Fixed-effects inference for longitudinal functional data

Ruonan Li [1], **Luo Xiao** [1], **Ekaterina Smirnova** [2], **Erjia Cui** [3], **Andrew Leroux** [4], **and Ciprian M. Crainiceanu** [3]

[1] Department of Statistics, North Carolina State University

[2] Department of Biostatistics, Virginia Commonwealth University

[3] Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

[4] Department of Biostatistics and Informatics, Colorado School of Public Health

---

Address for correspondence: Luo Xiao, Department of Statistics, North Carolina State University.

E-mail: `lxiao5@ncsu.edu`.

Phone: (+1) 919 513 4859.

Fax: (+1) 919 513 4859.

Abstract: We propose an inferential framework for fixed effects in longitudinal functional models and introduce tests for the correlation structures induced by the longitudinal sampling procedure. The framework provides a natural extension of standard longitudinal correlation models for scalar observations to functional observations. Using simulation studies, we compare fixed effects estimation under correctly and incorrectly specified correlation structures and provide explicit recommendations.

We also apply the proposed methods to a longitudinal functional data set.

---

1   Introduction

Due to technological advancement, longitudinal functional data have become common in scientific studies (Ramsay, 2005). For example, objective physical activity measurements obtained from wearable accelerometers usually consists of multiple daily profiles measured at every minute of the day for hundreds or thousands of study participants at a number of visits. This type of data structure gives rise to longitudinal functional data (Greven et al., 2010; Goldsmith et al., 2011), which resembles traditional longitudinal data, but has a function instead of a scalar as the basic unit of observation. The average activity measured at multiple visits gives rise to standard longitudinal data, whereas minute-by-minute activity profiles observed at multiple visits give rise to longitudinal functional data. Another example is brain imaging data, where through MRI scans, imaging data along brain tracts were collected for patients of multiple sclerosis and healthy controls at a number of visits. This new data structure requires specialized methodological tools to address scientific problems using the original structure of the data.

Longitudinal functional data often exhibit complex within- and between-function correlations. To accommodate such correlations, various functional mixed effects models have been proposed (Di et al., 2009; Greven et al., 2010; Chen and Müller, 2012;

Scheipl et al., 2015; Park and Staicu, 2015; Zhang et al., 2016; Chen et al., 2017; Zhu et al., 2019; Lee et al., 2019). These approaches, after data transformation or projections, allow inference on the fixed effects parameters using the linear mixed effects (LME) inferential machinery. Alternatively, bootstrap of study participants can be used to construct confidence bands and conduct hypothesis tests for fixed effect functions (Crainiceanu et al., 2012; Park et al., 2018; Zhu et al., 2019). In spite of the growing interest in this area of research, modeling longitudinal functional data continues to be daunting. On the one hand, it is often difficult to specify within-function correlation structures that account for the continuity of the functional domain and are compatible with various sampling resolutions. On the other hand, the between-function correlations may be modeled flexibly using structured specifications that represent the known longitudinal sampling structure (e.g., nesting of functions within individuals). In this paper we consider a unified inferential model that explicitly separates the two types of correlations and propose to: (a) study the effects of mis-specified longitudinal correlation structures on fixed effects inference for longitudinal functional data; (b) compare various longitudinal correlation structures proposed in the literature; and (c) propose testing methods for selecting the appropriate longitudinal correlation structure.

Our framework builds on the marginal decomposition model proposed by Park and Staicu (2015), which focused on modeling the marginal covariance structure of longitudinal functional data, but did not consider fixed effects inference. We show that this approach can be used both for fixed effects inference and for the selection of appropriate longitudinal correlation models via hypothesis testing. For the latter, we extend the bootstrap-based approach for covariance testing proposed by Chen et al. (2019). We will also show how this framework can be used to compare models

with different longitudinal correlation structures and evaluate the validity of marginal inferences under misspecified longitudinal correlation models.

In summary, the contributions of this paper are to, for longitudinal functional data: (a) generalize standard longitudinal correlation models; (b) introduce testing for longitudinal correlation structures; and (c) study the effect of correctly and incorrectly specified correlation structures on the fixed effects inference.

The paper is organized as follows. In Section 2, we introduce the general functional mixed effects model and the specific longitudinal correlation structures for modeling longitudinal functional data. In Section 3, we address the problem of testing longitudinal correlation structure. Section 4 assesses the properties of the statistical methods proposed using simulations. Section 5 applies our methods to real data. We conclude the paper with a discussion in Section 6.

2   Functional mixed effects models

Let $\mathcal{T} \subset \mathbb{R}$ be a compact interval and the index $s \in \mathcal{T}$ be called the functional index. Let $y_{ij}(s)$ denote the observed response variable at s for subject i, $1 \leq i \leq n$, at the jth, $1 \leq j \leq n_i$, visit that occurred at time T_{ij}. Denote by $\mathbf{x}_{ij} = [x_{ij1}, \ldots, x_{ijp}]^T$ the vector of covariates for subject i at visit j, and $\mathbf{x}_i = [\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}]^T$ the $n_i \times p$ matrix of covariates for subject i. Some of the covariates may depend on the visit number and some may not, but we use the index j for all variables to simplify notation. For one data application on physical activity, the covariates are age, body mass index (BMI) and sex, and thus $p = 3$. Let $\mathbf{y}_{ij} = [y_{ij}(s_1), \ldots, y_{ij}(s_m)]^T$, where $\{s_1, \ldots, s_m\}$ is a fixed grid in $\mathcal{T}$. Let $\mathbf{y}_i = [\mathbf{y}_{i1}^T, \ldots, \mathbf{y}_{in_i}^T]^T$ be the mn_i-dimensional column vector

obtained by vectorizing all observations for subject i.

We denote by $\mu_{ij}(\cdot)$ the conditional mean of the functional process $y_{ij}(\cdot)$ given the covariates $\mathbf{x}_{ij}$ and assume that

$$\mu_{ij}(s) = \beta_0(s) + \sum_{r=1}^{p} x_{ijr}\beta_r(s) , \tag{2.1}$$

where $\beta_r(\cdot)$ are smooth, unknown functions of s. Using the approach in Park and Staicu (2015), we decompose the functional residuals, $y_{ij}(s) - \mu_{ij}(s) = W_{ij}(s) + \epsilon_{ij}(s)$, where

$$W_{ij}(s) = \sum_{k \geq 1} \xi_{ijk}\phi_k(s) ,$$

and $\epsilon_{ij}(s)$ is white noise with zero mean and variance σ_ϵ^2. The functions $\{\phi_1(\cdot), \phi_2(\cdot), \ldots\}$ form an orthonormal basis and ξ_{ijk} are associated random coefficients. We assume that ξ_{ijk} are normally distributed with zero mean and covariance $G_k(T_{ij_1}, T_{ij_2}) = \mathrm{Cov}\{\xi_{ij_1k}, \xi_{ij_2k}|T_{ij_1}, T_{ij_2}\}$ and are uncorrelated over i and k. Then the covariance of the process is a four-dimensional operator with the following structure

$$\mathcal{C}(s_1, T_{ij_1}, s_2, T_{ij_2}) = \mathrm{Cov}\{W_{ij_1}(s_1), W_{ij_2}(s_2)|T_{ij_1}, T_{ij_2}\} = \sum_{k \geq 1} G_k(T_{ij_1}, T_{ij_2})\phi_k(s_1)\phi_k(s_2).$$

We also assume that the sampling visits T_{ij} are independent over i and j and have a smooth probability density function, $g(\cdot)$. Then the basis functions $\phi_k(\cdot)$ are eigenfunctions of the marginal covariance operator

$$\mathcal{C}_m(s_1, s_2) = \int \mathcal{C}(s_1, T, s_2, T)g(T)\mathrm{d}T = \sum_{k \geq 1} \lambda_k\phi_k(s_1)\phi_k(s_2),$$

where $\lambda_k = \int G_k(T, T)g(T)\mathrm{d}T$ are eigenvalues and are assumed to be ordered decreasingly, i.e., $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$, and the random coefficients ξ_{ijk} satisfy $\xi_{ijk} = \int W_{ijk}(s)\phi_k(s)\mathrm{d}s$.

Standard longitudinal models use working covariance often assume either independence, exchangeability, or leave them unspecified. A natural generalization of these

assumptions to longitudinal functional data is via the specification of the working covariance, $G_k(\cdot, \cdot)$, for every eigenfunction, $\phi_k(\cdot)$, indexed by k. We propose to model the within-function covariance nonparametrically and the between-function covariance using approaches inspired by standard and functional longitudinal data analysis. The rationale is that the functions are often densely observed and exhibit complex correlation structure, for which a nonparametric approach would provide the needed flexibility, while the longitudinal sampling visits are often sparse, for which a simpler correlation structure might be sufficient and also easy to interpret.

2.1 Longitudinal working covariances

We consider three common correlation models including two parametric correlation structures, but methods can easily be extended to other parametric correlation structures.

