Published in final edited form as:

Stat Med. 2024 November 10; 43(25): 4781–4795. doi:10.1002/sim.10207.

Multilevel Longitudinal Functional Principal Component Model

Wenyi Lin[1], **Jingjing Zou**[1], **Chongzhi Di**[2], **Cheryl L. Rock**[3], **Loki Natarajan**[1]

[1]Division of Biostatistics, Herbert Wertheim School of Public Health and Longevity Science, University of California San Diego, La Jolla, California, USA

[2]Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, Washington, USA

[3]Division of Preventive Medicine, Department of Family Medicine, School of Medicine, University of California San Diego, La Jolla, California, USA

Abstract

Sensor devices, such as accelerometers, are widely used for measuring physical activity (PA). These devices provide outputs at fine granularity (e.g., 10–100 Hz or minute-level), which while providing rich data on activity patterns, also pose computational challenges with multilevel densely sampled data, resulting in PA records that are measured continuously across multiple days and visits. On the other hand, a scalar health outcome (e.g., BMI) is usually observed only at the individual or visit level. This leads to a discrepancy in numbers of nested levels between the predictors (PA) and outcomes, raising analytic challenges. To address this issue, we proposed a multilevel longitudinal functional principal component analysis (mLFPCA) model to directly model multilevel functional PA inputs in a longitudinal study, and then implemented a longitudinal functional principal component regression (FPCR) to explore the association between PA and obesity-related health outcomes. Additionally, we conducted a comprehensive simulation study to examine the impact of imbalanced multilevel data on both mLFPCA and FPCR performance and offer guidelines for selecting optimal methods.

Keywords

functional principal component analysis; functional regression; unbalanced study design

1 | Introduction

Accelerometer-based devices are commonly used to characterize physical activity (PA) in health behavior research, including in longitudinal studies and clinical trials [1]. These devices, such as Actigraph GT3X, are able to provide estimates at minute-level or even

Correspondence: Loki Natarajan (lnatarajan@health.ucsd.edu).

Conflicts of Interest

The authors declare no conflicts of interest.

Supporting Information

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

Hz-level. In order to acquire sufficient information to characterize an individual's typical weekly activity, participants are required to wear the device at least 10 h per day for 5–7 consecutive days at each visit [2-6]. However, other characteristics, such as related health outcomes, are normally measured once per subject or visit. Therefore, the flexibility of accelerometer-based devices allows for a rich amount of data to be collected, but also brings challenges when more frequent high-dimensional PA data are acquired, yet outcomes are measured relatively infrequently. In this article, we aim to address this challenge. We propose a multilevel longitudinal functional principal component analysis (mLFPCA) model to address the case when PA is measured more frequently than related health outcomes in a longitudinal study.

The motivating dataset comes from the Metabolism, Exercise, and Nutrition at UCSD (MENU) trial, a 12-month behavioral intervention longitudinal study consisting of 245 overweight nondiabetic women [5, 6]. PA was recorded with GT3X+ Actigraph monitors, set to collect data at 30 Hz [7], for about a week per subject at each clinical visit. Each day's PA record for a subject and a visit is a densely and finely sampled function across the time interval. Therefore, these PA data are considered to have a three-level hierarchical structure under a longitudinal study design. Figure 1 displays an example of one subject's daily PA records on three random days obtained at baseline, 6 and 12 months. Scalar health outcomes, such as body mass index (BMI), were acquired at subject and visit level, thus having a two-level structure. In fact, the study is one example of many biomedical studies, either having cross-sectional or longitudinal study designs, where the predictors are more frequently observed than the outcomes. The goal of this study is to explore approaches for analyzing studies with these types of unbalanced study designs, typically with accelerometer-measured functional predictors and scalar outcomes.

There are several existing statistical approaches for analyzing accelerometer data. The most common one is to derive algorithms that can translate the dense signals into summary metrics, such as total (or average) daily time spent being active or minutes of activity with varying intensities [8-10]. For instance, sedentary behaviors are defined as activity with less than 100 counts/min [11] and moderate to vigorous physical activity (MVPA) time is defined as minutes with activity counts >2020 [9]. As such, it is common in health behavior studies to retain the original input data but aggregate the daily inputs over a week to obtain weekly averages or summary statistics, such as the average daily intake of dietary fat [12] or minutes of PA [13]. As one example, Bürgi et al. [14] investigated the cross-sectional and longitudinal relationship of PA with body fat and other health outcomes for preschool children, using total PA, moderate PA and vigorous PA summarized from at least 3 days of PA recording. Although these metrics provide useful summaries of overall activity, they lose the ability to capture the correlation over time within a subject's activity profile. Specifically, while this summary can provide a representation of an individual's overall behavior, it ignores the temporal patterns of PA accumulation and inevitably results in some loss of information.

