


REPLY

An Efficient Two-Dimensional Functional Mixed-Effect Model Framework for Repeatedly Measured Functional Data

Cheng Cao¹ | Jiguo Cao²  | Hao Pan³ | Yunting Zhang³ | Fan Jiang³ | Xinyue Li¹ 

¹Department of Data Science, City University of Hong Kong, Kowloon Tong, Hong Kong SAR | ²Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia, Canada | ³Department of Developmental and Behavioral Pediatrics, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

Correspondence: Xinyue Li (xinyueli@cityu.edu.hk)

Received: 20 December 2024 | **Revised:** 17 July 2025 | **Accepted:** 28 July 2025

Funding: This study was supported by City University of Hong Kong research grants (Grant Nos. 7005892, 7005967, and 9610473), The Institute of Digital Medicine, City University of Hong Kong, and The Hong Kong Innovation and Technology Commission (InnoHK Project CIMDA).

Keywords: functional mixed effect model | mental health | physical activity data | wearable device data

ABSTRACT

Advancements in wearable device technology have enabled accelerometers to continuously record minute-by-minute physical activity over consecutive days, yielding curves serially correlated in dense and regular longitudinal design. Motivated by a large-scale cohort of physical activity data throughout a week, the collected repeatedly measured functional data exhibits longitudinal (interday) and functional (intraday) interactions on fine grids. To accommodate this complex data structure and investigate the relationship between health assessment results and weekly physical activity patterns, we propose an innovative and efficient two-dimensional functional mixed-effect model (2dFMM), characterizing the longitudinal and functional cross-variability while incorporating two-dimensional fixed effects and four-dimensional correlation structure in marginal representation. We develop a fast three-stage estimation procedure to provide accurate fixed-effect inference for model interpretability and improve computational efficiency when encountering large datasets. We find strong evidence of intraday and interday varying significant associations between physical activity and mental health assessments among our cohort population, which sheds light on possible intervention strategies targeting daily physical activity patterns to improve school adolescent mental health. Our method is also used in environmental data to illustrate the wide applicability.

1 | Introduction

A growing amount of research suggests an essential relationship between adolescents' physical activity and mental health, and recent studies have further discovered that the context of when and where physical activity occurs is also an influential factor [1–3]. The development of wearable devices has enabled the continuous recording of an individual's physical activity

profiles over consecutive days, allowing for a longitudinal study design by segmenting the time series into daily profiles [4, 5]. Modeling the dependence arising from the temporal correlations between curves is the main concern of longitudinal functional data analysis [6, 7].

Our motivating dataset comes from a Shanghai school adolescent study, designed to investigate the association between students'

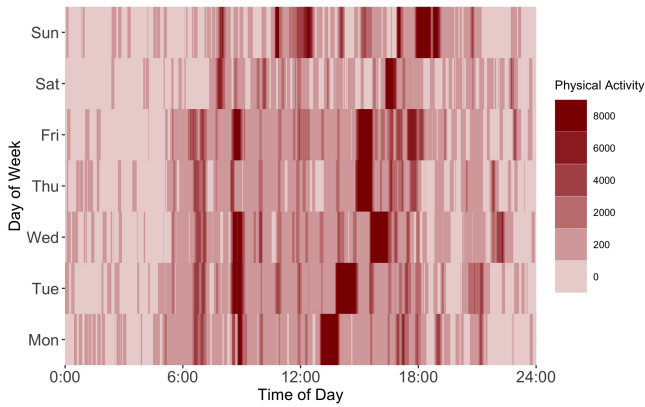


FIGURE 1 | Heatmap of a Shanghai school adolescent's activity profile showing minute-by-minute physical activity counts for a week. Data were obtained from wrist Actigraph.

weekly physical activity patterns and mental health outcomes, while adjusting for demographic covariates. In this study, a total of 2313 students aged between 11 and 18 years from several Shanghai schools wore ActiGraph accelerometers for seven consecutive days to obtain physical activity signals. The signals were then aggregated into minute-by-minute activity counts, yielding 1440 observation grids over a day throughout the week. To standardize longitudinal analyses, all participants' wear days were aligned to a Monday-Sunday sequence. Figure 1 illustrates an example of one participant's weekly activity profile, exemplifying repeatedly measured functional data (RMFD) we focus on in this work. This data exhibits both longitudinal (interday) and functional (intraday) variability within dense and regular sampling grids, requiring fixed-effect inference that preserves the serial dependencies between adjacent days rather than discretizing longitudinal indices or condensing the week's data to single curves. The challenge of modeling the intricate longitudinal and functional cross-variability in RMFD analysis presents an important area of independent scientific inquiry.

Current methods for analyzing RMFD commonly employ functional mixed effects models (FMEM) for fixed effect estimation and inference [8–10]. However, their underlying assumption of sparse and irregular sampling hinders effective modeling of fixed effects and longitudinal-functional correlation structures in densely and regularly sampled RMFD [11, 12]. To better capture longitudinal-functional interactions, bivariate fixed effects are preferable to univariate methods, enhancing understanding of underlying processes. While some studies propose bivariate models for symmetrical data (i.e., images), these often extend univariate approaches and can be computationally intensive [13–15]. Improving correlation structure is also essential for flexibility and universality. While the random effect component of FMEM incorporates longitudinal visits through a linear framework with additive assumptions [6, 12, 16], a four-dimensional covariance function that accounts for continuity along two domains with dense observation offers a more flexible alternative with minimal assumptions. Furthermore, many estimation procedures, including spline-based and functional principal component analysis (FPCA) methods, struggle with scalability even when encountering smaller sample sizes and functional sampling grids than our motivating data [4, 10]. This is primarily due

to the challenges of estimating the complex four-dimensional correlation structure and conducting fixed-effect inference. Although some techniques have been proposed to enhance computational efficiency [6, 8], the implementation in large-scale population studies remains challenging.

To address these issues, we propose a novel two-dimensional functional mixed-effect model (2dFMM) framework for RMFD. The 2dFMM incorporates two-dimensional fixed effects and subject-specific random intercepts, along with a nonparametric four-dimensional covariance structure in marginal representation. This flexible approach integrates longitudinal variation without relying on prespecified correlation assumptions, while maintaining optimality [11]. Model component estimation employs an efficient three-stage procedure that combines pointwise and smoothing techniques, ensuring both interpretability and fast computational performance. Additionally, statistical inference tools, including pointwise and simultaneous confidence bands derived from marginal decomposition, enhance the estimation of the four-dimensional covariance function. This framework leverages FPCA and basis splines to alleviate computational burdens.

The rest of this paper is organized as follows. In Section 2, we propose 2dFMM with an estimation and inference procedure. Asymptotic results of the proposed estimator are also provided. Extensive simulation studies are conducted in Section 3 to evaluate the performance of 2dFMM and compare it with existing approaches. Section 4 further applies 2dFMM to the Shanghai school adolescent study and Australia electricity demand data. Conclusion and discussion are in Section 5.

2 | Methods

2.1 | Two-Dimensional Functional Mixed-Effect Model

The functional response is denoted by $Y_i(s, t)$, the i th subject's profile at functional time $t \in \mathcal{T}$, repeatedly measured at longitudinal time $s \in \mathcal{S}$, where $i = 1, \dots, N$. We assume that the covariate-dependent mean function $\mu(s, t, \mathbf{X}_i)$ is the linear combination of P -dimensional time-invariant or time-varying covariates of interest $\mathbf{X}_i = (1, x_{i,1}(s, t), \dots, x_{i,P}(s, t))^T$ at time s and t , which is also known as a standard linear concurrent model. The proposed two-dimensional functional mixed-effect model (2dFMM) is

$$\begin{aligned} Y_i(s, t) &= \mu(s, t, \mathbf{X}_i) + \eta_i(s, t) + \epsilon_i(s, t) \\ &= \beta_0(s, t) + \sum_{p=1}^P x_{i,p}(s, t) \beta_p(s, t) + \eta_i(s, t) + \epsilon_i(s, t) \end{aligned} \quad (1)$$

where $\{\beta_0(s, t), \beta_1(s, t), \dots, \beta_P(s, t)\}$ are corresponding coefficient functions, $\eta_i(s, t)$ is a smooth bivariate random process with mean zero and four-dimensional covariance function $C(s, t; u, v)$, capturing the functional-longitudinal correlation structure, while random measurement error process $\epsilon_i(s, t)$ is mean zero with variance σ^2 and independent of $\eta_i(s, t)$. Thus, the functional response $Y_i(s, t)$ is *i.i.d.* Gaussian random process with mean $\mathbb{E}[Y_i(s, t)] = \mu(s, t, \mathbf{X}_i)$ and variance function of $Y_i(s, t)$ is $\sigma^2(s, t)$,

while for any $s \neq u$ and $t \neq v$, covariance function is $\sigma(s, t; u, v) = C(s, t; u, v) + \sigma^2 \mathbb{1}_{\{s=u, t=v\}}$, where $\mathbb{1}_{\{\cdot\}}$ is an indicator function.

The four-dimensional covariance function $C(s, t; u, v)$ removes the constraint of the linear framework in FMEM and allows a more flexible representation. Compared with FEME design, our model addresses the pre-specification difficulty: Whether using a random slope effect depends on hard-to-verify a priori assumptions about longitudinal correlation structure due to the discretization of longitudinal visits.

To ensure identifiability of univariate effects, for example, effects over the domain \mathcal{T} , we impose the identification condition $\mathbb{E}[\beta_p(s)] = 1$ and adopt the multiplicative decomposition $\beta_p(s, t) = \beta_p(s)\beta_p(t)$. This approach follows the common practice in nonparametric regression to resolve scaling ambiguities inherent in multiplicative structures [17–19]. The satisfaction of identification conditions implies $\beta_p(t) = \int_S \beta_p(s, t) f_S(s) ds$, where f_S is density function with bounded support. Hence, given an estimator of bivariate coefficient function $\hat{\beta}_p(s, t)$, the univariate function estimator can be obtained as $\hat{\beta}_p(t) = R^{-1} \sum_{r=1}^R \hat{\beta}_p(s_r, t)$ with the covariance estimator $\widehat{\text{Cov}}\{\hat{\beta}_p(t), \hat{\beta}_p(v)\} = R^{-2} \sum_{r_1} \sum_{r_2} \widehat{\text{Cov}}\{\hat{\beta}_p(s_{r_1}, t), \hat{\beta}_p(s_{r_2}, v)\}$.