The first is the working independence model, where $W_{ij}(\cdot)$ are assumed to be independent and identically distributed for all i and j. Under this framework the random scores $\xi_{ijk} = \zeta_{ijk}$, where ζ_{ijk} are uncorrelated across i and j for each k and $\mathrm{Var}(\zeta_{ijk}) = \sigma_{e,k}^2$. Then, we have $G_k(T_{ij_1}, T_{ij_2}) = \sigma_{e,k}^2 1_{\{j_1 = j_2\}}$.

The second is the exchangeable correlation model, that is $\xi_{ijk} = \xi_{ik} + \zeta_{ijk}$, where ξ_{ik} is uncorrelated with all ζ_{ijk} and ζ_{ijk} are mutually uncorrelated for all i, k and j. Denote $\sigma_{0,k}^2 = \mathrm{Var}(\xi_{ik})$ and $\sigma_{e,k}^2 = \mathrm{Var}(\zeta_{ijk})$, and then $G_k(T_{ij_1}, T_{ij_2}) = \sigma_{0,k}^2 + \sigma_{e,k}^2 1_{\{j_1 = j_2\}}$. The exchangeable model for longitudinal functional data is a particular case of the multilevel functional principal component model of Di et al. (2009). Indeed, we may write $W_{ij}(s)$ as $U_i(s) + V_{ij}(s)$ with $U_i(s) = \sum_{k \geq 1} \xi_{ik} \phi_k(s)$ and $V_{ij}(s) = \sum_{k \geq 1} \zeta_{ijk} \phi_k(s)$.

The third is the unspecified smooth correlation model, where $G_k(\cdot, \cdot)$ is modeled nonparametrically. In this case $\xi_{ijk} = \xi_{ik}(T_{ij}) + \zeta_{ijk}$, where $\xi_{ik}(T_{ij})$ is a zero-mean smooth process with covariance function $\text{Cov}\{\xi_{ik}(T_{ij_1}), \xi_{ik}(T_{ij_2})\} = G_{k0}(T_{ij_1}, T_{ij_2})$, and ζ_{ijk} are zero mean random errors with $\text{Var}(\zeta_{ijk}) = \sigma_{e,k}^2$. In particular, ζ_{ijk} are mutually uncorrelated for all i, j and k. The induced covariance function is $G_k(T_{ij_1}, T_{ij_2}) = G_{k0}(T_{ij_1}, T_{ij_2}) + \sigma_{e,k}^2 1_{\{j_1=j_2\}}$. We further decompose the covariance function $G_{k0}(T_{ij_1}, T_{ij_2})$ as $G_{k0}(T_{ij_1}, T_{ij_2}) = \sum_{\ell=1}^{\infty} \lambda_{k\ell} \psi_{k\ell}(T_{ij_1}) \psi_{k\ell}(T_{ij_2})$, where $\lambda_{k1} \geq \lambda_{k2} \geq \cdots \geq 0$ and $\psi_{k\ell}(\cdot)$ are orthonormal eigenfunctions that satisfy $\int_T \psi_{k\ell_1}(T) \psi_{k\ell_2}(T) \mathrm{d}T = 1_{\{\ell_1 \neq \ell_2\}}$. It follows that ξ_{ijk} can be written as $\sum_{\ell=1}^{\infty} \eta_{ik\ell} \psi_{k\ell}(T_{ij}) + \zeta_{ijk}$, where $\eta_{ik\ell}$ are uncorrelated random scores with zero mean and variance $\lambda_{k\ell}$.

The unspecified smooth covariance model contains the independent and exchangeable models as special cases. Indeed, the working independence model corresponds to $G_{k0}(T_{ij_1}, T_{ij_2}) = 0$ while the exchangeable correlation model corresponds to $G_{k0}(T_{ij_1}, T_{ij_2}) = \sigma_{0,k}^2$. Thus, for model estimation below, we will focus on the unspecified covariance model as the other cases can be treated as simpler cases.

2.2   Model estimation

We first consider estimation of $\mu_{ij}(\cdot)$ using B-spline expansion of the coefficient functions $\beta_r(\cdot)$. Specifically, let $\mathbf{B}(s) = [B_1(s), \ldots, B_c(s)]^T \in \mathbb{R}^c$ be a collection of B-spline basis functions in the unit interval with c being the number of basis functions. Then $\beta_r(s) = \sum_{d=1}^{c} \alpha_{rd} B_d(s)$, where $\boldsymbol{\alpha}_r = [\alpha_{r1}, \ldots, \alpha_{rc}]^T$ is the unknown coefficient vector. We use c and a relatively large number of equally spaced knots, i.e., a large c, so that the approximation bias is negligible. Let $\mathbf{B} = [\mathbf{B}(s_1), \ldots, \mathbf{B}(s_m)]^T \in \mathbb{R}^{m \times c}$, then $\boldsymbol{\mu}_{ij} = [\mu_{ij}(s_1), \ldots, \mu_{ij}(s_m)]^T = \mathbf{B}\boldsymbol{\alpha}_0 + \sum_{r=1}^{p} x_{ijr} \mathbf{B}\boldsymbol{\alpha}_r$. Denote $\mathbf{X}_{ij} = [\mathbf{B}, x_{ij1}\mathbf{B}, \ldots, x_{ijp}\mathbf{B}] \in$

$\mathbb{R}^{m \times ((p+1)c)}$, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_0^T, \ldots, \boldsymbol{\alpha}_p^T]^T \in \mathbb{R}^{(p+1)c}$. Then, we have $\boldsymbol{\mu}_{ij} = \mathbf{X}_{ij}\boldsymbol{\alpha}$. Let $\mathbf{X}_i = [\mathbf{X}_{i1}^T, \ldots, \mathbf{X}_{in_i}^T]^T \in \mathbb{R}^{(mn_i) \times ((p+1)c)}$ and $\boldsymbol{\mu}_i = [\boldsymbol{\mu}_{i1}^T, \ldots, \boldsymbol{\mu}_{in_i}^T]^T \in \mathbb{R}^{mn_i}$. Then, we have $\boldsymbol{\mu}_i = \mathbf{X}_i \boldsymbol{\alpha}$.

Now we consider $W_{ij}(s) = \sum_{k \geq 1} \xi_{ijk} \phi_k(s)$, where for each $k \geq 1$, $\xi_{ijk} = \sum_{\ell \geq 1} \eta_{ik\ell} \psi_{k\ell}(T_{ij}) + \zeta_{ijk}$. As both terms involve infinite sums, we propose to use a number of eigenfunctions $\phi_k(\cdot)$ that explains at least a given percentage of the variance explained (PVE) in the marginal covariance function, $\mathcal{C}_m$. Denote the selected number of $\phi_k(\cdot)$ by K. Next we determine the approximation of ξ_{ijk} for all $1 \leq k \leq K$ simultaneously by ranking $\lambda_{k\ell}$, the variances of scores $\eta_{ik\ell}$. We retain the $\lambda_{k\ell}$ that are larger than λ^*, where λ^* is the largest value that satisfies

$$\frac{\sum_{k=1}^K \sigma_{e,k}^2 + \sum_{k \geq 1} \sum_{\ell \geq 1} \lambda_{k\ell} 1_{\{\lambda_{k\ell} \geq \lambda^*\}}}{\sum_{k=1}^K \sigma_{e,k}^2 + \sum_{k \geq 1} \sum_{\ell \geq 1} \lambda_{k\ell}} > \text{PVE} .$$

This will discard $\eta_{ik\ell}$ with small variances and, according to our simulation studies, is a useful procedure for stabilizing the estimation. Let L_k be the number of eigenfunctions selected for ξ_{ijk}. Then approximately,

$$W_{ij}(s) = \sum_{1 \leq k \leq K} \left\{ \sum_{1 \leq \ell \leq L_k} \eta_{ik\ell} \psi_{k\ell}(T_{ij}) + \zeta_{ijk} \right\} \phi_k(s) .$$

Let $\boldsymbol{W}_{ij} = [W_{ij}(s_1), \ldots, W_{ij}(s_m)]^T \in \mathbb{R}^m$, $\boldsymbol{\phi}_k = [\phi_k(s_1), \ldots, \phi_k(s_m)]^T \in \mathbb{R}^m$, $\boldsymbol{\Psi}_{ijk} = [\psi_{k1}(T_{ij}), \ldots, \psi_{kL_k}(T_{ij})]^T \in \mathbb{R}^{L_k}$, and $\boldsymbol{\eta}_{ik} = [\eta_{ik1}, \ldots, \eta_{ikL_k}]^T \in \mathbb{R}^{L_k}$. Then

$$\boldsymbol{W}_{ij} = \sum_{1 \leq k \leq K} \left(\boldsymbol{\Psi}_{ijk}^T \boldsymbol{\eta}_{ik} + \zeta_{ijk} \right) \boldsymbol{\phi}_k .$$