Functional data analysis (FDA) is a set of powerful statistical techniques for exploiting the rich and dense time-series accelerometer outputs such as the minute-by-minute temporal patterns of PA. As was summarized in Ramsay and Silverman [15], FDA treats a sequence

of observations, such as the daily activity profile curve in our case, as a single unit rather than disjoint minutes spent in varying types of activity. Advancements in FDA methodology and applications have increased over the last decade given the novel imaging and wearable sensor technologies that are now available. Works by Wang, Chiou, and Müller [16] and Shang [17] provide insightful and comprehensive reviews. Our proposed model is based on functional principal component analysis (FPCA), which is performed on densely-sampled PA data to obtain the principal directions of variation and achieve dimension reduction [18, 19]. Current studies of FPCA have extended its application in modeling multilevel functional data [20, 21], longitudinal functional data [22] and longitudinal associations with scalar outcomes [23, 24], based on the corresponding study design or data structure. Our previous works [25, 26] have implemented these methods on the MENU study. For example, Xu et al. [25] applied a two-level FPCA model and explored the cross-sectional association between extracted principal component scores and health outcomes. Lin et al. [26] further explored the longitudinal association by means of longitudinal FPCA modeling. However, neither of these approaches is optimal for the study design, since the PA data have repeated (daily) measures across multiple visits. The issue was previously addressed by taking the average of the daily records or selecting one single day at each visit, which may cause a loss of information within the corresponding level. Therefore, we propose the mLFPCA approach by extending the longitudinal FPCA model in Greven et al. [22] to accommodate the situation where additional levels of predictor inputs were observed in a longitudinal study.

Another research question of interest in this article is to assess the regression performance of functional predictors with health outcomes when functional predictors are measured more frequently than scalar outcomes. In data applications, models are often misspecified due to lack of background information or the preference for simplified models (e.g., taking averages or summary statistics of PA). In particular, misspecification may happen at the stage of model construction or data preprocessing. To our knowledge in the context of longitudinal multilevel functional data, no prior work has addressed the effects of misspecification on model performance, in both cross-sectional and longitudinal studies. In this paper, we provide a comprehensive investigation of this question using both mathematical derivation and simulation studies.

We organize this article as follows. Section 3 describes existing methods in the field of FPCA modeling, Section 4 then provides our proposed mLFPCA. Sections 5 and 6 illustrate the estimation procedure of our model and its comparison with misspecified models. Section 7 shows the performance of our model and existing methods with extensive simulation studies. Section 8 presents the application of the multilevel longitudinal FPCA methods to the MENU study, Section 9 concludes the article with a discussion.

2 | Notation

The observed PA records are functional data $\boldsymbol{X}_{ijk} = \{X_{ijk}(t), t \in \mathcal{D}\}$, which are random functions in $L^2[0, 1]$ measured at minute-level time t on a set of grid points $\mathcal{D}$ with length D, for subject $i = 1, 2, \ldots, n$ at visit $j = 1, 2, \ldots, n_i$ and day $k = 1, 2, \ldots, n_{ij}$. The total number of observations are denoted as $I = \sum_{i,j} n_{ij}$ and the number of observations for subject i is

$I_i = \sum_j n_{ij}$. In general, this subject-visit-day structure can be extended to other types of data which have similar nested structures.

3 | Overview of Existing FPCA Models

In this section, we summarize previous work by Di et al. [20], Greven et al. [21], and Shou et al. [22], to illustrate the development of FPCA models when more complex data structures were involved. In later sections, we also implemented these models in our simulation studies and compared them with our proposed method. FPCA plays an important role in FDA, whose basic purpose is to decompose the functional curves into principal directions of variation. For notational simplicity, we use $X_i(t)$, $X_{ij}(t)$ and $X_{ijk}(t)$, $t \in \mathcal{D}$, to denote the one-, two-, and three-level functional inputs, respectively, where the hierarchical structure of the data can be analogous to subject i, visit j and day k.