2.2 | Dimension Reduction for Covariance Structure

Traditional approaches decompose $\eta_i(s, t)$ using two-dimensional FPCA via a Karhunen-Loève expansion with orthonormal $L^2(S \times \mathcal{T})$ bases [20]. While this method explains a large proportion of the four-dimensional covariance $C(s, t; u, v)$ using the top eigencomponents, it faces two critical limitations: Computational intractability and poor suitability for asymmetric bivariate structures commonly observed in real-world functional data (RMFD).

To address this issue, in Model (1), we decompose $\eta_i(s, t)$ by exploiting the technique of marginal covariance functions to separate the bivariate process for efficiency as follows [6, 11, 16],

$$\eta_i(s, t) = \sum_{j=1}^{\infty} \xi_{i,j}(t) \psi_j(s)$$

where $\psi_j(s)$ be eigenfunction of marginal covariance function, such that $C_S(s, u) = \sum_{j=1}^{\infty} \tau_j \psi_j(s) \psi_j(u)$, the marginal covariance with respect to S [21, 22]. The smooth random coefficient functions $\{\xi_{i,j}(t) : j \geq 1\}$ represents the projection of $\eta_i(\cdot, t)$ subject onto the direction $\psi_j(s)$, that is, $\langle \eta_i(\cdot, t), \psi_j \rangle_S$, ensuring $\mathbb{E}[\xi_{i,j}(t)] = 0$ for $t \in \mathcal{T}$ and $\mathbb{E}[\langle \xi_{i,j}, \xi_{i,h} \rangle_{\mathcal{T}}] = \tau_j \mathbb{1}_{\{j=h\}}$.

We further define the covariance function $\Theta_j(t, v) = \mathbb{E}[\xi_{i,j}(t) \xi_{i,j}(v)]$, which depends on the j th eigenfunction of marginal covariance $C_S(s, u)$. Therefore, for any $s \neq u$ and $t \neq v$, we can obtain an expression of the four-dimensional covariance function as follows,

$$\begin{aligned} C(s, t; u, v) &= \mathbb{E}[(Y_i(s, t) - \mu(s, t, \mathbf{X}_i))(Y_i(u, v) - \mu(u, v, \mathbf{X}_i))] \\ &= \sum_{j=1}^{\infty} \psi_j(s) \psi_j(u) \Theta_j(t, v) \end{aligned} \quad (2)$$

The four-dimensional covariance function in Equation (2) employs a decomposition that relaxes the strong separability assumption, which requires the covariance to factorize into purely spatial and temporal components, that is, $C(s, t; u, v) = C_S(s, u)C_{\mathcal{T}}(t, v)$ [21, 22]. Instead, our model assumes a component-wise separability: For each eigenfunction $\psi_j(s)$, the temporal covariance $\Theta_j(t, v)$ is allowed to vary across eigencomponents. This more flexible approach better captures real-world phenomena while maintaining computational feasibility, and can achieve near-optimal performance under appropriate conditions [11, 23].

2.3 | Estimation of Model Components

Suppose design points s_r and t_l satisfy densities in Remark 3, we let a pointwise response vector denoted by \mathbf{Y}_{s_r, t_l} and design matrix $\mathbf{X}_{s_r, t_l} = (\mathbf{X}_{1, s_r, t_l}^T, \dots, \mathbf{X}_{N, s_r, t_l}^T)^T$, where $\mathbf{X}_{i, s_r, t_l} = (1, x_{i,1}(s_r, t_l), \dots, x_{i,p}(s_r, t_l))^T$. We propose a computationally efficient three-stage estimation procedure as follows.

Step I: Bivariate Pointwise Estimation. We reform a linear model $\mathbf{Y}_{s_r, t_l} = \mathbf{X}_{s_r, t_l} \boldsymbol{\beta}_{s_r, t_l} + \mathbf{e}_{s_r, t_l}$, where $\boldsymbol{\beta}_{s_r, t_l} = (\beta_0(s_r, t_l), \dots, \beta_p(s_r, t_l))^T$ and $\mathbf{e}_{s_r, t_l} \sim N(\mathbf{0}, \sigma_{s_r, t_l}^2 \mathbf{I}_N)$, under the mutual independence assumption, which is commonly adopted in other fixed-effect estimation approach [7, 8, 24]. We use ordinary least squares estimator $\tilde{\boldsymbol{\beta}}_{s_r, t_l} = (\mathbf{X}_{s_r, t_l}^T \mathbf{X}_{s_r, t_l})^{-1} \mathbf{X}_{s_r, t_l}^T \mathbf{Y}_{s_r, t_l}$ for initial estimation because it shares nice statistical properties and low computational cost. Considering the bivariate pointwise conditions, the covariance matrix estimates of coefficient functions can be obtained for any $s_{r_1}, s_{r_2}, t_{l_1}, t_{l_2}$,

$$\text{Cov}\{\tilde{\boldsymbol{\beta}}_{s_{r_1}, t_{l_1}}, \tilde{\boldsymbol{\beta}}_{s_{r_2}, t_{l_2}}\} = \sigma_{s_{r_1}, t_{l_1}; s_{r_2}, t_{l_2}} \mathbf{H}_{s_{r_1}, t_{l_1}} \mathbf{H}_{s_{r_2}, t_{l_2}}^T \quad (3)$$

where $\mathbf{H}_{s_r, t_l} = (\mathbf{X}_{s_r, t_l}^T \mathbf{X}_{s_r, t_l})^{-1} \mathbf{X}_{s_r, t_l}^T$ and the estimates for $\sigma_{s_{r_1}, t_{l_1}; s_{r_2}, t_{l_2}}$ will be provided in step III later. Note that for pairwise observation (s_r, t_l) , the standard error estimate of the linear model which is denoted by $\tilde{\sigma}_{s_r, t_l}^2$ is the raw estimate to σ_{s_r, t_l}^2 .

Step II: Bivariate Smoothing. As the raw estimate is obtained, bivariate smoothing is required to refine it by integrating neighboring temporal information. Here, we illustrate the use of the sandwich smoother due to its computational efficiency with nice asymptotic properties. Denote the raw p th estimated bivariate coefficient function by matrix $\tilde{\boldsymbol{\beta}}_p = (\tilde{\beta}_p(s_r, t_l))_{R \times L}$, the refined estimator is expressed by $\hat{\boldsymbol{\beta}}_p = \mathbf{S}_2 \tilde{\boldsymbol{\beta}}_p \mathbf{S}_1$, where \mathbf{S}_1 and \mathbf{S}_2 are smoother matrices for \mathcal{T} and S respectively, utilizing P-splines in different prespecified number of knots K_R and K_L . The procedure is also known as the bivariate pointwise and post-smoothing (pointwise-smoothing) estimation.

Additionally, with the help of the bivariate smoother, the variability of the covariance estimator $\tilde{\sigma}_{s_r, t_l}^2$ can be further diminished. Applying the sandwich smoother on standard error matrix $\tilde{\mathbf{R}} = (\tilde{\sigma}_{s_r, t_l}^2)_{R \times L}$ gives the final covariance estimator $\hat{\mathbf{R}} = (\hat{\sigma}_{s_r, t_l}^2)_{R \times L}$. However, there is no guarantee that the resulting estimators are all non-negative. The issue can be handled by trimming the negative values at zero [8, 25, 26]. In this study, we also provide tensor product smooths in the practical implementation of bivariate smoothing. The performance of two smoothers will be shown in detail in simulation studies.

Step III: Covariance Estimation. To reduce the computing burden of covariance estimation, we employ FPCA on marginal covariance and B-splines to approximate score functions under marginal representation, tailored to dimensional imbalance between domains. The efficient estimation procedure also consists of three stages.

Firstly, we use the centered data, $\tilde{Y}_i(s_r, t_l) = Y_i(s_r, t_l) - \hat{\mu}(s_r, t_l, \mathbf{X}_i)$, to estimate the marginal covariance function $C_S(s, u)$ and obtain the estimates of semi-definite covariance function estimates $\hat{C}_S(s_{r_1}, s_{r_2})$, eigenfunctions $\psi_j(s)$ and score functions $\xi_{i,j}(t)$ by FPCA. Specifically, we pool $\{\tilde{Y}_i(\cdot, t_l), i = 1, \dots, N, l = 1, \dots, L\}$ and have sample covariance $\hat{C}_S(s_{r_1}, s_{r_2}) = (|\mathcal{T}|/NL) \sum_{i=1}^N \sum_{l=1}^L \tilde{Y}_i(s_{r_1}, t_l) \tilde{Y}_i(s_{r_2}, t_l)$, where $1 \leq r_1 \leq r_2 \leq R$. Because the number of longitudinal grids is relatively small, we obtain eigenfunctions estimates $\{\hat{\psi}_j(s) : j = 1, \dots, J\}$ and score function estimates $\hat{\xi}_{i,j}(t_l) = \int_S \tilde{Y}_i(s, t_l) \hat{\psi}_j(s) ds$ by kernel-based local linear smoothing in PACE algorithm [26]. The number of components J is determined by a fraction of variance explained (FVE), the threshold of which is set to 0.99.