Let $\boldsymbol{\Psi}_{ij} = \text{blockdiag}(\boldsymbol{\Psi}_{ij1}^T, \ldots, \boldsymbol{\Psi}_{ijK}^T) \in \mathbb{R}^{K \times L}$, where $L = \sum_{1 \leq k \leq K} L_k$, and $\boldsymbol{\eta}_i = [\boldsymbol{\eta}_{i1}^T, \ldots, \boldsymbol{\eta}_{iK}^T]^T \in \mathbb{R}^L$. Note that $\boldsymbol{\eta}_{ik}$ has a diagonal covariance matrix $\boldsymbol{\Lambda}_k = \text{diag}(\lambda_{k1}, \ldots, \lambda_{kL_k})$ and $\boldsymbol{\eta}_i$ has a diagonal covariance matrix $\boldsymbol{\Lambda} = \text{blockdiag}(\boldsymbol{\Lambda}_1, \ldots, \boldsymbol{\Lambda}_K)$. Let $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_K] \in \mathbb{R}^{m \times K}$ and $\boldsymbol{\zeta}_{ij} = [\zeta_{ij1}, \ldots, \zeta_{ijK}]^T \in \mathbb{R}^K$. Then

$$\boldsymbol{W}_{ij} = \boldsymbol{\Phi} \left(\boldsymbol{\Psi}_{ij} \boldsymbol{\eta}_i + \boldsymbol{\zeta}_{ij} \right) .$$

Let $\boldsymbol{W}_i = [\boldsymbol{W}_{i1}^T, \ldots, \boldsymbol{W}_{in_i}^T]^T \in \mathbb{R}^{mn_i}$, $\boldsymbol{\Psi}_i = [\boldsymbol{\Psi}_{i1}^T, \ldots, \boldsymbol{\Psi}_{in_i}^T]^T \in \mathbb{R}^{(Kn_i) \times L}$, and $\boldsymbol{\zeta}_i = [\boldsymbol{\zeta}_{i1}^T, \ldots, \boldsymbol{\zeta}_{n_i i}^T]^T \in \mathbb{R}^{Kn_i}$. Then

$$\boldsymbol{W}_i = (\mathbf{I}_{n_i} \otimes \boldsymbol{\Phi}) (\boldsymbol{\Psi}_i \boldsymbol{\eta}_i + \boldsymbol{\zeta}_i) \ .$$

It is easy to show that $\boldsymbol{\zeta}_{ij}$ has a diagonal covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{e,1}^2, \ldots, \sigma_{e,K}^2)$ and that $\boldsymbol{\zeta}_i$ has a diagonal covariance matrix $\mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}$. For the random error term $\epsilon_{ij}(s)$, let $\boldsymbol{\epsilon}_{ij} = [\epsilon_{ij}(s_1), \ldots, \epsilon_{ij}(s_m)]^T \in \mathbb{R}^m$ and $\boldsymbol{\epsilon}_i = [\boldsymbol{\epsilon}_{i1}^T, \ldots, \boldsymbol{\epsilon}_{in_i}^T]^T \in \mathbb{R}^{mn_i}$.

Finally, we obtain the mixed model representation

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\alpha} + (\mathbf{I}_{n_i} \otimes \boldsymbol{\Phi}) (\boldsymbol{\Psi}_i \boldsymbol{\eta}_i + \boldsymbol{\zeta}_i) + \boldsymbol{\epsilon}_i,$$

$$\begin{pmatrix} \boldsymbol{\eta}_i \\ \boldsymbol{\zeta}_i \\ \boldsymbol{\epsilon}_i \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \mathbf{0}_L \\ \mathbf{0}_{Kn_i} \\ \mathbf{0}_{mn_i} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Lambda} & 0 & 0 \\ 0 & \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma} & 0 \\ 0 & 0 & \sigma_\epsilon^2 \mathbf{I}_{mn_i} \end{pmatrix} \right], \tag{2.2}$$

where $[\boldsymbol{\eta}_i^T, \boldsymbol{\zeta}_i^T, \boldsymbol{\epsilon}_i^T]^T$ are independent across subjects. Therefore, given the eigenfunctions $\boldsymbol{\Psi}_i$ and $\boldsymbol{\Phi}$, the fixed effects $\boldsymbol{\alpha}$ and variance components $\lambda_{k\ell}$, $\sigma_{e,k}^2$, and σ_ϵ^2 can be estimated using standard mixed effects software. To avoid overfitting, we impose a quadratic penalty on each $\boldsymbol{\alpha}_r$. Let $\mathcal{P}$ be a second order difference penalty matrix (Eilers and Marx, 1996). To be specific we use the penalty $\sum_{r=0}^p \tau_r \boldsymbol{\alpha}_r^T \mathcal{P} \boldsymbol{\alpha}_r$, where τ_r is the smoothing parameter for the rth function $\beta_r(\cdot)$. The model can be fitted using the *gam* function (Wood et al., 2016) in the R package *mgcv* (Wood, 2019) and we use generalized cross validation to select the smoothing parameters.

2.3   Longitudinal covariance estimation

We will use an iterative procedure between estimating the mean and covariance structure, respectively. First, assume that $\boldsymbol{\alpha}$ is estimated, and thus $\beta_r(\cdot)$s, are known.

Then the residuals $r_{ij}(s) = y_{ij}(s) - \mu_{ij}(s)$ can be calculated. An initial estimate of $\boldsymbol{\alpha}$ can be obtained, for example, by fitting model (2.2) without including the term $(\mathbf{I}_{n_i} \otimes \boldsymbol{\Phi})(\boldsymbol{\Psi}_i \boldsymbol{\eta}_i + \boldsymbol{\zeta}_i)$. This is equivalent to assume that all observations $y_{ij}(s)$ are mutually independent.

Once the residuals are estimated, the eigenfunctions $\phi_k(\cdot)$ can be estimated by using the fast *face* algorithm (Xiao et al., 2016) of the sandwich smoother (Xiao et al., 2013) on the set of pooled residuals, $r_{ij}(s)$. Specifically, let $\boldsymbol{R}_{ij} = [r_{ij}(s_1), \ldots, r_{ij}(s_m)]^T \in \mathbb{R}^m$, $\boldsymbol{R}_i = [\boldsymbol{R}_{i1}, \ldots, \boldsymbol{R}_{in_i}] \in \mathbb{R}^{m \times n_i}$ and $\boldsymbol{R} = [\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n] \in \mathbb{R}^{m \times N}$, where $N = \sum_{i=1}^n n_i$. Then we apply the *face* algorithm to $N^{-1}\boldsymbol{R}\boldsymbol{R}^T$.

Given $\phi_k(\cdot)$, we calculate $\int r_{ij}(s)\phi_k(s)\mathrm{d}s$ as the prediction for ξ_{ijk}. For the working independence model, λ_k can be estimated along with $\phi_k(\cdot)$ directly. For the exchangeable model, we have $\xi_{ijk} = \xi_{ik} + \zeta_{ijk}$ and a one-way nested ANOVA model can be used to estimate $\sigma_{0,k}^2$ and $\sigma_{e,k}^2$. Finally, when $G_k(\cdot, \cdot)$ is unspecified, we use the fast covariance estimation method for sparse functional data (Xiao et al., 2018) to ξ_{ijk} to obtain estimates of $G_{k0}(\cdot, \cdot)$ and $\sigma_{e,k}^2$ for each k.

3   Tests of longitudinal covariances

We use the goodness of fit test in Chen et al. (2019) to test the structure of the covariance $G_{k0}(\cdot, \cdot)$ in the following scenarios: (1) test the independent model against the unspecified covariance model; and (2) test the exchangeable model against the unspecified covariance model. Tests with specified parametric alternatives in a similar setting have been considered in Park and Staicu (2015). We will consider two types of tests: test of an individual covariance or joint test of multiple covariances. To

illustrate the idea, consider the case when the null hypothesis assumes independence. For the individual test, we may test $H_{0k} : G_{k0}(\cdot, \cdot) = 0$ against $H_{1k} : G_{k0}(\cdot, \cdot)$ is unspecified for each k. For the joint test, we may test $H_0 : G_{k0}(\cdot, \cdot) = 0$ for all $k \geq 1$ against $H_1 : G_{k0}(\cdot, \cdot)$ is unspecified for at least one $k \geq 1$. The null model for the joint test corresponds to the assumption that the longitudinal functions $W_{ij}(\cdot)$ are uncorrelated across visits, j, within study participants.

3.1   Individual longitudinal covariance test

The hypothesis test can be written as

$$H_{0k} : G_{k0}(T, T') = \widetilde{G}_{k0}(T, T') \quad \text{versus} \quad H_{1k} : G_{k0}(T, T') \neq \widetilde{G}_{k0}(T, T') \,, \qquad (3.1)$$

where $\widetilde{G}_{k0}(T, T')$ is a parametric covariance function with a finite number of unknown parameters. For example, to test the null hypothesis of an exchangeable covariance model, we have $G_k(T_{ij_1}, T_{ij_2}) = \widetilde{G}_{k0}(T_{ij_1}, T_{ij_2}) + 1_{\{j_1=j_2\}}\sigma_{e,k}^2$, where $\widetilde{G}_{k0}(T_{ij_1}, T_{ij_2}) = \sigma_{0,k}^2$. For testing the null hypothesis of working independence, everything stays the same, except that $\widetilde{G}_{k0}(T_{ij_1}, T_{ij_2}) = 0$.

