variables generally have higher FVEs. The results indicates that using the fraction of variance explained criterion for $m\mathrm{FPC}_n$ can adequately select the number of functional principal components.

S3 Additional Proofs

S3.1 Proof of Lemma 5.1

Proof. For (a), we refer to the proofs of Theorem 3.1 in Li and Hsing (2010), which applies one- and two-dimensional local linear regression with convergence rates depending on the bandwidths, sample sizes, and the number of recording times. It follows that, for any $1 \le k \le p$, $\sup_{t\in\mathcal{T}} |\hat{\mu}_k(t) - \mu_k(t)| = O\left(\tau_{n1}(b_{\mu_k})\right)$ *a.s.*

As for (b), we provide a sketch of the proof and point out the differences, with more details in relation to the proof of Theorem 3.3 in Li and Hsing (2010). Let

$$R_{pq} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{j\neq j'}\widetilde{G}_{kk}(T_{ij},T_{ij'})\left(\frac{T_{ij}-s}{b_{G_k}}\right)^p\left(\frac{T_{ij'}-t}{b_{G_k}}\right)^q K\left(\frac{T_{ij}-s}{b_{G_k}}\right)K\left(\frac{T_{ij'}-t}{b_{G_k}}\right).$$

We can write $\hat{G}_{kk}(s,t)$ explicitly, that is

$$\hat{G}_{kk}(s,t) = (\mathcal{A}_1R_{00} - \mathcal{A}_2R_{10} - \mathcal{A}_3R_{01})\mathcal{B}_0^{-1},$$

where $\mathcal{A}_1 = S_{20}S_{02} - S_{11}^2$, $\mathcal{A}_2 = S_{10}S_{02} - S_{01}S_{11}$, $\mathcal{A}_3 = S_{01}S_{20} - S_{10}S_{11}$ and $\mathcal{B}_0 = \mathcal{A}_1S_{00} - \mathcal{A}_2S_{10} - \mathcal{A}_3S_{01}$, with $S_{pq} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{j\neq j'}\left(\frac{T_{ij}-s}{b_{G_k}}\right)^p\left(\frac{T_{ij'}-t}{b_{G_k}}\right)^q K\left(\frac{T_{ij}-s}{b_{G_k}}\right)K\left(\frac{T_{ij'}-t}{b_{G_k}}\right)$. Define $R^*_{pq} = R_{pq} - G_{kk}(s,t)S_{pq} - b_{G_k}\frac{\partial}{\partial s}G_{kk}(s,t)S_{p+1,q} - b_{G_k}\frac{\partial}{\partial t}G_{kk}(s,t)S_{p,q+1}$. It is straightforward to show that

$$\left(\hat{G}_{kk} - G_{kk}\right)(s,t) = (\mathcal{A}_1R^*_{00} - \mathcal{A}_2R^*_{10} - \mathcal{A}_3R^*_{01})\mathcal{B}_0^{-1}. \tag{S3.1}$$

By (5.22) in Li and Hsing (2010), we have $\mathcal{A}_1 = \left[f(s)f(t)\nu_2\right]^2 + O(\delta_{n2}(b_{G_k}) + b_{G_k})$ *a.s.*, $\mathcal{A}_2 = \mathcal{A}_3 = O(\delta_{n2}(b_{G_k}) + b_{G_k})$ *a.s.*, and $\mathcal{B}_0 = f^3(s)f^3(t)\nu_2^2 + O(\delta_{n2}(b_{G_k}) + b_{G_k}^2)$ *a.s.*, where $\delta_{n2}(b_{G_k}) = \left[\left\{1 + (b_{G_k}\gamma_{n1})^{-1} + \left(b_{G_k}^2\gamma_{n2}\right)^{-1}\right\}(\log n/n)\right]^{1/2}$ and $\nu_2^2 = \int_{-1}^{1} t^2K(t)dt$. It remains to investigate the order of R^*_{00}. By definition,

$$\begin{aligned} R^*_{00} &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{i\neq j}\left\{\widetilde{G}_{kk}(T_{ij},T_{ij'}) - G_{kk}(s,t) - \left[\frac{\partial}{\partial s}G_{kk}(s,t)\right](T_{ij}-s)\right. \\ &\qquad \left. - \left[\frac{\partial}{\partial t}G_{kk}(s,t)\right](T_{ij'}-t)\right\}K\left(\frac{T_{ij}-s}{b_{G_k}}\right)K\left(\frac{T_{ij'}-t}{b_{G_k}}\right). \end{aligned}$$