Secondly, we estimate the marginal covariance functions $\Theta_j(t, v)$ by “observed” functional data $\tilde{\xi}_{i,j}(t)$. Suppose each functional data has B-splines basis expansion $\tilde{\xi}_{i,j}(t_l) = \mathbf{B}_j^T(t_l) \mathbf{b}_{i,j}$, where $\mathbf{B}_j(t_l) = (B_{j1}(t_l), \dots, B_{jK}(t_l))^T$ and $\mathbf{b}_{i,j} = (b_{i,j1}, \dots, b_{i,jK})^T$, $B_{jk}(t_l)$ is the k th B-splines basis function of the j th principal component, K is the number of basis functions. Let K be the same for all j ; the basis functions $\mathbf{B}(t_l)$ do not rely on j . The covariance estimator of $\tilde{\xi}_{i,j}(t_l)$, denoted by $\hat{\Theta}_j(t_{l_1}, t_{l_2})$, can be obtained nonparametrically

$$\begin{aligned} \hat{\Theta}_j(t_{l_1}, t_{l_2}) &= N^{-1} \sum_{i=1}^N \hat{\xi}_{i,j}(t_{l_1}) \hat{\xi}_{i,j}(t_{l_2}) \\ &= N^{-1} \sum_{i=1}^N \mathbf{B}^T(t_{l_1}) (\hat{\mathbf{b}}_{i,j}^T \otimes \hat{\mathbf{b}}_{i,j}) \mathbf{B}(t_{l_2}) \end{aligned}$$

The choice of the value of K depends on a tradeoff between capturing variations adequately and ensuring computational efficiency. To ensure that the majority of variations are captured by a large enough number of basis functions, we also consider the computational efficiency of the basis function expansion. Compared with FPCA, which has the computational complexity $O(NL^2 + L^3)$ for each $\tilde{\xi}_{i,j}(t)$ [27, 28], the usage of B-splines requires $O(NLK^2)$ computations. It implies that when $K < (L + L^2/N)^{1/2}$, our approach offers a lighter computational cost. Naturally, we can set $K = K_L$ and $K_L = \min\{L/2, 35\}$ is recommended by [29]. Thus, in practical usage, we suggest $K = \min\{(L + L^2/N)^{1/2}, L/2, 35\}$.

Finally, given Equation (2), we obtain the estimator of four-dimensional covariance function $C(s_{r_1}, t_{l_1}; s_{r_2}, t_{l_2})$ as follows.

$$\hat{C}(s_{r_1}, t_{l_1}; s_{r_2}, t_{l_2}) = N^{-1} \mathbf{B}^T(t_{l_1}) \left\{ \sum_{j=1}^J \hat{\psi}_j(s_{r_1}) \hat{\psi}_j(s_{r_2}) \sum_{i=1}^N \hat{\mathbf{b}}_{i,j}^T \otimes \hat{\mathbf{b}}_{i,j} \right\} \mathbf{B}(t_{l_2}) \quad (4)$$

and therefore, $\hat{\sigma}_{s_{r_1}, t_{l_1}; s_{r_2}, t_{l_2}} = \hat{C}(s_{r_1}, t_{l_1}; s_{r_2}, t_{l_2}) + \hat{\sigma}_{s_{r_1}, t_{l_1}}^2 \mathbb{1}_{\{s_{r_1} = s_{r_2}, t_{l_1} = t_{l_2}\}}$.

The estimation with randomly sampled designs is discussed in Section 5. Our approach also lends itself well to parallelization, further accelerating the entire process.

2.4 | Inference Procedure

Here we show the construction of pointwise and simultaneous confidence bands. Let $\tilde{\beta}_p = \text{vec}(\tilde{\beta}_p)$ indicate a matrix is stacked by column and $\widehat{\text{Cov}}\{\tilde{\beta}_p, \tilde{\beta}_p\}$ be a $RL \times RL$ covariance matrix estimates of the p th bivariate coefficient function constructed by the p th diagonal element of $\widehat{\text{Cov}}\{\tilde{\beta}_{s_{r_1}, t_{l_1}}, \tilde{\beta}_{s_{r_2}, t_{l_2}}\}$ via Equations (3) and (4), for all $s_{r_1}, s_{r_2}, t_{l_1}, t_{l_2}$. By tensor product properties $\hat{\beta}_p = (\mathbf{S}_1 \otimes \mathbf{S}_2) \tilde{\beta}_p$, the ultimate four-dimensional covariance estimator of the p th bivariate coefficient function is

$$\widehat{\text{Var}}\{\hat{\beta}_p(s_r, t_l)\} = \mathbf{e}_{s_r, t_l}^T (\mathbf{S}_1 \otimes \mathbf{S}_2) \widehat{\text{Cov}}\{\tilde{\beta}_p, \tilde{\beta}_p\} (\mathbf{S}_1 \otimes \mathbf{S}_2)^T \mathbf{e}_{s_r, t_l} \quad (5)$$

where \mathbf{e}_{s_r, t_l} denotes RL -dimensional unit vector with 1 at the $(l-1)L + r$ th entry. The covariance estimator of the refined coefficient is coincidentally equivalent to the covariance smoothing approach [30, 31].

The analytic inference for two-dimensional fixed effects using confidence bands is straightforward. To construct confidence bands for the two-dimensional fixed effects, we estimate the pointwise variability at each grid point (s_r, t_l) using Equation (5) for practical inference. The 95% pointwise confidence bands (PCBs) are then defined as:

$$\hat{\beta}_p(s_r, t_l) \pm 2\widehat{\text{Var}}\{\hat{\beta}_p(s_r, t_l)\}^{1/2}$$

While PCBs are computationally efficient, they ignore the inherent correlations across functional domains and inflate false positive rates [8, 32].

To address this issue, we employ simultaneous confidence bands (SCBs) constructed via nonparametric bootstrap of the maximum absolute statistic [33]. This bootstrap method is generally applicable to SCBs in functional data analysis, which does not require parametric assumptions [8, 32, 34, 35]. Specifically, resampling entire subjects—also known as the subject-level bootstrap algorithm—is well-suited for our model, as it preserves within-subject functional-longitudinal correlations while respecting between-subject independence. This approach avoids restrictive assumptions about the error process $\epsilon_i(s, t)$ and aligns with our flexible covariance framework. Algorithm 1 details our implementation.

In the algorithm, when generating $\hat{\beta}_{m,p}(s_r, t_l)$ for computing the maximum absolute statistic, we reduce the computational complexity of multivariate normal simulations by leveraging FPCA and B-spline basis expansions [34]. Instead of simulating directly in the high-dimensional functional domain, we generate parameters in a lower-dimensional spline space. This dimension reduction ensures tractable computation while maintaining accuracy. The number of components J_B and the number of B-spline basis functions K_B are determined using criteria analogous to those for selecting J and K .

- 1: **for** $b = 1, \dots, B$ **do**
- 2: Resample the subject indexes from the index set $\{1, \dots, N\}$ with replacement and define I_b be the set of indices;
- 3: Denote the b th bootstrap sample as $\{Y_{I_b}(s_r, t_l), \mathbf{X}_{I_b}(s_r, t_l)\}$;
- 4: Use Step I and Step II of the estimation procedure to obtain $\hat{\beta}_p^{(b)}(s_r, t_l)$;
- 5: **end for**
- 6: Perform FPCA and marginal decomposition technique on $\{\hat{\beta}_p^{(1)}(s_r, t_l), \dots, \hat{\beta}_p^{(B)}(s_r, t_l)\}$ to obtain $\{\hat{\psi}_j(s_r), \hat{\mathbf{b}}_j, j = 1, \dots, J_B\}$, where $\hat{\psi}_j$ is the j th eigenfunction function and $\hat{\mathbf{b}}_j$ is the B-spline coefficient vector of length K_B . Derive the mean function $\bar{\beta}_p(s_r, t_l)$ and obtain the sample covariance of $\hat{\mathbf{b}}_j$, denoted by $\bar{\mathbf{V}}_j$;
- 7: **for** $m = 1, \dots, M$ **do**
- 8: Generate a random variable $\mathbf{u}_{m,j}$ from the multivariate normal with mean $\mathbf{0}$ and covariance matrix $\bar{\mathbf{V}}_j$;
- 9: Derive $\hat{\beta}_{m,p}(s_r, t_l) = \bar{\beta}_p(s_r, t_l) + \sum_{j=1}^{J_B} \mathbf{B}_j^T(t_l) \mathbf{u}_{m,j} \hat{\psi}_j(s_r)$;
- 10: Compute $q_m^* = \max_{s_r, t_l} \{|\hat{\beta}_{m,p}(s_r, t_l) - \hat{\beta}_p(s_r, t_l)| / \sqrt{\widehat{\text{Var}}\{\hat{\beta}_p(s_r, t_l)\}}^{1/2}\}$;
- 11: **end for**
- 12: Obtain the $100(1 - \alpha)\%$ empirical quantile of $\{q_1^*, \dots, q_M^*\}$, denoted by $q_{1-\alpha}$;
- 13: The $100(1 - \alpha)\%$ simultaneous confidence bands are given by $\hat{\beta}_p(s_r, t_l) \pm q_{1-\alpha} \sqrt{\widehat{\text{Var}}\{\hat{\beta}_p(s_r, t_l)\}}^{1/2}$.

2.5 | Asymptotic Results

In this section, we derive the asymptotic distribution of our pointwise-smoothing estimator by showing the asymptotic bias and variance structure. Asymptotics of our estimator is established based on the properties of the least square estimator and sandwich smoother, which is equivalent to a bivariate kernel regression estimator with a product kernel, $(RLh_R h_L)^{-1} \sum_{r,l} \beta_p(s_r, t_l) H_{m_R} \{h_R^{-1}(s - s_r)\} H_{m_L} \{h_L^{-1}(t - t_l)\}$, where H_m is the equivalent kernel for univariate penalized splines, h_R and h_L are the bandwidths [29, 36]. The kernel function H_m is symmetric and bounded. For simplicity, our results are for the case of equally spaced design points and knots. For notation convenience, $a \sim b$ means a/b converges to 1. Some conditions are required and stated in the [Supporting Information](#).

We first derive the asymptotic bias in the interior points. Let m_R and m_L are difference orders of differencing matrices, $m_T = 4m_R m_L + m_R + m_L$ for notation simplicity.