We now briefly describe the test proposed in Chen et al. (2019). Suppose that the null covariance is G_0 and the alternative is G_A. First, we estimate the unknown parameters under the null hypothesis and denote the resulting estimate by $\widehat{G}_0$. Under the alternative hypothesis, we estimate G_A nonparametrically using, for example, smoothing based on tensor product of B-splines. If the estimator under the alternative hypothesis is $\widehat{G}_A$ then the test statistic is the Hilbert-Schmidt norm distance

$$T_n = \left\| \widehat{G}_A - \mathcal{K}\widehat{G}_0 \right\|_{HS} \,,$$

where $\|f\|_{HS} = \sqrt{\iint f(T, T')^2 \mathrm{d}T \mathrm{d}T'}$ and $\mathcal{K}\widehat{G}_0$ is the smoothed null covariance esti-

mate using again tensor-product of B-splines. The rationale for smoothing $\widehat{G}_0$ is to remove the potential bias induced by smoothing procedure used for obtaining $\widehat{G}_A$. Larger values of the test statistic, T_n, will provide evidence against the null hypothesis. The null distribution of T_n is obtained via wild bootstrap (Wu, 1986), i.e., the estimated error variance under the alternative hypothesis is used to generate bootstrap samples.

3.2   Joint test of multiple longitudinal covariances

To explore the overall structure of longitudinal covariances for modeling longitudinal functional data, we propose a joint test of $G_{k0}(\cdot, \cdot)$ for all $1 \leq k \leq K$. The approach we propose is straightforward: conduct individual covariance tests and use a Bonferroni correction to control the family-wise error rate.

4   Simulation studies

We use simulations to compare the performance of inferential approaches for the longitudinal functional models using independent, exchangeable, and unspecified covariance assumptions. We will focus on estimation and inference for fixed effects as well as size and power of the covariance tests introduced in Section 3. The longitudinal functional models are compared with the standard functional data method implemented in the *pffr* function (Ivanescu et al., 2015; Scheipl et al., 2015, 2016) of the R package *refund* (Goldsmith et al., 2018). This approach was designed for functional data without residual correlation and will be referred to as the naive method, in the sense that the method is naive to the longitudinal correlation assumptions.

The methods implemented in the *pffr* function work very well in practice when residuals are not correlated. We also compare methods with the bootstrap method (Park et al., 2018) to obtain the confidence bands for the coefficient functions. The point estimators for the bootstrap approach are obtained using the *pffr* function.

4.1 Simulation settings

Data are generated from the model $y_{ij}(s) = \mu_{ij}(s) + W_{ij}(s) + \epsilon_{ij}(s)$, $s \in [0, 1]$. The functional mean is $\mu_{ij}(s) = \beta_0(s) + x_{i1}\beta_1(s) + x_{ij2}\beta_2(s)$, where the number of covariates is $p = 2$, $x_{i1} \overset{iid}{\sim} N(0, 1)$, and $x_{ij2} = aT_{ij} + e_{ij}$, where $e_{ij} \sim N(\rho e_{i(j-1)}, 1)$ with $e_{i0} = 0$. We let $a = 1$ and $\rho = 0.7$. The coefficient functions are $\beta_0(s) = 1 + \sqrt{3}\cos(3\pi s)$, $\beta_1(s) = 2 + \cos(2\pi s)$ and $\beta_2(s) = 2 + \sin(\pi s)$. The number of visits per subject, m_i, will be specified later while the visit times, T_{ij}, are generated uniformly from the unit interval and then ordered. Using the orthonormal functions $\phi_1(s) = 1$ and $\phi_2(s) = \sqrt{2}\sin(2\pi s)$, $W_{ij}(\cdot)$ is generated from one of the three models below:

1. Independent covariance: $W_{ij}(s) = \sum_{k=1}^{2} \zeta_{ijk}\phi_k(s)$.

2. Exchangeable covariance: $W_{ij}(s) = \sum_{k=1}^{2} (\xi_{ik} + \zeta_{ijk})\phi_k(s)$.

3. Unspecified covariance: $W_{ij}(s) = \sum_{k=1}^{2} \left\{ \sum_{\ell=1}^{2} \eta_{ik\ell}\psi_{k\ell}(T_{ij}) + \zeta_{ijk} \right\} \phi_k(s)$.

For the independent covariance, $\zeta_{ij1} \overset{iid}{\sim} N(0, 4.5)$ and $\zeta_{ij2} \overset{iid}{\sim} N(0, 3)$. For the exchangeable covariance, $\xi_{i1} \overset{iid}{\sim} N(0, 3)$, $\xi_{i2} \overset{iid}{\sim} N(0, 2)$, $\zeta_{ij1} \overset{iid}{\sim} N(0, 1.5)$ and $\zeta_{ij2} \overset{iid}{\sim} N(0, 1)$. For the unspecified covariance, $\psi_{11}(T) = \sqrt{2}\cos(2\pi T)$, $\psi_{12}(T) = \sqrt{2}\sin(2\pi T)$, $\psi_{21}(T) = \sqrt{2}\cos(4\pi T)$, and $\psi_{22}(T) = \sqrt{2}\sin(4\pi T)$. We generate the random terms as $\eta_{i11} \overset{iid}{\sim} N(0, 2)$, $\eta_{i12} \overset{iid}{\sim} N(0, 1.5)$, $\eta_{i21} \overset{iid}{\sim} N(0, 1.5)$, $\eta_{i22} \overset{iid}{\sim} N(0, 1)$, $\zeta_{ij1} \overset{iid}{\sim} N(0, 1)$ and

$\zeta_{ij2} \overset{iid}{\sim} N(0, 0.5)$. The functional arguments $\{s_1, \ldots, s_m\}$ form a grid of 101 equidistant points in the unit interval.

The mutually independent random errors $\epsilon_{ij}(s)$ are generated from $N(0, \sigma_\epsilon^2)$. The value of σ_ϵ^2 is 1.5, which is determined by a signal to noise ratio (SNR) with a value of 5. For the independent and exchangeable covariances, we defined SNR in the functional data as $\sum_{k \geq 1}(\sigma_{0,k}^2 + \sigma_{e,k}^2)/\sigma_\epsilon^2$; for the unspecified covariance, we defined SNR as $\sum_{k \geq 1}(\sum_{\ell \geq 1}\lambda_{k\ell} + \sigma_{e,k}^2)/\sigma_\epsilon^2$.

For simulations, we used a factorial design with three factors: (a) the number of subjects is either 100 or 200; (b) the number of visits m_i are generated from either $\mathrm{Unif}[3, \ldots, 6]$ or $\mathrm{Unif}[8, \ldots, 12]$; and (c) the covariance model is either independent, exchangeable or unspecified. There are 12 model conditions and for each condition we simulate 500 datasets.

For tests of longitudinal covariances, data are generated similarly. However, under the alternative we add terms to ensure that the null model is misspecified. Specifically, we consider

1. Deviation from independent covariance: $\xi_{ijk} \rightarrow \xi_{ijk} + \Delta z_{ik}(T_{ij})$.

2. Deviation from exchangeable covariance: $\xi_{ik} + \zeta_{ijk} \rightarrow \xi_{ik} + \zeta_{ijk} + \Delta z_{ik}(T_{ij})$.

The scalar Δ controls the magnitude of the deviation from the null model. The scores ξ_{ik} and ζ_{ijk} are independent from the non-linear random function $z_{ik}(\cdot)$, where $z_{ik}(T) = \sum_{\ell=1}^2 \eta_{ik\ell}\psi_{k\ell}(T)$. The terms $\eta_{ik\ell}$ and $\psi_{k\ell}(T)$ are specified as before.

4.2   Model estimation and evaluation criteria

The initial mean estimation is conducted using the *pffr* function using 10 cubic B-splines bases and the second order difference penalty. The marginal eigenfunctions $\phi_k(s)$ are estimated by fitting the functional residuals via the *fpca.face* function (Xiao et al., 2016) in the R package *refund* with 20 knots. The unspecified covariance model components $\psi_{k\ell}(T)$ and $\xi_{ik\ell}$ are estimated using the *face.sparse* function in the R package *face* with 15 knots. The truncation parameters K and L_k (for the unspecified covariance model) are all determined using the pre-specified level percentage of variance explained (PVE) with a value of 0.95. The final mixed model is fitted using *bam* function (Wood et al., 2015) in the R package *mgcv* (Wood, 2019).