Let $\eta^*_{kijj'} = \widetilde{G}_{kk}(T_{ij},T_{ij'}) - G_{kk}(T_{ij},T_{ij'})$. By the Taylor's expansion,

$$R^*_{00} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\sum_{i\neq j}\eta^*_{kijj'}K\left(\frac{T_{ij}-s}{b_{G_k}}\right)K\left(\frac{T_{ij'}-t}{b_{G_k}}\right) + O(b_{G_k}^2).$$

Since $\sup_{t\in\mathcal{T}}|\mu_k(t) - \hat{\mu}_k(t)| = O(\tau_{n1}(b_{\mu_k}))$ *a.s.* in (a),

$$\begin{aligned} E(\eta^*_{kijj'}|T_{ij},T_{ij'}) &= E\left[\left\{Y_{kij} - \mu_k(T_{ij}) + \left(\mu_k(T_{ij}) - \hat{\mu}_k(T_{ij})\right)\right\}\left\{Y_{kij'} - \mu_k(T_{ij'})\right.\right. \\ &\qquad \left.\left. + \left(\mu_k(T_{ij'}) - \hat{\mu}_k(T_{ij'})\right)\right\}\Big|T_{ij},T_{ij'}\right] - G_{kk}(T_{ij},T_{ij'}) \\ &= G_{kk}(T_{ij},T_{ij'}) + O(\tau_{n1}(b_{\mu_k})) - G_{kk}(T_{ij},T_{ij'}). \end{aligned}$$

We then have $E(R^*_{00}) = E(E(R^*_{00}|T_{ij}, T_{ij'})) = O(\tau_{n1}(b_{\mu_k}))$ and $R^*_{00} = O\left(\tau_{n2}(b_{G_k}) + \tau_{n1}(b_{\mu_k})\right)$ *a.s.*. We note $E(R^*_{00})$ and R^*_{00} are different from those of Li and Hsing (2010) since the raw data $\widetilde{G}_{kk}(T_{ij}, T_{ij'})$ contains the unobservable term $\hat{\mu}_k$. Thus, $\left|\hat{G}_{kk}(s,t) - G_{kk}(s,t)\right| = O\left(\tau_{n2}(b_{G_k}) + \tau_{n1}(b_{\mu_k})\right)$ *a.s.* uniformly in $\mathcal{T}^2$. The results of (b) follows directly by the definition of $\|\cdot\|_2$.

As for (c), recall (3.8) that $\hat{\sigma}^2_k = (2/|\mathcal{T}|)\int_{\mathcal{T}_1}\{\hat{\mathrm{W}}_k(t) - \hat{\mathrm{G}}_{kk}(t,t)\}dt$. We have

$$\left|\hat{\sigma}^2_k - \sigma^2_k\right| \leq \sup_{t\in\mathcal{T}}\left|\hat{\mathrm{W}}_k(t) - \mathrm{W}_k(t)\right| + \sup_{t\in\mathcal{T}}\left|\hat{\mathrm{G}}_{kk}(t,t) - \mathrm{G}_{kk}(t,t)\right|.$$

Following the proof in (a), it is easy to show that when $h_{\mathrm{W}_k} \asymp h_1$,

$$\sup_{t\in\mathcal{T}}\left|\hat{\mathrm{W}}_k(t) - \mathrm{W}_k(t)\right| = O\left(\tau_{n1}(h_{\mathrm{W}_k})\right) a.s.$$

Combining the result of (b), we have $\left|\hat{\sigma}^2_k - \sigma^2_k\right| = O\left(\tau_{n1}(h_{\mathrm{W}_k}) + \tau_{n1}(b_{\mu_k}) + \tau_{n2}(b_{\mathrm{G}_k})\right)$ *a.s.* and therefore, the result of (c) follows directly by the definition of $\|\cdot\|_2$. □