Proposition 1. *Suppose conditions (a)–(d) and (g) are satisfied, further assume $K_R \sim c_R(RL)^{b_1}$, $K_L \sim c_L(RL)^{b_2}$, with $b_1 > (m_R + 1)m_L/m_T$, $b_2 > (m_L + 1)m_R/m_T$, $h_R \sim d_R(RL)^{-m_L/m_T}$, $h_L \sim d_L(RL)^{-m_R/m_T}$ for some positive constants c_R, c_L, d_R, d_L . Then, for any $(s, t) \in (0, 1)^2$, we have*

$$\begin{aligned} \text{bias}\{\hat{\beta}_p(s, t)\} &= (-1)^{m_R+1} d_R^{2m_R} \frac{\partial^{2m_R}}{\partial s^{2m_R}} \beta_p(s, t) \\ &\quad + (-1)^{m_L+1} d_L^{2m_L} \frac{\partial^{2m_L}}{\partial t^{2m_L}} \beta_p(s, t) + o(h_L^{2m_L}) \end{aligned}$$

The bias remains same with the sandwich smoother [29], containing components from \mathcal{S} and \mathcal{T} . Noted that the bias converges at a slower rate at the boundary than in the interior [37], the proof of which is ignored here.

To derive the asymptotic variance of the estimator, we assume the covariates are identically and independently distributed as well as

time-invariant. For simplicity of illustration, we also assume no missing values.

Proposition 2. *Suppose conditions (a)–(g) are satisfied and under the conditions of Proposition 1, when $N \rightarrow \infty$, we have*

$$\text{var}\{\hat{\beta}_p(s, t)\} = 2(RLh_R h_L N)^{-1} \omega_p \sigma^2(s, t) \kappa(H_{m_R}) \kappa(H_{m_L}) + o(h_L^{4m_L})$$

where ω_p is the (p, p) th entry of $\mathbb{E}(\mathbf{X}_{1,s_r,t_l} \mathbf{X}_{1,s_r,t_l}^T) \rightarrow \Omega$ and $\kappa(H_m) = \int H_m^2(u) du$. The proposition implies that the asymptotic variance structure of our estimator has an extra component because of the existence of a pointwise least square estimator, compared to the sandwich smoother only [29]. Additionally, it also shows that the correlation influence of two points can be ignored, similarly with kernel regression estimator [38].

Based on the asymptotic bias and variance structures, and Proposition 1 in (Xiao et al.) [29], the corresponding asymptotic distribution of our estimator is given by

$$\begin{aligned} (N h_R h_L)^{1/2} (RL)^{(4m_R m_L + m_T)/2m_T} (\hat{\beta}_p(s, t) \\ - \beta_p(s, t)) \rightarrow N(\alpha_p(s, t), V_p(s, t)) \end{aligned}$$

in distribution as $R \rightarrow \infty$ and $L \rightarrow \infty$, $N \rightarrow \infty$, where $\alpha_p(s, t) = (-1)^{m_R+1} d_R^{2m_R} \frac{\partial^{2m_R}}{\partial s^{2m_R}} \beta_p(s, t) + (-1)^{m_L+1} d_L^{2m_L} \frac{\partial^{2m_L}}{\partial t^{2m_L}} \beta_p(s, t)$ and $V_p(s, t) = 2\omega_p \sigma^2(s, t) \kappa(H_{m_R}) \kappa(H_{m_L})$.

3 | Simulation

We conduct extensive simulation studies to evaluate the performance of the proposed estimation and inference procedure. The bivariate functional model is simulated as follows:

$$\begin{aligned} Y_i(s, t) &= \beta_0(s, t) + X_i(s) \beta_1(s, t) + \gamma_{i0}(t) + z_i(s) \gamma_{i1}(t) \\ &\quad + \epsilon_i(s, t), \quad (s, t) \in [0, 1]^2 \end{aligned}$$

The fixed-effect covariates are generated from $X_i(s) \sim N(0, 4)$ and $z_i(s) = 6(s - 0.5)^2 + N(0, \rho^2)$, where ρ represents how noisy the signal of repeatedly measured visits is. For bivariate coefficient functions, we take into account two different types, as shown in Figure S1 of the Supporting Information. The first scenario (S1) presents a continuous non-differentiable bivariate function with local zero regions (sparse), while the second scenario (S2) presents a smooth bivariate function. The simulations of random effect and measurement error are also provided in the Supporting Information.

The performance of the method is evaluated from three aspects reflecting the accuracy of estimation and inference, as well as computational efficiency. First, the estimation error is assessed by integrated squared error (ISE) to measure the difference between the estimate and the underlying truth, defined as $ISE(\hat{\beta}_p) = |S \times \mathcal{T}|^{-1} \int_S \int_{\mathcal{T}} (\hat{\beta}_p(s, t) - \beta_p(s, t))^2 dt$, where $|S \times \mathcal{T}|$ denotes the area of the entire domain. Secondly, the proportion of pointwise surfaces wrapping the true plane in the sandwich form is computed for the bivariate functional slope to evaluate inferential performance on fine grids. We use the empirical coverage probability of 95% PCB, defined as $|S \times \mathcal{T}|^{-1} \int_S \int_{\mathcal{T}} \mathbb{1}_{\hat{\beta}_p(s, t) \in \text{PCB}_p(s, t)} ds dt$. A similar approach is used

for SCB. Additionally, to measure the width of confidence bands we also report integrated actual width (IAW), defined as $|S \times \mathcal{T}|^{-1} \int_S \int_{\mathcal{T}} \{\widehat{\text{UB}}_p(s, t) - \widehat{\text{LB}}_p(s, t)\} ds dt$, where $\widehat{\text{UB}}_p(\cdot, \cdot)$ and $\widehat{\text{LB}}_p(\cdot, \cdot)$ are pointwise estimates of upper bound and lower bound respectively.

3.1 | Bivariate Comparison

We compare our proposed approach (2dFMM) with the concurrent bivariate functional regression method denoted by 2dGAM, which is developed by tensor product smooths [13]. The following several simulation scenarios are considered. We let sample size $N \in \{50, 75, 100\}$, the number of functional grids $L \in \{100, 150, 200\}$, and the number of longitudinal grids $R \in \{10, 15, 20\}$. The baseline setting is $N = 50, R = 10$, and $L = 100$, where all other sample-generating parameters are fixed at their baseline values when one is changed. The noise argument of the longitudinal signal is set to $\rho = 0.5$. A total of 100 replicates is independently simulated.

Figure 2 demonstrates that the proposed 2dFMM generally outperforms 2dGAM in estimation accuracy under the first

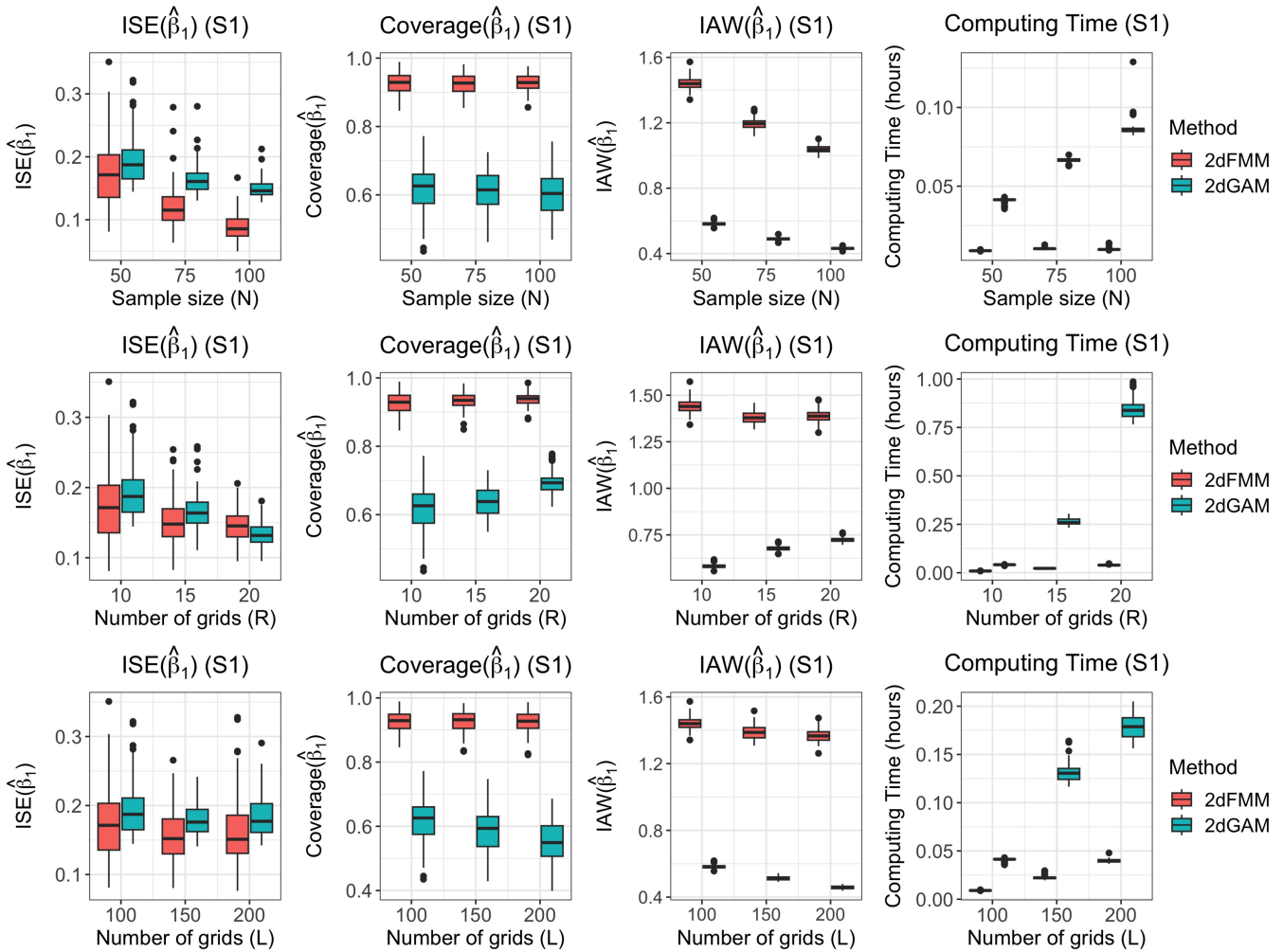


FIGURE 2 | The comparison of ISE, coverage probability of 95% PCB (Coverage), IAW, and computing time for 2dFMM and 2dGAM under the first scenario (S1) of $\beta_1(s, t)$. The baseline setting is $N = 50, R = 10$, and $L = 100$. When one parameter is changed, all other sample-generating parameters are fixed at their baseline values.

scenario. This is due to the sandwich smoother's ability to handle non-differentiable functions. However, as the dimensional difference between longitudinal and functional domains decreases, 2dGAM slightly surpasses 2dFMM because tensor product spline basis expansions suit symmetric grids. The 95% PCB coverage probability for 2dFMM is close to the nominal level, whereas 2dGAM falls short due to ignoring the four-dimensional correlation structure, resulting in narrower confidence bands. Additionally, 2dGAM's reliance on the generalized additive model representation leads to significantly longer computation times and storage issues with large matrices, often causing memory shortages even with moderate sample sizes.