The estimation accuracy for functional parameters is assessed using the mean integrated squared errors (MISE) and the properties of the pointwise confidence intervals. Specifically, for a function β with estimate $\hat{\beta}$, $\mathrm{MISE}(\hat{\beta}) = \frac{1}{N_{sim}} \sum_{i_{sim}=1}^{N_{sim}} \int \{\hat{\beta}^{i_{sim}}(s) - \beta(s)\}^2 ds$. For the pointwise confidence intervals, we report the integrated actual pointwise coverage (IAC) and integrated actual width (IAW), where IAC is defined as $\frac{1}{N_{sim}} \sum_{i_{sim}=1}^{N_{sim}} \int I\{\hat{\beta}^{i_{sim}}(s) \in CI_p(s)\} ds$ and $CI_p(s)$ is the pointwise approximate confidence interval for the functional parameter $\beta(s)$. For example, an approximate 95% pointwise confidence interval $CI_p(s_0)$ for $\beta_1(s_0)$ can be constructed as $\beta_1(s_0) \pm 1.96 \times \hat{\mathrm{SD}}\{\hat{\beta}_1(s_0)\}$. The length of the confidence interval at s_0 is $3.92 \times \hat{\mathrm{SD}}\{\hat{\beta}_1(s_0)\}$ and $\mathrm{IAW} = \frac{1}{N_{sim}} \sum_{i_{sim}=1}^{N_{sim}} \int 3.92 \times \hat{\mathrm{SD}}\{\hat{\beta}^{i_{sim}}(s)\} ds$.

4.3   Simulation results for estimation

Tables 1, 2 and 3 display the results for estimating the coefficient functions when the data generating mechanism uses independent, exchangeable, or unspecified co-variances, respectively. We first focus on the point estimation, which is partially captured by the MISE in the three tables. Results indicate that as the number of subjects and visits increase, MISEs decrease for all methods; see the columns labeled $\sqrt{\mathrm{MISE}}$ in Tables 1, 2 and 3. Moreover, the MISE obtains the smallest value when the correlation is correctly specified. In particular, the reduction in MISE is substantial for data with exchangeable and unspecified covariance structures when the correct model is used compared to one that does not account for correlation. For example, in Table 2 which corresponds to data simulated from an exchangeable correlation structure, for $n = 200$ and $m_i \sim \mathrm{Unif}[8, 12]$, the $\sqrt{\mathrm{MISE}}$ was 0.12 for $\beta_2(s)$ for the method which accounted for exchangeable correlation, while it was 0.28 for the Park et al. method which did not. Finally, the unspecified correlation model consistently provides good estimators, with the corresponding MISE comparable to that obtained from the correctly specified model.

For confidence bands, all methods outperform the naive method in terms of cover-age. The confidence bands from the naive method are on average too narrow and are far below the nominal coverage level. The correctly-specified and the unspecified covariance models provide confidence bands that are close to the nominal, though, slightly conservative sometimes. The bootstrap method provides confidence bands that close to the nominal, though they require more computational times and have higher average width for estimating slope functions and lower average width for esti-mating intercept functions. When models are misspecified (e.g., exchangeable model

for unspecified covariance data), the average coverage of the confidence bands tends to be close to the nominal, but at the expense of much wider confidence bands.

In terms of running time, the bootstrap method is the slowest and the naive method is the fastest. Computational time increases with more complex longitudinal correlation structures. However, when the dataset has a large number of subjects and a large number of visits per subject ($n = 200$ and $m_i \sim \mathrm{Unif}[8, 12]$), the computation time for fitting the proposed methods can be comparable to that of the bootstrap.

In summary, the unspecified covariance models perform well in terms of inference for the fixed effects in a variety of scenarios. However, in practice one may still want to know whether a simpler correlation structure is reasonable for the data, as the correct model slightly outperforms the unspecified covariance model. In the next section, we provide results for testing the longitudinal covariance.

4.4   Simulation results for longitudinal covariance tests

For the covariance tests introduced in Section 3, we report the empirical type I error rate (estimated size) for nominal levels $\alpha = 0.05$ and 0.10 based on 5000 simulated datasets, and the power at the $\alpha = 0.05$ level with 1000 simulated datasets. For the joint test, we use the Bonferroni correction to account for the multiplicity of hypotheses tests. For the individual covariance tests, the power curves are presented as functions of the deviation from the null, defined as

$$\frac{\Delta^2 \int \mathrm{Var}\{z_{ik}(T)\}dT}{\int \mathrm{Var}\{\xi_{ik}(T)\}dT + \Delta^2 \int \mathrm{Var}\{z_{ik}(T)\}dT}$$

for each k. For the joint test, the deviation from the null is defined as

$$\frac{\Delta^2 \sum_k \int \mathrm{Var}\{z_{ik}(T)\}dT}{\sum_k \int \mathrm{Var}\{\xi_{ik}(T)\}dT + \Delta^2 \sum_k \int \mathrm{Var}\{z_{ik}(T)\}dT} \ .$$

Table 1: $\sqrt{\text{MISE}}$, IAC for 95% pointwise confidence interval and IAW for estimating $\beta_0(s)$, $\beta_1(s)$ and $\beta_2(s)$ based on independent covariance data using five methods across 500 simulations.

Method	Time(secs)	$\beta_0(s)$			$\beta_1(s)$			$\beta_2(s)$		
		$\sqrt{\text{MISE}}$	IAW	IAC	$\sqrt{\text{MISE}}$	IAW	IAC	$\sqrt{\text{MISE}}$	IAW	IAC
$n = 100$ & $m_i \sim \text{Unif}[3,6]$										
Naive	2	0.143	0.23	0.58	0.133	0.15	0.45	0.094	0.09	0.40
Park et al.	417	0.143	0.54	0.94	0.133	0.51	0.94	0.094	0.35	0.94
Independent	7	0.142	0.81	0.97	0.129	0.51	0.96	0.085	0.33	0.95
Exchangeable	12	0.142	0.82	0.98	0.129	0.52	0.96	0.085	0.34	0.95
Unspecified	12	0.143	0.84	0.98	0.130	0.52	0.96	0.086	0.34	0.95
$n = 100$ & $m_i \sim \text{Unif}[8,12]$										
Naive	3	0.093	0.16	0.60	0.087	0.10	0.46	0.060	0.07	0.44
Park et al.	557	0.093	0.36	0.94	0.087	0.34	0.95	0.060	0.23	0.95
Independent	61	0.092	0.54	0.97	0.086	0.34	0.96	0.056	0.23	0.96
Exchangeable	89	0.092	0.55	0.97	0.085	0.35	0.96	0.056	0.23	0.96
Unspecified	61	0.093	0.56	0.98	0.086	0.35	0.96	0.056	0.23	0.96
$n = 200$ & $m_i \sim \text{Unif}[3,6]$										
Naive	3	0.096	0.16	0.60	0.091	0.11	0.46	0.063	0.07	0.42
Park et al.	535	0.096	0.38	0.95	0.091	0.36	0.94	0.063	0.25	0.95
Independent	43	0.095	0.56	0.97	0.090	0.36	0.95	0.058	0.24	0.96
Exchangeable	93	0.095	0.57	0.98	0.090	0.36	0.96	0.058	0.24	0.96
Unspecified	46	0.096	0.58	0.98	0.090	0.36	0.95	0.058	0.24	0.96
$n = 200$ & $m_i \sim \text{Unif}[8,12]$										
Naive	5	0.064	0.11	0.63	0.062	0.07	0.48	0.044	0.05	0.43
Park et al.	854	0.064	0.25	0.95	0.062	0.24	0.94	0.044	0.17	0.94
Independent	562	0.064	0.38	0.97	0.062	0.24	0.94	0.042	0.16	0.95
Exchangeable	815	0.064	0.39	0.98	0.062	0.24	0.95	0.042	0.16	0.96
Unspecified	533	0.064	0.38	0.98	0.062	0.24	0.94	0.042	0.16	0.96

Table 2: $\sqrt{\text{MISE}}$, IAC for 95% pointwise confidence interval and IAW for estimating $\beta_0(s)$, $\beta_1(s)$ and $\beta_2(s)$ based on exchangeable covariance data using five methods across 500 simulations.