S3.2 Proof of Lemma 5.2

Proof. Note that $|v_k(t) - \hat{v}_k(t)| = \left|v_k^{1/2}(t) - \hat{v}_k^{1/2}(t)\right|\left|v_k^{1/2}(t) + \hat{v}_k^{1/2}(t)\right|$. Lemma 5.1(b) implies that $\hat{v}_k(t)$ is bounded and bounded away from 0 *a.s.* for $1 \leq k \leq p$ and $t \in \mathcal{T}$. Hence $\left|v_k^{1/2}(t) - \hat{v}_k^{1/2}(t)\right| = O\left(\tau_{n2}(b_{\mathrm{G}_k}) + \tau_{n1}(b_{\mu_k})\right)$ *a.s.* Given $0 < m_{v_k} \leq v_k(t) \leq M_{v_k}$ for all $t \in \mathcal{T}$, we have $m_{v_k} - \delta_0 \leq \hat{v}_k(t) \leq M_{v_k} + \delta_0$ *a.s.* for some fixed $\delta_0 > 0$ as $n \gg 0$. There exist M_{μ_k} and M_{Y_k} such that $0 \leq |\mu_k(t)| \leq M_{\mu_k}$ for all $t \in \mathcal{T}$, and $0 \leq \left|\mathrm{Y}_{kij}\right| \leq M_{\mathrm{Y}_k}$ *a.s.*, where the existence of M_{Y_k} is assured by (C5) or (C6). It follows that

$$\begin{aligned}
\max_{1\leq j\leq m_i} |\widetilde{\mathrm{U}}_{kij} - \mathrm{U}_{kij}| &= \max_{1\leq j\leq m_i} \frac{1}{\hat{v}_k^{1/2}(T_{ij})v_k^{1/2}(T_{ij})}\Big|v_k^{1/2}(T_{ij})(\mathrm{Y}_{kij} - \hat{\mu}_k(T_{ij})) \\
&\qquad - \hat{v}_k^{1/2}(T_{ij})(\mathrm{Y}_{kij} - \mu(T_{ij}))\Big| \\
&\leq \frac{1}{m_{v_k}(m_{v_k} - \delta_0)} \sup_{t\in\mathcal{T}}\Big\{(M_{\mathrm{Y}_k} + M_{\mu_k})\left|v_k^{1/2}(t) - \hat{v}_k^{1/2}(t)\right| \\
&\qquad + M_{v_k}\left|\mu_k(t) - \hat{\mu}_k(t)\right|\Big\} \ a.s. \\
&= O\left(\tau_{n2}(b_{\mathrm{G}_k}) + \tau_{n1}(b_{\mu_k})\right) \ a.s.
\end{aligned}$$

□

S3.3 Proof of Lemma 6.1

Proof. (a) Using the notations m_{v_k} and M_{μ_k} for the lower bound of $v_k(t)$ and the upper bound of $\mu_k(t)$, for all $t \in \mathcal{T}$, as in the proof of Lemma 5.2, we have

$$|\mathrm{Z}_k(t)|^\lambda \leq m_{v_k}^{-1}\left\{|\mathrm{X}_k(t)| + M_{\mu_k}\right\}^\lambda = m_{v_k}^{-1}\sum_{s=0}^{\lambda}\binom{\lambda}{s}|\mathrm{X}_k(t)|^s M_{\mu_k}^{\lambda-s}.$$

To show $\mathrm{E}(\sup_{t\in\mathcal{T}} |\mathrm{Z}_k(t)|^s) < \infty$, it is sufficient to show that $\mathrm{E}(\sup_{t\in\mathcal{T}} |\mathrm{X}_k(t)|^s) < \infty$ for $s < \lambda$. Noting that $\sup_{t\in\mathcal{T}} |\mathrm{X}_k(t)|^s = (\sup_{t\in\mathcal{T}} |\mathrm{X}_k(t)|)^s$ and given the probability density function g of $\sup_{t\in\mathcal{T}} |\mathrm{X}_k(t)|$, we have

$$\mathrm{E}(\sup_{t\in\mathcal{T}} |\mathrm{X}_k(t)|^s) = \int_{|x|\le 1} |x|^s\, g(x)dx + \int_{|x|>1} |x|^s\, g(x)dx \le 1 + \sup_{t\in\mathcal{T}} |\mathrm{X}_k(t)|^\lambda < \infty.$$