Assuming no measurement error, in the simplest one-level setting, $X_i(t)$ can be represented as,

$$X_i(t) = \mu(t) + U_i(t) \tag{1}$$

where $\mu(t)$ is the overall population mean function at t and $U_i(t)$ is the subject-specific deviation from the overall mean function. Specifically, $\mu(t)$ is a deterministic function and $U_i(t)$ are assumed to be i.i.d. stochastic processes with mean zero and covariance function $K_U(s,t) = cov\{U_i(s), U_i(t)\}$. By Mercer's theorem [27], the spectral decomposition is provided as $K_U(s,t) = \sum_{l=1}^{\infty} \lambda_l \phi_l^U(s)\phi_l^U(t)$, where $\lambda_1 \geq \lambda_2 \geq \ldots$ are ordered non-negative eigenvalues and ϕ_l^U are corresponding orthogonal eigenfunctions. Using the Karhunen–Loève (KL) expansion [28], Model (1) becomes $X_i(t) = \mu(t) + \sum_{l=1}^{\infty} \xi_{il}\phi_l^U(t)$, where $\xi_{il} = \int U_i(s)\phi_l^U(t)dt$ are uncorrelated principal component scores with mean zero and variance λ_l.

Di et al. [20] expanded the one-level FPCA model to a two-level FPCA when the data $X_{ij}(t)$ are measured at both subject- and visit-level. The decomposition has the form

$$X_{ij}(t) = \mu(t) + U_i(t) + V_{ij}(t) \tag{2}$$

where $U_i(t)$ is the subject-specific (Level 1) deviation from the overall mean function and $V_{ij}(t)$ is the subject- and visit-specific (Level 2) deviation from the subject-mean function. It is assumed that $U_i(t)$ and $V_{ij}(t)$ are uncorrelated stochastic processes with zero mean and continuous covariance functions. $K_U(s,t) = cov\{U_i(s), U_i(t)\}$ and $K_V(s,t) = cov\{V_{ij}(s), V_{ij}(t)\}$ are covariance functions for the above random processes. Therefore, the proposed two-level FPCA model extended the original FPCA methods to fit data with multilevel structure.

Greven et al. [22] further extended the two-level FPCA to longitudinal FPCA (LFPCA-lv2), analogous to a classical longitudinal model but in functional format, which has the form,

$$X_{ij}(t) = \mu(t) + U_{i0}(t) + U_{i1}(t)T_{ij} + V_{ij}(t) \tag{3}$$

where $U_{i0}(t)$ is the random functional intercept and $U_{i1}(t)$ is the random functional slope for subject i, respectively, and T_{ij} is the time at visit j for subject i, which can be either the visit indicator with $T_{ij} = j$ or a continuous time variable. One major difference between the longitudinal FPCA and multilevel FPCA is the construction of the subject-specific variation $K_U(s,t)$, which in the longitudinal setting is the covariance function between the bivariate process $\boldsymbol{U}_i(t) = (U_{i0}(t), U_{i1}(t))$ and has two parts: the auto-covariance $K_{U_0}(s,t)$, $K_{U_1}(s,t)$ and the cross-covariance $K_{U_{01}}(s,t)$. By incorporating both the random intercept and slope processes, the longitudinal FPCA model can be considered as the longitudinal generalization of previous models.

Additional levels of data structure were considered by Shou et al. [21], referred to as structured FPCA or multilevel FPCA (MFPCA). Specifically, with three-level data $X_{ijk}(t)$, the three-way FPCA model decomposes the data into three parts, subject-specific process $U_i(t)$, visit-specific process $V_{ij}(t)$ and day-specific process $W_{ijk}(t)$, which can be written as,

$$X_{ijk}(t) = \mu(t) + U_i(t) + V_{ij}(t) + W_{ijk}(t) \tag{4}$$

where $U_i(t)$ is the subject-specific (Level 1) process, $V_{ij}(t)$ is the subject-visit-specific (Level 2) deviation and $W_{ijk}(t)$ quantifies the daily (Level 3) deviation from the subject- and visit-mean function. $U_i(t)$, $V_{ij}(t)$, and $W_{ijk}(t)$ are mutually uncorrelated random processes with mean zero and covariance functions K_U, K_V and K_W respectively. Therefore, the variability of $X_{ijk}(t)$ is fully determined by processes $U_i(t)$, $V_{ij}(t)$, and $W_{ijk}(t)$, that is, $K_X = K_U + K_V + K_W$. The MFPCA models provided a set of efficient tools to analyze data with any number of levels of nested designs.