Method	Time(secs)	$\beta_0(s)$ $\sqrt{\text{MISE}}$	IAW	IAC	$\beta_1(s)$ $\sqrt{\text{MISE}}$	IAW	IAC	$\beta_2(s)$ $\sqrt{\text{MISE}}$	IAW	IAC
$n = 100$ & $m_i \sim \text{Unif}[3,6]$										
Naive	2	0.248	0.22	0.35	0.248	0.15	0.24	0.138	0.09	0.29
Park et al.	419	0.248	0.97	0.95	0.248	0.94	0.94	0.138	0.51	0.94
Independent	7	0.246	0.80	0.86	0.242	0.50	0.69	0.125	0.33	0.81
Exchangeable	11	0.240	1.43	0.97	0.225	0.89	0.95	0.072	0.29	0.95
Unspecified	15	0.242	1.43	0.98	0.229	0.87	0.94	0.073	0.29	0.95
$n = 100$ & $m_i \sim \text{Unif}[8,12]$										
Naive	3	0.233	0.15	0.27	0.227	0.10	0.18	0.108	0.07	0.26
Park et al.	561	0.233	0.92	0.95	0.227	0.90	0.95	0.108	0.39	0.93
Independent	64	0.233	0.53	0.73	0.224	0.34	0.56	0.102	0.23	0.75
Exchangeable	86	0.230	1.37	0.97	0.209	0.87	0.97	0.043	0.17	0.96
Unspecified	87	0.236	1.37	0.98	0.217	0.85	0.95	0.043	0.17	0.95
$n = 200$ & $m_i \sim \text{Unif}[3,6]$										
Naive	3	0.177	0.16	0.36	0.177	0.11	0.24	0.094	0.07	0.30
Park et al.	535	0.177	0.69	0.95	0.177	0.66	0.93	0.094	0.37	0.95
Independent	43	0.175	0.57	0.86	0.175	0.36	0.69	0.088	0.24	0.83
Exchangeable	86	0.168	1.02	0.97	0.166	0.64	0.95	0.054	0.21	0.95
Unspecified	90	0.169	1.02	0.99	0.167	0.63	0.94	0.054	0.21	0.95
$n = 200$ & $m_i \sim \text{Unif}[8,12]$										
Naive	5	0.168	0.11	0.28	0.160	0.07	0.20	0.075	0.05	0.27
Park et al.	845	0.168	0.65	0.94	0.160	0.63	0.95	0.075	0.28	0.94
Independent	538	0.168	0.38	0.73	0.159	0.24	0.57	0.073	0.16	0.75
Exchangeable	751	0.165	0.98	0.97	0.151	0.62	0.96	0.030	0.12	0.96
Unspecified	732	0.166	0.96	0.97	0.153	0.61	0.95	0.030	0.12	0.96

Table 3: $\sqrt{\text{MISE}}$, IAC for 95% pointwise confidence interval and IAW for estimating $\beta_0(s)$, $\beta_1(s)$ and $\beta_2(s)$ based on unspecified covariance data using five methods across 500 simulations.

Method	Times(secs)	$\beta_0(s)$			$\beta_1(s)$			$\beta_2(s)$		
		$\sqrt{\text{MISE}}$	IAW	IAC	$\sqrt{\text{MISE}}$	IAW	IAC	$\sqrt{\text{MISE}}$	IAW	IAC
$n = 100$ & $m_i \sim \text{Unif}[3,6]$										
Naive	2	0.134	0.23	0.59	0.133	0.15	0.44	0.090	0.09	0.40
Park et al.	416	0.134	0.52	0.95	0.133	0.49	0.93	0.090	0.36	0.95
Independent	8	0.133	0.81	0.99	0.129	0.51	0.96	0.082	0.33	0.96
Exchangeable	13	0.133	0.82	0.99	0.129	0.52	0.96	0.082	0.34	0.96
Unspecified	28	0.093	0.63	1.00	0.084	0.31	0.94	0.060	0.22	0.93
$n = 100$ & $m_i \sim \text{Unif}[8,12]$										
Naive	3	0.092	0.16	0.64	0.088	0.11	0.47	0.070	0.07	0.39
Park et al.	559	0.092	0.37	0.96	0.088	0.34	0.94	0.070	0.26	0.92
Independent	72	0.091	0.56	0.98	0.086	0.36	0.96	0.066	0.24	0.93
Exchangeable	96	0.091	0.56	0.98	0.087	0.36	0.96	0.066	0.24	0.93
Unspecified	168	0.058	0.43	1.00	0.058	0.21	0.93	0.042	0.15	0.94
$n = 200$ & $m_i \sim \text{Unif}[3,6]$										
Naive	3	0.090	0.1600	0.60	0.081	0.10	0.48	0.072	0.07	0.35
Park et al.	533	0.090	0.34	0.94	0.081	0.31	0.95	0.072	0.28	0.95
Independent	47	0.089	0.54	0.98	0.079	0.34	0.97	0.068	0.23	0.91
Exchangeable	109	0.089	0.54	0.98	0.079	0.34	0.97	0.068	0.23	0.91
Unspecified	237	0.050	0.36	1.00	0.046	0.18	0.96	0.034	0.13	0.96
$n = 200$ & $m_i \sim \text{Unif}[8,12]$										
Naive	5	0.062	0.11	0.63	0.055	0.07	0.52	0.051	0.05	0.39
Park et al.	850	0.062	0.24	0.95	0.055	0.22	0.96	0.051	0.20	0.95
Independent	600	0.061	0.38	0.98	0.054	0.24	0.97	0.049	0.16	0.90
Exchangeable	848	0.061	0.38	0.98	0.054	0.24	0.98	0.049	0.16	0.90
Unspecified	1562	0.034	0.26	1.00	0.032	0.12	0.95	0.024	0.09	0.95

The deviation from the null hypothesis is a number between 0 and 1.

Table 4 reports the empirical type I errors of the individual and joint tests, indicating that all tests have empirical levels close to their nominal levels. Tests can be slightly conservative for exchangeable covariance with a larger number of visits per subject.

Figure 1 presents the power curves for the proposed tests, indicating similar patterns for the two individual and the joint tests. The test for independent covariance has higher power than the one for exchangeable covariance data. For both types of the null hypothesis, power increases with the number of subjects and observation per subject. Power seems to be particularly sensitive to the number of visits, m_i, per curve. Indeed, with 100 study participants and exchangeable covariance, for all three tests, the power for $m_i \sim \text{Unif}[8, 12]$ is close to one under a deviation around 0.25. In contrast, for $m_i \sim \text{Unif}[3, 6]$, the deviation needs to be closer to 0.7 to reach the same power. As expected, power depends on how closely the true model matches the specified alternative assumed by the test.

In summary, in all settings, the proposed tests maintain proper size and have the power to detect alternative longitudinal covariance both when the null hypothesis assumes independent covariance and exchangeable covariance.

5   Data application

Wearable devices such as accelerometers provide objective and detailed measurements of physical activity and enable researchers to study how physical activity changes with age, personal characteristics, and health status. In an earlier study based on

Table 4: Empirical type I error of covariance tests at the nominal $\alpha = 0.05$ and 0.10 levels based on 5000 datasets, by sample size (n) and observations per subject (m_i).

Null	m_i	Layer	$n = 100$		$n = 200$	
			$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$
Exchangeable	Unif[3, 6]	First layer	0.055	0.105	0.046	0.098
		Second layer	0.050	0.099	0.045	0.093
		Joint test	0.053	0.103	0.043	0.090
	Unif[8, 12]	First layer	0.042	0.094	0.039	0.085
		Second layer	0.037	0.080	0.040	0.092
		Joint test	0.037	0.078	0.033	0.078
Independent	Unif[3, 6]	First layer	0.056	0.115	0.057	0.110
		Second layer	0.051	0.108	0.052	0.103
		Joint test	0.053	0.105	0.056	0.106
	Unif[8, 12]	First layer	0.053	0.103	0.057	0.108
		Second layer	0.054	0.103	0.055	0.102
		Joint test	0.056	0.104	0.058	0.108

cross-sectional data, Xiao et al. (2015) analyzed the systematic and random circadian rhythms of physical activity as functions of time of day and age; see also Goldsmith et al. (2015).

In this analysis, our objective is to quantify the circadian rhythm of physical activity over multiple days and how they relate to age, sex and body mass index (BMI) and provide associated inference. Data were collected from the National Health and Nutrition Examination Survey (NHANES), a large cohort study conducted in two-year waves by the US Centers for Disease Control and Prevention (CDC) to assess
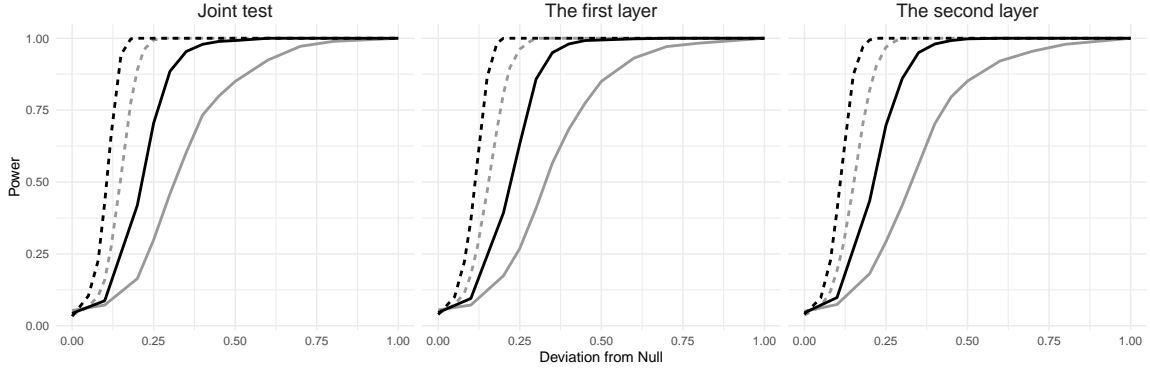
(a) Null: exchangeable covariance
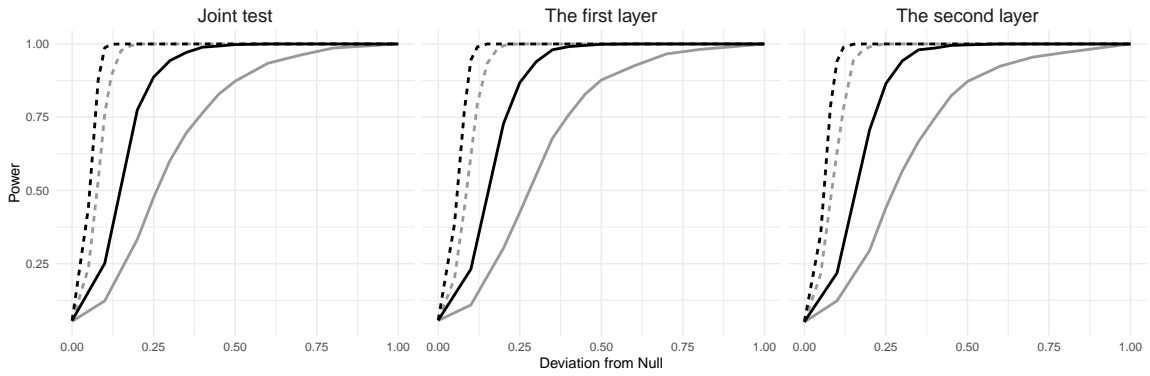
(b) Null: independent covariance