Further, since $\varepsilon_{kij} = \epsilon_{kij}/v_k(t_{ij})^{1/2}$, boundedness of $\mathrm{E}(\left|\varepsilon_{kij}\right|^{2\lambda_{h_2}})$ follows by the boundedness of $v_k(t)$, which completes the proof of (a). The result of (b) can be shown analogously. □

S3.4 Proof of Corollary 5.1

Proof. (a) For any fixed $t \in \mathcal{T}$, $\|\hat{\mathbf{Z}}_i^{L,WLS}(t) - \mathbf{Z}_i(t)\|_2 \le \|\hat{\mathbf{Z}}_i^{L,WLS}(t) - \mathbf{Z}_i^L(t)\|_2 + \|\mathbf{Z}_i^L(t) - \mathbf{Z}_i(t)\|_2$. Note that $\|\mathbf{Z}_i^L(t) - \mathbf{Z}_i(t)\|_2 \to 0$ in probability as $L \to \infty$, by the Karhunen-Loève theorem. It remains to discuss the asymptotic behavior of $\hat{\mathbf{Z}}_i^{L,WLS}(t) - \mathbf{Z}_i^L(t)$. By Theorem 5.3, the limiting distribution of $(\hat{\boldsymbol{\xi}}_{i,L}^{WLS} - \boldsymbol{\xi}_{i,L})$ is $N(\mathbf{0}, \boldsymbol{\Omega}_{i,L}^{WLS})$ for each L. Since $\hat{\boldsymbol{\phi}}_{L,t} \to \boldsymbol{\phi}_{L,t}$ *a.s.* as $n \to \infty$, we have $\left\{\hat{\mathbf{Z}}_i^{L,WLS}(t) - \mathbf{Z}_i^L(t)\right\}$ convergence in distribution to $N(\mathbf{0}, \omega_{i,L}^{WLS}(t,t))$. It remains to show that

$$\lim_{L\to\infty} \lim_{n\to\infty} \hat{\omega}_{i,L}^{WLS}(t,t) = \omega_i^{WLS}(t,t) \ a.s.$$

We note that $|\hat{\omega}_{i,L}^{WLS}(t,t) - \omega_i^{WLS}(t,t)| \le |\hat{\omega}_{i,L}^{WLS}(t,t) - \omega_{i,L}^{WLS}(t,t)| + |\omega_{i,L}^{WLS}(t,t) - \omega_i^{WLS}(t,t)|$. For a fixed L, $\lim_{n\to\infty} |\hat{\omega}_{i,L}^{WLS}(s,t) - \omega_{i,L}^{WLS}(s,t)| = 0$ *a.s.* by the consistency properties of the estimates of λ_r's, $\boldsymbol{\phi}_r(t)$'s and $\mu_k(t)$'s along with the Slusky's theorem. Further, $\lim_{L\to\infty} |\omega_{i,L}^{WLS}(t,t) - \omega_i^{WLS}(t,t)| = 0$ *a.s.* under (C7), which completes the proof of (a).

(b) By the consistency properties of $\hat{\boldsymbol{\xi}}_{i,L}^{WLS}$ and $\hat{\boldsymbol{\phi}}_{L,t}$, it is sufficient to examine the asymptotic behavior of $\left\{\mathbf{Z}_i^{L,WLS}(t) - \mathbf{Z}_i^L(t)\right\}$, where $\mathbf{Z}_i^{L,WLS}(t) = \boldsymbol{\phi}_{L,t}^\top \boldsymbol{\xi}_{i,L}^{WLS}$. We observe that $\boldsymbol{a}^\top \left\{\mathbf{Z}_i^{L,WLS}(t) - \mathbf{Z}_i^L(t)\right\} = (\boldsymbol{\phi}_{L,t}\boldsymbol{a})^\top \left(\boldsymbol{\xi}_{i,L}^{WLS} - \boldsymbol{\xi}_{i,L}\right)$, where $(\boldsymbol{\phi}_{L,t}\boldsymbol{a})$ is an L-vector. Hence, it reduces to the form of a linear combination of the FPC scores, which is similar to that in Corollary 2 and Theorem 5 of Yao, Müller, and Wang (2005) and, thus, the result follows. □