Figure 3 displays the results of the second scenario of the slope coefficient function, while we use tensor product smooths as a post-smoother instead. The comparable performance of the two methods regarding estimation results indicates that the smoother is sufficient to compensate for the violation of the independence assumption underneath the pointwise technique when encountering continuous functions. The choice of the tensor product smooths also solves the problem of the symmetric numbers of

longitudinal and functional grids that the sandwich smoother has. Despite similar estimation accuracy, our method is still plausible given the nice inferential behaviors and low computational cost.

We evaluate the empirical coverage probabilities of 95% PCB and SCB under different scenarios, with results shown in Table 1. As expected, SCB exhibits wider intervals and greater robustness compared to PCB, as its construction accounts for the inherent correlation structure in functional data, thereby reducing false positive rates.

3.2 | Univariate Comparison

Our method is examined not only in bivariate but also in univariate perspectives, as other competing FMEM estimation methods often focus solely on fixed effects over the functional domain. We compare our approach with established methods, including fast univariate inference (FUI) and fixed-effect inference for longitudinal functional data (FILF), which accounts for between-visit correlations [6, 8].

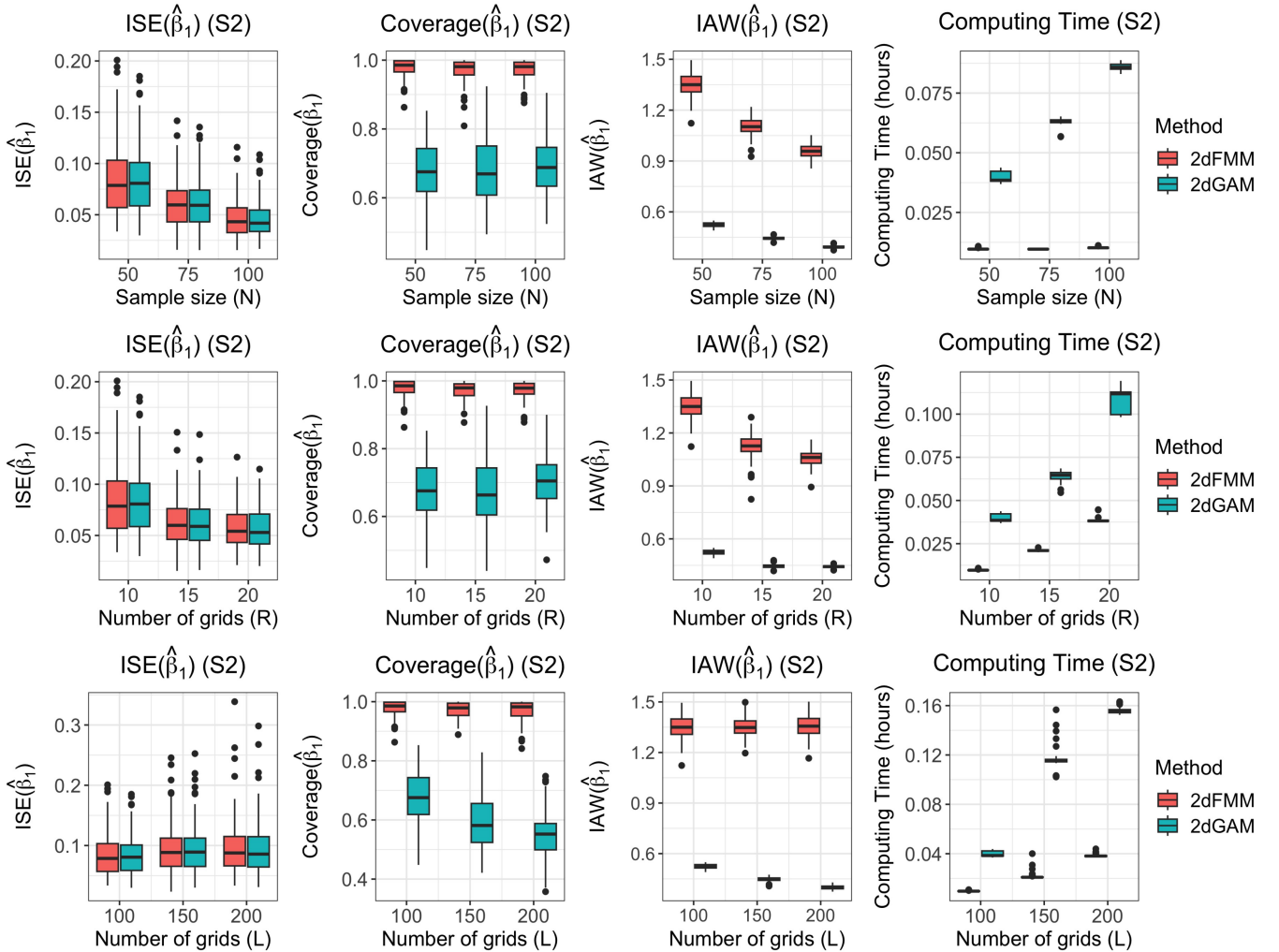


FIGURE 3 | The comparison of ISE, coverage probability of 95% PCB (Coverage), IAW, and computing time for 2dFMM and 2dGAM under the second scenario (S2) of $\beta_1(s, t)$. The baseline setting is $N = 50$, $R = 10$, and $L = 100$. When one parameter is changed, all other sample-generating parameters are fixed at their baseline values.

TABLE 1 | The average empirical coverage of 95% PCB and SCB under **S1** and **S2** with varying sample parameters among 100 simulation replicates. The value in the parentheses corresponds to IAW. The baseline setting is $N = 50$, $R = 10$, and $L = 100$. When one parameter is changed, all other sample-generating parameters are fixed at their baseline values.

		Sample size (N)		
Type	95% CB	50	75	100
S1	PCB	0.93 (1.45)	0.93 (1.19)	0.93 (1.03)
	SCB	0.95 (1.62)	0.95 (1.33)	0.95 (1.14)
S2	PCB	0.97 (1.36)	0.97 (1.11)	0.96 (0.96)
	SCB	0.97 (1.42)	0.97 (1.15)	0.97 (0.99)
Number of grids (R)				
		10	15	20
S1	PCB	0.93 (1.45)	0.93 (1.38)	0.94 (1.39)
	SCB	0.95 (1.62)	0.96 (1.58)	0.97 (1.59)
S2	PCB	0.97 (1.36)	0.97 (1.12)	0.97 (1.05)
	SCB	0.97 (1.42)	0.97 (1.18)	0.97 (1.09)
Number of grids (L)				
		100	150	200
S1	PCB	0.93 (1.45)	0.93 (1.39)	0.93 (1.36)
	SCB	0.95 (1.62)	0.96 (1.58)	0.95 (1.53)
S2	PCB	0.97 (1.36)	0.97 (1.36)	0.96 (1.35)
	SCB	0.97 (1.42)	0.98 (1.44)	0.97 (1.42)

In longitudinal functional data analysis, fixed effects in the FMEM framework are often evaluated only over the functional domain. Several methods, including functional additive mixed models (FAMM), fast univariate inference (FUI), and fixed-effect inference for longitudinal functional data (FILF), allowing between-visit correlations, are considered for comparisons, while FUI and FILF are designed for simpler computation [6, 8, 39]. They are prespecified to incorporate random slope covariates using longitudinal time points. However, the performance of FAMM is not shown because it is similar to FUI while taking dramatically longer computing time and narrower confidence bands [6, 8].

All approaches are evaluated in two cases: (i) Bivariate functional slope is retained in the true model, but only the marginal effect over the functional domain is examined. (ii) Bivariate functional intercept and slope are shrunk to univariate ones in the true model, that is, $\beta_p(t) = |S|^{-1} \int_S \beta_p(s, t) ds, p = 0, 1$. Let the noise argument $\rho \in \{0.5, 2, 6\}$ control the magnitude of the longitudinal correlation in the true model. A total of 100 replicates is independently simulated. When comparing with other methods for FMEM from the univariate perspective, we accommodate all metrics with only \mathcal{T} direction.

Figure 4a presents that, under case (i), 2dFMM outperforms other methods in estimation accuracy across all noise levels due to its two-dimensional effects. FMEM methods perform worse with increasing noise, indicating a lack of longitudinal correlation and resulting in functional random slope model misspecification. For confidence bands, 2dFMM maintains robust coverage and width, while FILF achieves high coverage with wider bands. In contrast, Figure 4b highlights 2dFMM’s disadvantage under case (ii), where strong longitudinal correlation and univariate effects

align with the FMEM framework. The lower estimation accuracy is due to misinterpreting strong random longitudinal signals as fixed effects. Consequently, 2dFMM performs better with heavier longitudinal noise, while others decline due to misspecification. Despite this, 2dFMM still achieves decent coverage close to the nominal level, though slightly conservative in unfavorable conditions. These patterns persist with larger sample generating parameters, which are shown in Figure S2–S4. Comparisons of computing time for the entire estimation procedure are given in Table S1 in the Supporting Information.