Figure 1: Power curves (type I error $\alpha = 0.05$) for individual and joint tests of covariances, under deviation from the null. Shown are: $m_i \sim \text{Unif}[3, 6]$ (solid line) and $m_i \sim \text{Unif}[8, 12]$ (dashed line), for $n = 100$ (gray) and $n = 200$ (black).

the health and nutritional status of US population. We focus on the NHANES 2011-2012 and 2013-2014 data, where each study participant was asked to wear the wrist-worn device all the time for seven consecutive days. The physical activity data were collected and summarized in minute-level Monitor-Independent Movement Summary (MIMS) units, a physical activity intensity unit optimized to capture normal human motion (John et al., 2019). For each day of the study participant, our initial data are MIMS measured at 1440 minutes from midnight to midnight. For simplicity, the

data are summarized by taking the half-hourly average, resulting in 48 observations that gives the daily activity profile as a function of the day's hour. We include sex (denoted by sex_i), age (age_i), and BMI (BMI_i) as the covariates in this analysis.

The final dataset consists of 513 study participants (208 men and 305 women) with an average age of 57 years and an average BMI value of 28. For each study participant, we further subset the days with at least 95% estimated wear time, leading to a total of 3187 days of wear with an average of 6.21 days of wear per study participant. For subject i at day j, the average MIMS at hour s of a day is our longitudinal functional outcome, denoted by $y_{ij}(s)$.

We specify the mean function as $\mu_{ij}(s) = \beta_0(s) + \text{sex}_i\beta_{\text{sex}}(s) + \text{age}_i\beta_{\text{age}}(s) + \text{BMI}_i\beta_{\text{BMI}}(s)$. The variable sex_i is 0 for males and 1 for females. Both age and BMI are centered so that $\beta_0(s)$ is the population mean MIMS corresponding to males of age 57 and with BMI of 28. The longitudinal covariate T_{ij} is the day of wear. Considering the coefficient function can be thought of as on a circle where the end points of it should actually be equivalent(hour of day), the mean estimation is conducted using cyclic P-splines. To investigate the longitudinal correlation structure of the data, we first estimate the marginal covariance function, which is estimated via the *fpca.face* R function using cubic B-spline basis functions with 10 knots. Using the procedure described in Section 2.2, we identify that seven eigenfunctions explain 90% of the marginal variance. Using these seven eigenfunctions, we conduct a joint test for the null hypothesis that all longitudinal covariances $G_{k0}(\cdot, \cdot)$ can be induced by an exchangeable covariance between daily profiels within each subject. The p-values are 0.042, 0.259, 0.039, 0.500, 0.454, 0.204 and 0.593, respectively. Thus, with the Bonferroni correction, we do not reject the null hypothesis at the 0.05 level. Thus, for

model fit, we use exchangeable covariance structure for all longitudinal covariances.

The estimated coefficient functions are displayed in Figure 2, along with the 95% pointwise confidence bands. The top left panel in Figure 2 displays the estimated mean activity profile $\beta_0(s)$ for males with age 57 and BMI 28. The shape is consistent with that of published average activity plots over the course of the day. On average, individuals are less active during night. Then, their activities increase sharply between 5AM and 11AM, sustain between 11AM and 7PM, and decrease rapidly after 7PM. The top right panel indicates that females in this age group (50 to 70) have significantly more activities than males throughout most of the day except for midnight. The bottom left panel displays the estimated coefficient function for age, indicating a loss of activity for older individuals from afternoon to 1AM. There is no substantial age effect in activity in the morning while the most considerable age effect happens in the early night. These findings are consistent with common sense that older individuals sleep early at night and wake up early in the morning, while younger individuals are more energetic at night. The bottom right panel indicates that activity patterns tend to decrease with increasing BMI, especially in the daytime.

Figure 3 displays the estimated mean activity profiles for females and males, with different age and BMI combinations. From the top panels, we see that for a given BMI value, there is a pronounced decrease in activity as a function of age, both for females and males. Moreover, younger individuals have on average two activity peaks, while older individuals tend to lose the second activity peak. It indicates that the change of physical activity patterns with age is due to less activity in the afternoon and at nigh. The bottom panels indicate that the average activity profile decreases with increasing BMI. The negative effect of age only works after noon while
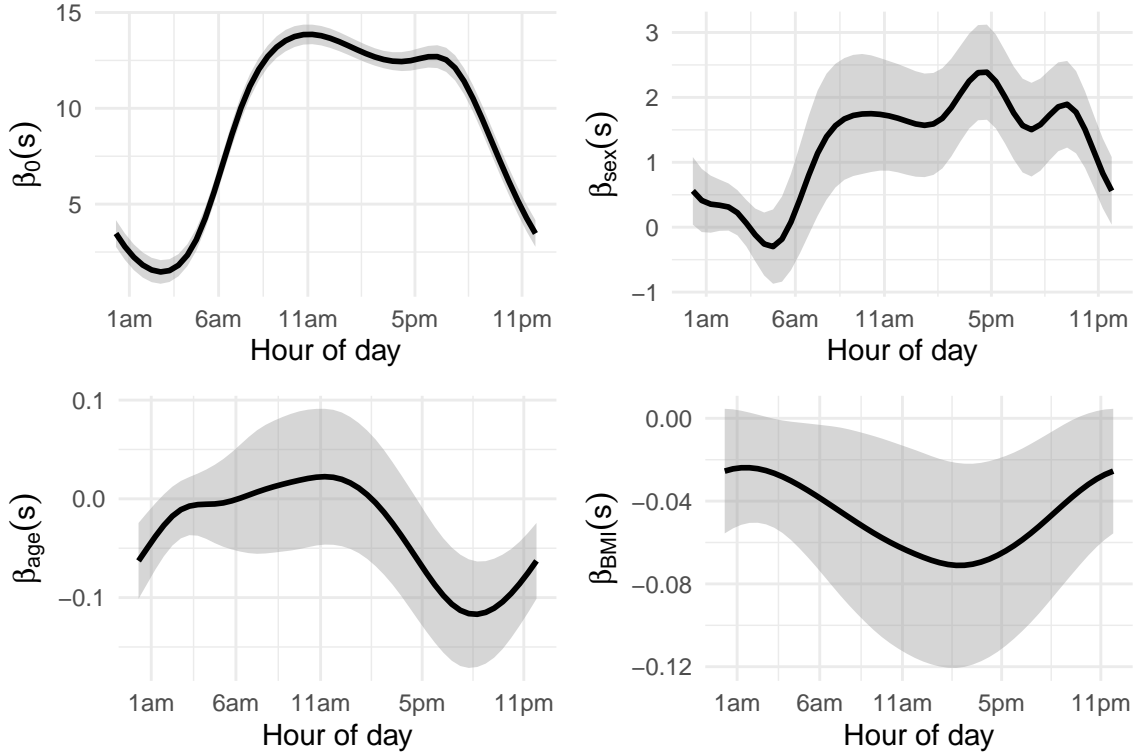
Figure 2: Estimated coefficient functions for the NHANES dataset along with point-wise 95% confidence bands under exchangeable longitudinal covariance structure.

individuals with larger BMI have less activity during all daytime. Moreover, the effect of two units on the BMI scale is roughly equivalent to two years of age in terms of overall activity loss.