We note that $\mu(t)$, the overall population mean in the models described in Equations (2), (3), and (4) could be generalized to a mean surface that incorporates the nested levels (e.g., $\mu(t, T_{ij})$ representing the mean surface over t and visit), but for simplicity we suppress this term, although it can be easily incorporated as needed when fitting models.

All of the above mentioned methods and related papers provided an extensive exploration of functional data with multilevel or longitudinal structures. However, these models do not address the issue where both longitudinal and multilevel (more than two levels) data coexist in the study. We proposed our mLFPCA model by extending the previously introduced FPCA models, for accommodating data structures that are common in health behavior studies (including our MENU study), that is, a longitudinal study with repeated day-level records, that is, day k at visit j, for each subject i. In the proposed model, we incorporated both subject-level random intercept and slope processes as longitudinal components, as well as uncorrelated visit- and day-level processes as additional hierarchical components.

4 | Multilevel Longitudinal Functional Principal Component Model

In this section, we describe the analytic details of the proposed mLFPCA model. Let $\boldsymbol{U}_i(t) = (U_{i0}(t), U_{i1}(t))$, $V_{ij}(t)$ and $W_{ijk}(t)$ be mutually uncorrelated random processes with mean zero as described in Section 3.

We assume that $U_{i0}(t)$ and $U_{i1}(t)$ have covariance functions $K_{U_0}(s,t)$ and $K_{U_1}(s,t)$, respectively, and cross-covariance function $K_{U_{01}}(s,t)$ with $K_U(s,t)$, representing the covariance function of $U_i(t)$; $V_{ij}(t)$ has covariance function $K_V(s,t)$ and $W_{ijk}(t)$ has covariance function $K_W(s,t)$. The proposed mLFPCA model is,

$$X_{ijk}(t) = \mu(t) + U_{i0}(t) + U_{i1}(t)T_{ij} + V_{ij}(t) + W_{ijk}(t) \tag{5}$$

which in fact is a natural generalization of the longitudinal FPCA and MFPCA models, since the proposed model simultaneously decomposes the multilevel data both longitudinally and hierarchically. Similarly, we obtain the KL expansion as,

$$X_{ijk}(t) = \mu(t) + \sum_{l=1}^{\infty} (1, T_{ij})\xi_{il}\phi_l^U(t) + \sum_{m=1}^{\infty} \zeta_{ijm}\phi_m^V(t) + \sum_{r=1}^{\infty} \eta_{ijkr}\phi_r^W(t) \tag{6}$$

where $\phi_l^U(t) = (\phi_l^{U_0}(t), \phi_l^{U_1}(t))'$, $\phi_m^V(t)$, and $\phi_r^W(t)$ are the eigenfunctions of the covariance operators associated with the covariance functions K_U, K_V, and K_W, with corresponding eigenvalues λ_l^U, λ_m^V, and λ_r^W, respectively. $\xi_{il} = \int U_{i0}(s)\phi_l^{U_0}(s)ds + \int U_{i1}(s)\phi_l^{U_1}(s)ds$, $\zeta_{ijm} = \int V_{ij}(s)\phi_m^V(s)ds$, and $\eta_{ijkr} = \int W_{ijk}(s)\phi_r^W(s)ds$ are uncorrelated random variables with mean zero and variance λ_l^U, λ_m^V, and λ_r^W, respectively. Since writing out an infinite expansion is not feasible, we consider finite-dimensional approximations of processes $\boldsymbol{U}$, V, and W, such that most of the variability of each process is captured by the first N_U, N_V, and N_W principal components,

$$X_{ijk}(t) = \mu(t) + \sum_{l}^{N_U} (1, T_{ij})\xi_{il}\phi_l^U(t) + \sum_{m}^{N_V} \zeta_{ijm}\phi_m^V(t) + \sum_{r}^{N_W} \eta_{ijkr}\phi_r^W(t) \tag{7}$$

5 | Estimation

Having defined the mLFPCA model, the next step is estimation of the various model components, which we describe in the next few subsections. We assume $X_{ijk}(t)$ are measured on a set of grid points $\mathcal{D}$ with finite length D. Missing data, either in terms of visits or days, can be easily handled with our method. Estimation can be achieved via the following steps, and more details are provided in the next few sections.