4 | Application

In this section, we apply our proposed method to two studies for illustration. The first study uses motivating accelerometer data to examine the intraday and interday dynamic associations between adolescents’ physical activity and their demographic characteristics, family socioeconomic status, and physical and mental health assessments. The second study uses a public environmental dataset to assess the association between electricity demand and temperature. To present the statistical inferences of bivariate coefficient functions, we define a new metric $\hat{I}_p(s, t) = \hat{\beta}_p(s, t) \mathbb{1}_{\{\widehat{LB}_p(s, t) > 0 \text{ or } \widehat{UB}_p(s, t) < 0\}}$ to quantify and interpret the dynamics of the associations. Heatmaps of the significance evaluation $\hat{I}_p(s, t)$ will be displayed.

4.1 | Application to Shanghai School Adolescent Physical Activity Data

In the Shanghai school adolescent study, we collected each subject’s demographic information, including a binary indicator of

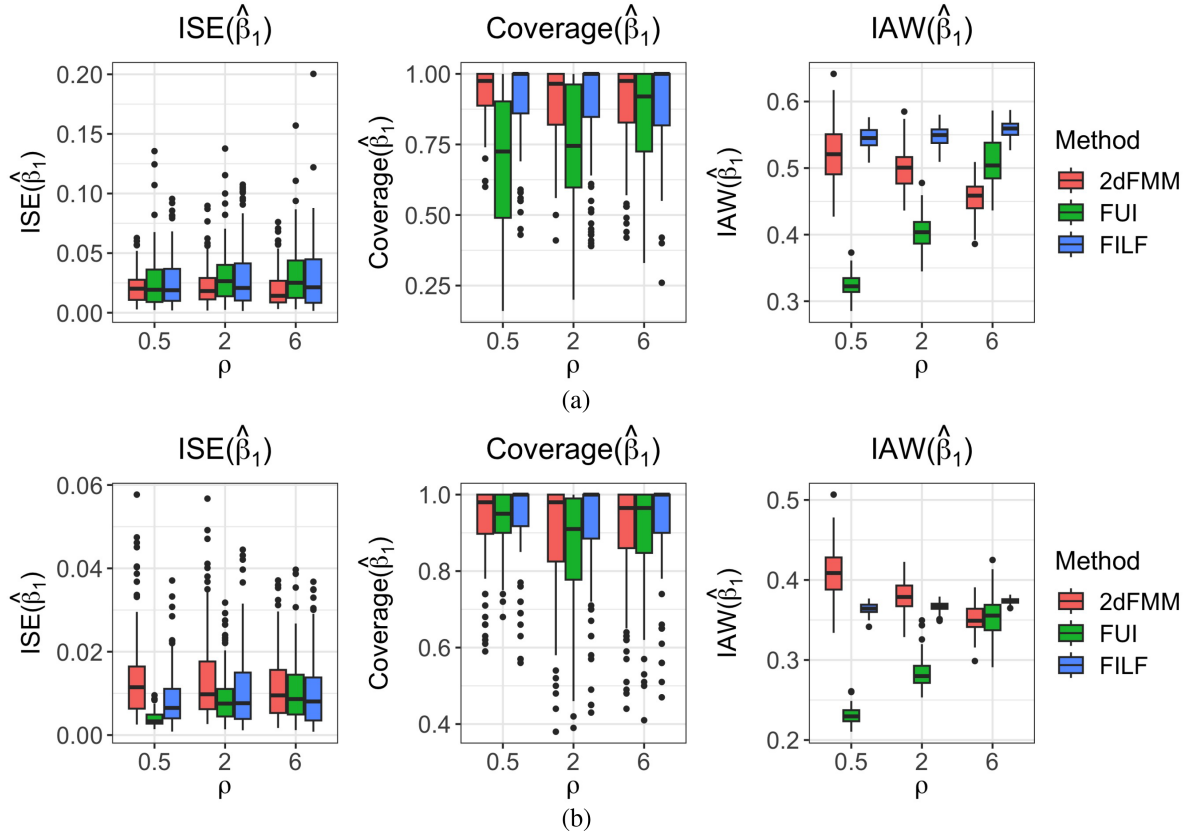


FIGURE 4 | The comparisons of ISE, coverage probability of 95% PCB (Coverage), and IAW for FUI, FILF, and our method with reducing to functional direction with the setting $N = 50$, $R = 10$, and $L = 100$. (a) Case (i): The bivariate functional slope $\beta_1(s, t)$ under the second scenario (S2) is retained in the true model, (b) Case (ii): The bivariate functional slope $\beta_1(s, t)$ is shrunk to univariate $\beta_1(t)$ under the second scenario (S2) in the true model.

gender, grade from 7th to 12th, annual family income level, and mother's education level. Annual family income contains 7 levels from “ ≤ 10 k” to “ ≥ 300 k”, while mother's education level includes 8 levels from “not graduated from primary school” to “at least master's degree”, both of which are treated as ordinal variables. Mental health screening questionnaires were also conducted during the measurement periods. All health assessment results are numeric to present the subjective psychological conditions within different aspects, such as anxiety disorder and happiness. We first regress the subjects' profiles on demographic and socioeconomic covariates in the baseline model. To avoid the collinearity of mental health outcomes, each covariate is added to the baseline model separately. In this study, activity profiles were analyzed over a standardized Monday-Sunday sequence to analyze longitudinal trends as calendar day-of-week effects. The functional response consists of a 16191×1440 dimensional matrix, where each row corresponds to “day of week” and each column corresponds to “time of day”.

Figure 5 presents the baseline model results, incorporating only demographic and socioeconomic data. The heatmap of the estimated intercept function aligns with overall student physical activity patterns. Girls tend to sleep later and wake earlier, while boys are more active during the day. Interestingly, girls show more intense activity on Saturday night and Sunday. Grade, indicative of age and school workload, shows that older students are more sedentary, except at 6 a.m., 6 p.m., and midnight, likely

due to longer study hours. Among socioeconomic factors, family income has a weaker effect than the mother's education level. Higher maternal education correlates with students waking up and going to bed later, possibly due to more demanding assignments or better material well-being, such as private car commutes and more evening entertainment options. Additional univariate effects for weekdays and weekends are shown in Figure S5 in the [Supporting Information](#).

Figure 6 shows inferential heatmaps for self-reported health assessments from well-known questionnaires. Physical activity patterns are similar for depression, anxiety, and stress (DASS) scores. Students with severe symptoms are more active at midnight on weekends but more sedentary at 6 p.m. after school or on weekends, likely due to reluctance to exercise. The Anticipatory and Consummatory Interpersonal Pleasure Scale (AICPS) indicates that students with higher scores are more active during the day, especially on weekends, suggesting that happier students are more motivated and energetic. Emotion regulation strategies (ERQ) measure the tendency to adjust emotions, including cognitive reappraisal (ERQ(CR)) and expressive suppression (ERQ(ES)). Heavy reliance on expressive suppression significantly reduces physical activity on weekends, while cognitive reappraisal mainly affects behavior during weekend daytime. Additional univariate effects for health assessments by time of day for weekdays and weekends are shown in Figure S6 in the [Supporting Information](#).

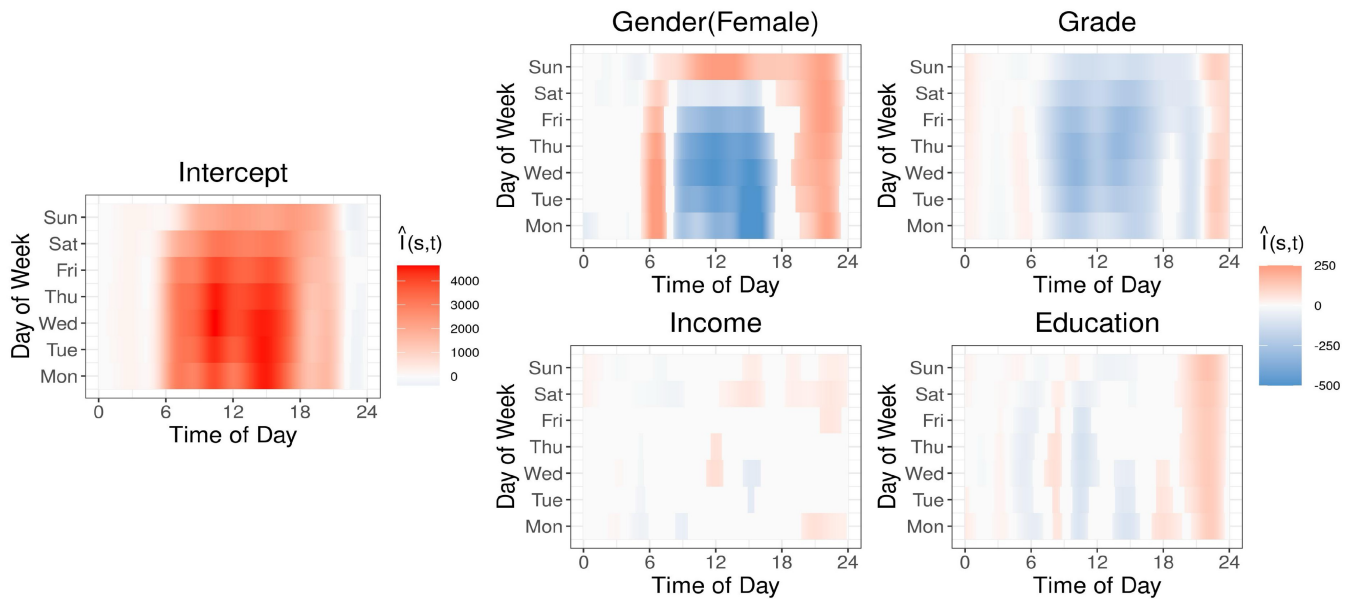


FIGURE 5 | Statistical inference evaluation $\hat{I}_p(s, t)$ of intercept, demographic and socioeconomic covariates. White regions indicate no significant effects, red indicates significantly positive effects, and blue indicates significantly negative effects, while the darkness of the colors represents the magnitude of the effects.