6   Discussion

We introduced an inferential framework for estimating fixed effects functions for longitudinal functional data with complex correlation structures. Methods are natural extensions of standard longitudinal data methods with scalar observations and use the flexible marginal decomposition model framework in Park and Staicu (2015).
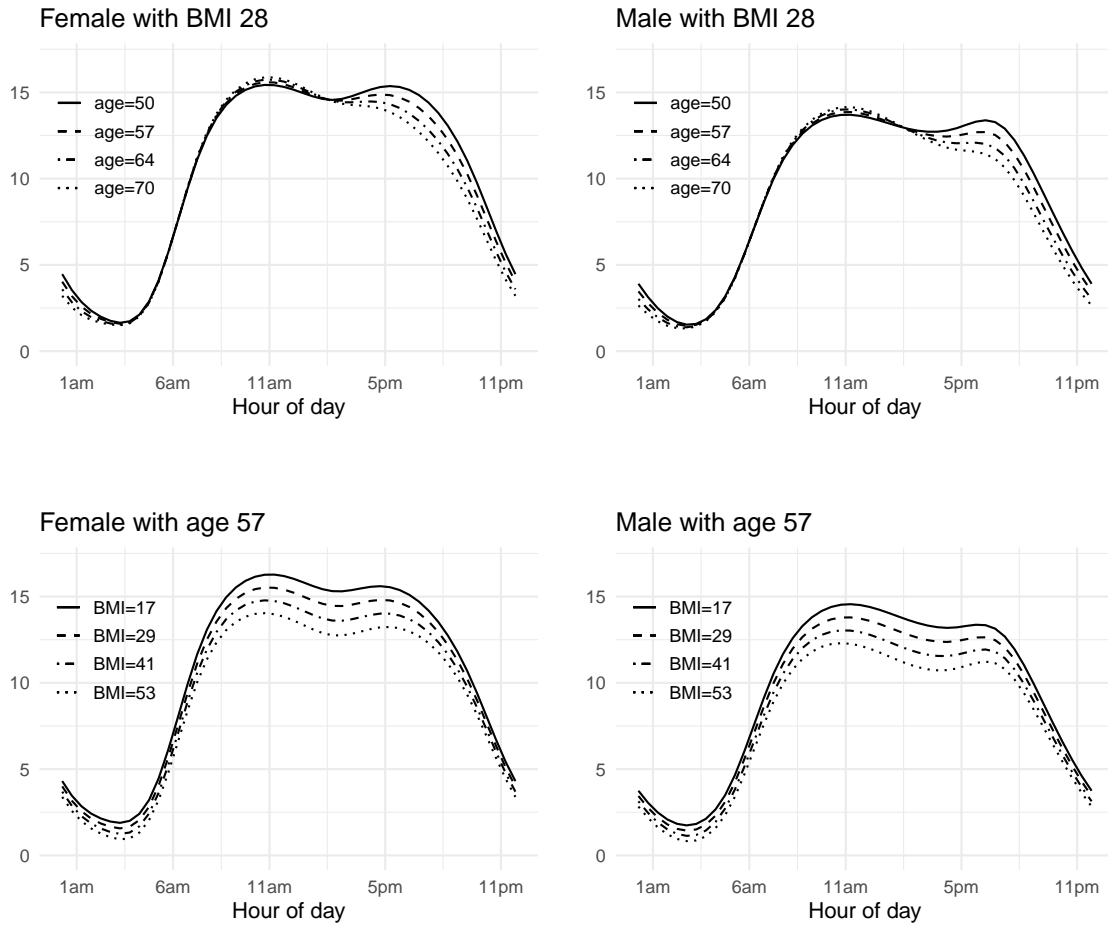
Figure 3: Top panels: estimated mean activity profiles of females and males with BMI 28 at four different ages; bottom panels: estimated mean activity profiles of females and males with age 57 and with four different BMI values.

Simulation studies indicate that the estimation procedure works remarkably well even when the longitudinal correlations are misspecified. However, the confidence bands obtained under misspecified correlation models may have smaller than nominal coverage or are wider than necessary. Our results indicate that using unspecified correlation matrices is the safer approach when it is computationally feasible. More work will be necessary to improve the computational feasibility of these methods for larger sample sizes in terms of the number of study participants, visits, and observations per curve.

We also introduced a testing framework for longitudinal correlations for longitudinal functional data, which extended the testing method in Chen et al. (2019). The proposed tests provide insight into what correlation structures are appropriate and may help reduce the computational cost when a simpler longitudinal correlation is chosen.

The computer code to demonstrate the proposed method and a RMD file with details can be obtained from `https://github.com/rli20ST758/FILF` .

References

Chen, K. and Müller, H.-G. (2012). Modeling repeated functional observations. *Journal of the American Statistical Association*, **107**(500), 1599–1609.

Chen, K., Delicado, P., and Müller, H.-G. (2017). Modelling function-valued stochastic processes, with applications to fertility dynamics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **79**(1), 177–196.

Chen, S., Xiao, L., and Staicu, A.-M. (2019). A smoothing-based goodness-of-fit test of covariance for functional data. *Biometrics*, **75**(2), 562–571.

Crainiceanu, C. M., Staicu, A. M., Ray, S., and Punjabi, N. (2012). Bootstrap-based inference on the difference in the means of two correlated functional processes. *Stat Med*, **31**(26), 3223–3240.

Di, C.-Z., Crainiceanu, C. M., Caffo, B. S., and Punjabi, N. M. (2009). Multilevel functional principal component analysis. *Ann. Appl. Stat.*, **3**(1), 458–488.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical Science*, **11**(2), 89–121.

Goldsmith, J., Bobb, J., Crainiceanu, C. M., Caffo, B., and Reich, D. (2011). Penalized functional regression. *Journal of Computational and Graphical Statistics*, **20**(4), 830–851.

Goldsmith, J., Zipunnikov, V., and Schrack, J. (2015). Generalized multilevel function-on-scalar regression and principal component analysis. *Biometrics*, **71**(2), 344–353.

Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Gellar, J., Harezlak, J., McLean, M. W., Swihart, B., Xiao, L., Crainiceanu, C., and Reiss, P. T. (2018). *refund: Regression with Functional Data.*

Greven, S., Crainiceanu, C., Caffo, B., and Reich, D. (2010). Longitudinal functional principal component analysis. *Electron. J. Statist.*, **4**, 1022–1054.

Ivanescu, A. E., Staicu, A.-M., Scheipl, F., and Greven, S. (2015). Penalized function-on-function regression. *Computational Statistics.*, **30**(2), 539–568.

John, D., Tang, Q., Albinali, F., and Intille, S. (2019). An open-source monitor-independent movement summary for accelerometer data processing. *Journal for the Measurement of Physical Behaviour*, **2**(4), 268–281.

Lee, W., Miranda, M. F., Rausch, P., Baladandayuthapani, V., Fazio, M., Downs, J. C., and Morris, J. S. (2019). Bayesian semiparametric functional mixed models for serially correlated functional data, with application to glaucoma data. *Journal of the American Statistical Association*, **114**(526), 495–513.

Park, S. Y. and Staicu, A.-M. (2015). Longitudinal functional data analysis. *Statistics*, **4**(1), 212–226.

Park, S. Y., Staicu, A.-M., Xiao, L., and Crainiceanu, C. M. (2018). Simple fixed-effects inference for complex functional models. *Biostatistics*, **19**(2), 137–152.

Ramsay, J. (2005). *Functional Data Analysis*. American Cancer Society.

Scheipl, F., Staicu, A.-M., and Greven, S. (2015). Functional additive mixed models. *Journal of Computational and Graphical Statistics*, **24**(2), 477–501.

Scheipl, F., Gertheiss, J., and Greven, S. (2016). Generalized functional additive mixed models. *Electronic Journal of Statistics*, **10**(1), 1455–1492.

Wood, S. (2019). *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation.*

Wood, S. N., Goude, Y., and Shaw, S. (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **64**(1), 139–155.

Wood, S. N., Pya, N., and Safken, B. (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, **111**(516), 1548–1575.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics*, **14**(4), 1261–1295.

Xiao, L., Li, Y., and Ruppert, D. (2013). Fast bivariate *P*-splines: the sandwich smoother. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **75**(3), 577–599.

Xiao, L., Huang, L., Schrack, J., Ferrucci, L., Zipunnikov, V., and Crainiceanu, C. (2015). Quantifying the lifetime circadian rhythm of physical activity: A covariate-dependent functional approach. *Biostatistics*, **16**, 352–367.

Xiao, L., Zipunnikov, V., Ruppert, D., and Crainiceanu, C. (2016). Fast covariance estimation for high-dimensional functional data. *Statistics and Computing*, **26**(1), 409–421.

Xiao, L., Li, C., Checkley, W., and Crainiceanu, C. (2018). Fast covariance estimation for sparse functional data. *Statistics and Computing*, **28**(3), 511–522.

Zhang, L., Baladandayuthapani, V., Zhu, H., Baggerly, K. A., Majewski, T., Czerniak, B. A., and Morris, J. S. (2016). Functional car models for large spatially correlated functional datasets. *Journal of the American Statistical Association*, **111**(514), 772–786.

Zhu, H., Chen, K., Luo, X., Yuan, Y., and Wang, J.-L. (2019). FMEM: Functional mixed effects models for longitudinal functional responses. *Statistica Sinica*, **29**(4), 2007–2033.