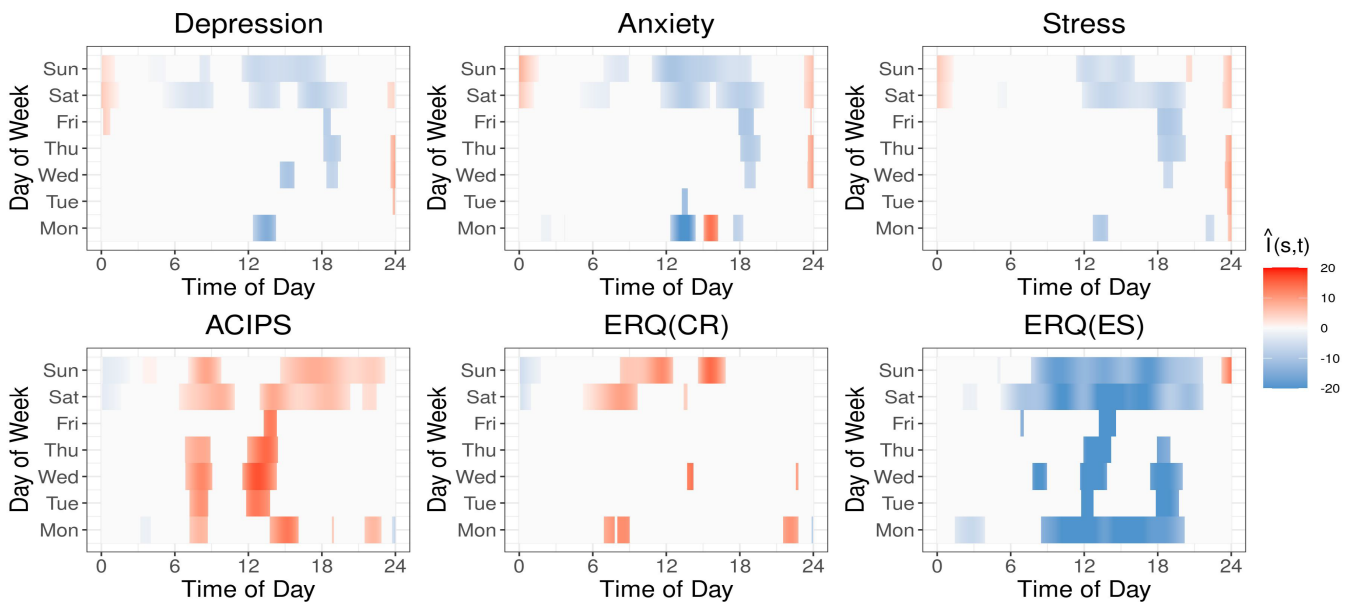


FIGURE 6 | Statistical inference evaluation $\hat{I}_p(s, t)$ of mental health assessment results: DASS including Depression, Anxiety, and Stress, ACIPS (an interpersonal pleasure scale), ERQ scores of cognitive reappraisal (ERQ(CR)) and expressive suppression (ERQ(ES)). White regions indicate no significant effects, red indicates significantly positive effects, and blue indicates significantly negative effects, while the darkness of the colors represents the magnitude of the effects.

4.2 | Application to Electricity Demand Data

In Adelaide, Australia, summer electricity demand is highly volatile and strongly correlated with temperature, as shown in various studies [13, 40]. The data, available through the R `feds` package [41], covers July 6, 1997, to March 31, 2007, with half-hourly recordings for each day. Our study examines the association between electricity demand and temperature, and the impact of weekends, using our proposed model. The

dataset includes 63 two-dimensional samples, each representing a weekday of a year, such as Monday of 1998. The response variable is electricity demand, measured in megawatts, recorded at half-hourly intervals within each week of the year, resulting in 48 functional grids per day and 52 longitudinal grids per year.

Figure 7 shows inferential heatmaps for intercept, temperature, and weekend effects. The intercept indicates higher electricity

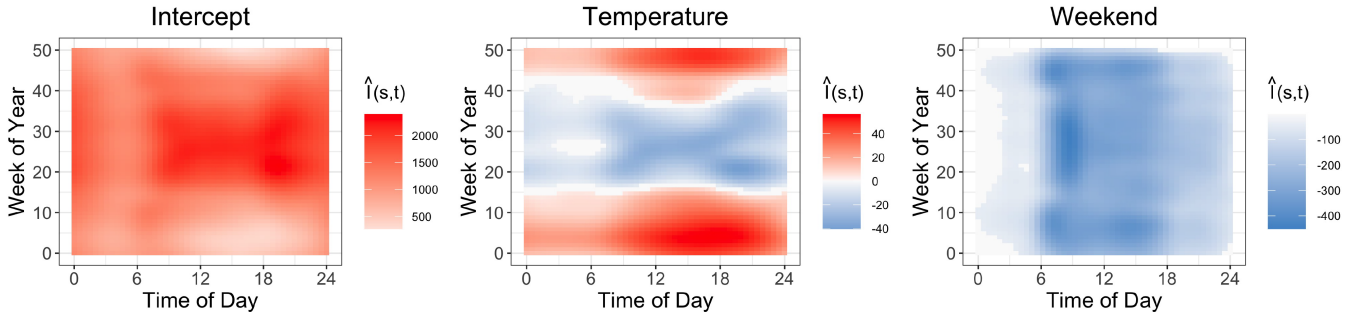


FIGURE 7 | Statistical inference evaluation $\hat{I}_p(s, t)$ of intercept, temperature, and weekend. White regions indicate no significant effects, red indicates significantly positive effects, and blue indicates significantly negative effects, while the darkness of the colors represents the magnitude of the effects.

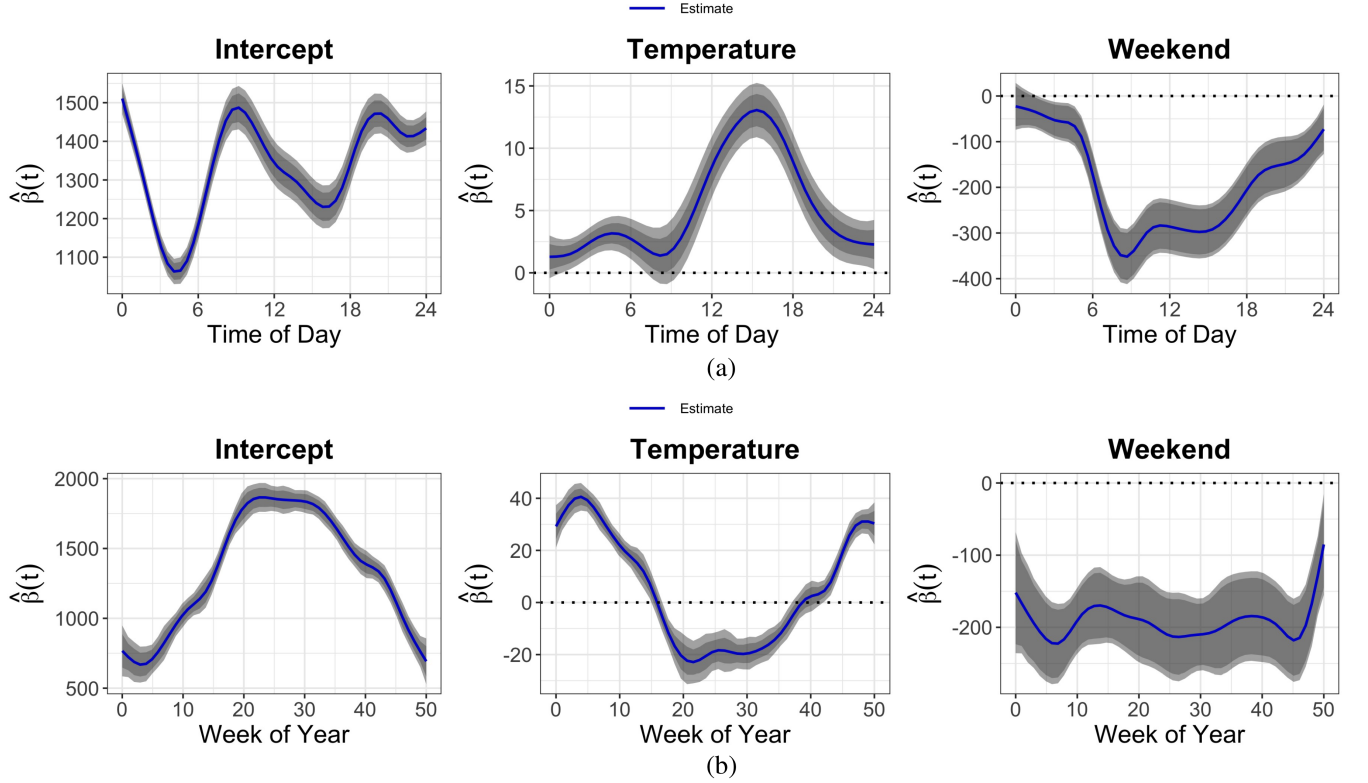


FIGURE 8 | Fixed-effects estimates (dashed blue line), 95% PCB (dark gray shaded area), and 95% SCB (light gray shaded area) of physical and intercept, temperature, and weekend. (a) Univariate effect over time of day, (b) Univariate effect over week of year.

demand during mid-year weeks (Australian winter). In summer, temperature positively affects demand, especially from 10 a.m. to 8 p.m., due to residential cooling needs and daily activity patterns. In winter, rising temperatures decrease demand, likely due to Adelaide's mild winter reducing heating needs. Weekends show lower electricity demand compared to weekdays, particularly around 6 a.m., reflecting later wake-up times. Figure 8 presents univariate effects of time of day and week of year separately, supporting the bivariate effects. However, the interplay of two temporal directions is only evident in the bivariate perspective of temperature, such as the prolonged positive effect around week 10 and week 40 at 3 p.m., possibly due to work-related activities offsetting climate effects.

5 | Conclusion

Motivated by daily activity profiles from wearable devices, our work presents an effective method for repeatedly measured functional data, common in longitudinal studies. This type of data is crucial in large-scale medical research, biomedicine, and environmental sciences. Efficient and flexible statistical tools are needed to handle ultra-high-dimensional and complex longitudinal-functional structures. We introduce a two-dimensional functional mixed-effect model with efficient fixed-effect inference. Our approach is demonstrated through analyses of adolescent physical activity profiles and hourly electricity demand data, examining their associations with various covariates.

Our fast three-stage estimation procedure significantly reduces computing costs for large-scale samples and dense functional grids. Although focused on dense, regular longitudinal sampling designs, the method can adapt to irregularly sampled data by averaging within equal-sized rectangular bins. The covariance function can be estimated using local-linear smoothing and functional principal components analysis, while maintaining two-dimensional fixed-effect inference. Our method supports the use of wearable devices in health research and offers broader applications for longitudinal studies and spatial analysis. Its flexibility and computational efficiency make it a valuable tool for researchers handling complex, high-dimensional functional data across multiple disciplines.

Acknowledgments

The authors thank all participants who took part in the study and the researchers who assisted with data collection. The study was supported by the City University of Hong Kong Grant Nos. 7005892, 7005967, and 9610473, The Institute of Digital Medicine, City University of Hong Kong and The Hong Kong Innovation and Technology Commission (InnoHK Project CIMDA).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All code for model implementation and simulation is available at <https://github.com/Cheng-0621/2dFMM>.

References

1. V. Carson, S. Hunter, N. Kuzik, et al., "Systematic Review of Sedentary Behaviour and Health Indicators in School-Aged Children and Youth: An Update," *Applied Physiology, Nutrition, and Metabolism* 41 (2016): S240–S265.
2. J. P. Chaput, J. Willumsen, F. Bull, et al., "2020 WHO Guidelines on Physical Activity and Sedentary Behaviour for Children and Adolescents Aged 5–17 Years: Summary of the Evidence," *International Journal of Behavioral Nutrition and Physical Activity* 17 (2020): 141.
3. B. G. G. da Costa, B. Bruner, G. H. Raymer, et al., "Association of Daily and Time-Segmented Physical Activity and Sedentary Behaviour With Mental Health of School Children and Adolescents From Rural North-eastern Ontario, Canada," *Frontiers in Psychology* 13 (2022): 1025444.
4. J. Goldsmith, V. Zipunnikov, and J. Schrack, "Generalized Multilevel Function-On-Scalar Regression and Principal Component Analysis," *Biometrics* 71 (2015): 344–353.
5. L. Xiao, L. Huang, J. A. Schrack, L. Ferrucci, V. Zipunnikov, and C. M. Crainiceanu, "Quantifying the Lifetime Circadian Rhythm of Physical Activity: A Covariate-Dependent Functional Approach," *Biostatistics* 16, no. 2 (2015): 352367.
6. R. Li, L. Xiao, E. Smirnova, E. Cui, A. Leroux, and C. M. Crainiceanu, "Fixed-Effects Inference and Tests of Correlation for Longitudinal Functional Data," *Statistics in Medicine* 41, no. 17 (2022): 33493364.
7. S. Y. Park, A. M. Staicu, L. Xiao, and C. M. Crainiceanu, "Simple Fixed-Effects Inference for Complex Functional Models," *Biostatistics* 19, no. 1 (2018): 137–152.
8. E. Cui, A. Leroux, E. Smirnova, and C. M. Crainiceanu, "Fast Univariate Inference for Longitudinal Functional Models," *Journal of Computational and Graphical Statistics* 31, no. 1 (2021): 219–230.
9. J. S. Morris and R. J. Carroll, "Wavelet-Based Functional Mixed Models," *Journal of the Royal Statistical Society, Series B* 68, no. 2 (2006): 179–199.
10. H. Zhu, K. Chen, X. Luo, J.-L. Yuan, and Y. Wang, "Fmem: Functional Mixed Effects Models for Longitudinal Functional Responses," *Statistica Sinica* 29 (2019): 20072033.
11. K. Chen, P. Delicado, and H. Müller, "Modeling Functional-Valued Stochastic Processes, With Applications to Fertility Dynamics," *Journal of the Royal Statistical Society, Series B* 79, no. 1 (2017): 177–196.
12. K. Chen and H. Müller, "Modeling Repeated Functional Observations," *Journal of the American Statistical Association* 107 (2012): 1599–1609.
13. E. Ivanescu, "Function-On-Function Regression for Two Dimensional Functional Data," *Communications in Statistics - Simulation and Computation* 47, no. 9 (2018): 2656–2669.
14. J. S. Morris, V. Baladandayuthapani, R. C. Herrick, P. Sanna, and H. Gutstein, "Automated Analysis of Quantitative Image Data Using Isomorphic Functional Mixed Models With Application to Proteomics Data," *Annals of Applied Statistics* 5, no. 2A (2011): 894–923.
15. J. Zhang, M. K. Clayton, and P. A. Townsend, "Functional Concurrent Linear Regression Model for Spatial Images," *Journal of Agricultural, Biological, and Environmental Statistics* 16, no. 1 (2011): 894130.
16. S. Y. Park and A. M. Staicu, "Longitudinal Functional Data Analysis," *Stat* 4, no. 1 (2015): 212–226.
17. L. Hu, T. Huang, and J. You, "Estimation and Identification of a Varying Coefficient Additive Model for Locally Stationary Processes," *Journal of the American Statistical Association* 114, no. 527 (2019): 1191–1204.
18. L. Hu, T. Huang, and J. You, "Robust Inference in Varying-Coefficient Additive Models for Longitudinal/Functional Data," *Statistica Sinica* 31, no. 2 (2021): 773–796.
19. S. Liu, J. You, and L. Hu, "Unified Statistical Inference for a Nonlinear Dynamic Functional/Longitudinal Data Model," *Journal of Statistical Planning and Inference* 219 (2022): 175–188.
20. J. Ramsay and B. Silverman, *Functional Data Analysis* (Springer, 2005).
21. D. Liang, H. Huang, Y. Guan, and F. Yao, "Test of Weak Separability for Spatially Stationary Functional Field," *Journal of the American Statistical Association* 118, no. 543 (2022): 1606–1619.
22. B. Lynch and K. Chen, "A Test of Weak Separability for Multi-Way Functional Data, With Application to Brain Connectivity Studies," *Biometrika* 105 (2018): 815–831.
23. J. A. Aston, D. Pigoli, and S. Tavakoli, "Tests for Separability in Non-parametric Covariance Operators of Random Surfaces," *Annals of Statistics* 45, no. 4 (2017): 1431–1461.
24. J. Fan and J.-T. Zhang, "Two-Step Estimation of Functional Linear Models With Applications to Longitudinal Data," *Journal of the Royal Statistical Society, Series B, Statistical Methodology* 62, no. 2 (2000): 303–322.
25. S. Greven, C. M. Crainiceanu, B. Caffo, and D. Reich, "Longitudinal Functional Principal Component Analysis," *Electronic Journal of Statistics* 4 (2010): 10221054.
26. F. Yao, H. Müller, and J. Wang, "Functional Data Analysis for Sparse Longitudinal Data," *Journal of the American Statistical Association* 100, no. 470 (2005): 577–590.
27. S. Chung and R. Kontar, "Functional Principal Component Analysis for Extrapolating Multistream Longitudinal Data," *IEEE Transactions on Reliability* 70, no. 4 (2021): 1321–1331.
28. E. Cui, R. Li, C. M. Crainiceanu, and L. Xiao, "Fast Multilevel Functional Principal Component Analysis," *Journal of Computational and Graphical Statistics* 32, no. 2 (2023): 366377.

29. L. Xiao, Y. Li, and D. Ruppert, “Fast Bivariate p-Splines: The Sandwich Smoother,” *Journal of the Royal Statistical Society, Series B* 75, no. 3 (2013): 577–599.
30. L. Xiao, C. Li, W. Checkley, and C. M. Crainiceanu, “Fast Covariance Estimation for Sparse Functional Data,” *Statistics and Computing* 28 (2018): 511–522.
31. L. Xiao, D. Ruppert, V. Zipunnikov, and C. M. Crainiceanu, “Fast Covariance Function Estimation for High-Dimensional Functional Data,” *Statistics and Computing* 26 (2016): 409–421.
32. C. M. Crainiceanu, J. Goldsmith, A. Leroux, and E. Cui, *Functional Data Analysis With R* (Chapman and Hall/CRC, 2024).
33. D. Ruppert, M. P. Wand, and R. J. Carroll, *Semiparametric Regression* (Cambridge University Press, 2003).
34. C. M. Crainiceanu, A.-M. Staicu, S. Ray, and N. Punjabi, “Bootstrap-Based Inference on the Difference in the Means of Two Correlated Functional Processes,” *Statistics in Medicine* 31 (2012): 3223–3240.
35. D. A. Degras, “Simultaneous Confidence Bands for Nonparametric Regression With Functional Data,” *Statistica Sinica* 21 (2011): 1735–1765.
36. X. Wang, J. Shen, and D. Ruppert, “On the Asymptotics of Penalized Spline Smoothing,” *Electronic Journal of Statistics* 5 (2011): 1–17.
37. Y. Li and D. Ruppert, “On the Asymptotics of Penalized Splines,” *Biometrika* 95, no. 2 (2008): 415–436.
38. M. P. Wand and M. C. Jones, *Kernel Smoothing* (Chapman & Hall, 1995).
39. F. Scheipl, A.-M. Staicu, and S. Greven, “Functional Additive Mixed Models,” *Journal of Computational and Graphical Statistics* 24 (2015): 477–501.
40. L. Magnano, J. Boland, and R. Hyndman, “Generation of Synthetic Sequences of Half-Hourly Temperature,” *Environmetrics* 19, no. 9 (2008): 818–835.
41. H. L. Shang and R. J. Hyndman, “FDS: Functional Data Sets [Computer Software Manual]. R Package,” (2016).

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1.** Supporting Information.