Step 1. Estimating the mean function $\mu(t)$ by the sample average $\hat{\mu}(t) = \frac{1}{I}\sum_{i,j,k} X_{ijk}(t)$. Denote the centered data as $\widetilde{X}_{ijk}(t) = X_{ijk}(t) - \hat{\mu}(t)$.

Step 2. Estimating the covariance function $\widehat{K}_W$ for the W process from $\widetilde{X}_{ijk}(t)$ via method of moment (MoM) estimators [21, 28].

Step 3. Estimating the covariance functions $\widehat{K}_U$ for $\boldsymbol{U}_i = (U_{i0}, U_{i1})$ and $\widehat{K}_V$ for V_{ij} via mixed linear regression models.

Step 4. Performing eigen decompositions of the estimated covariance functions to provide bases for representing $\boldsymbol{U}_i = (U_{i0}, U_{i1})$, V_{ij} and W_{ijk}.

Step 5. Estimating the best linear unbiased prediction (BLUP) to provide subject-, visit- and day-specific principal component scores.

5.1 | Estimation of the Mean and Covariance Functions

The fixed population mean function $\mu(t)$ is estimated by taking the sample mean in our implementation. More generally, when the observations across visits and subjects are relatively dense, a bivariate smoother in t (time in minutes) and T (visit) may be considered to form the mean surface $\mu(t, T)$, such as penalized spline smoothers [22]. Similarly, for sparser collection of T_{ij}, $\mu(t)$ can be approximated via the univariate smoother $\mu_j(t)$ for visit j. With the estimated mean function $\hat{\mu}(t)$ from any of the aforementioned methods, data are centered via $X_{ijk}(t) - \hat{\mu}(t)$ and without loss of generality, we assume that $X_{ijk}(t)$ has mean zero.

The main challenge of our proposed method is to estimate the covariance functions $K_U = \begin{pmatrix} K_{U_0} & K_{U_{01}} \\ K_{U_{01}} & K_{U_1} \end{pmatrix}$, K_V, and K_W. Under the setup and assumptions of Model (5), for all i, j, j', k, k', s, t, we have

$$\begin{aligned} Cov(X_{ijk}(s), X_{ij'k'}(t)) = {} & E(X_{ijk}(s)X_{ij'k'}(t)) \\ = {} & Cov(U_{i0}(s), U_{i0}(t)) \\ & + T_{ij}Cov(U_{i0}(s), U_{i1}(t)) \\ & + T_{ij'}Cov(U_{i0}(t), U_{i1}(s)) \\ & + T_{ij}T_{ij'}Cov(U_{i1}(s), U_{i1}(t)) \\ & + Cov(V_{ij}(s), V_{ij'}(t)) \\ & + Cov(W_{ijk}(s), W_{ij'k'}(t)) \end{aligned} \tag{8}$$

The estimation of these covariance functions is not straightforward, and we cannot simply apply the method of Greven et al. [22], that is, linearly regressing the left side "outcome" $X_{ijk}(s)X_{ij'k'}(t)$ on the right side "covariates." This is because the total number of the observations $I = \sum_{i,j} n_{ij}$ in the three-level model can be much larger than that in a two-level setup. Let δ denote the Kronecker's delta, defined as $\delta_{kk'} = \begin{cases} 1, & \text{if } k = k' \\ 0, & \text{otherwise} \end{cases}$. Specifically, to compute least squared estimates for these covariance matrices, a matrix $\boldsymbol{H} = \boldsymbol{F}'(\boldsymbol{F}\boldsymbol{F}')^{-1}$ needs to be constructed, where $\boldsymbol{F}_{6\times m}$ has the column equal to $\boldsymbol{f}_{ijj'kk'} = (1, T_{ij}, T_{ij'}, T_{ij}T_{ij'}, \delta_{jj'}, \delta_{kk'})'$ and $m = \sum_{i=1}^{n} n_i^2 \sum_{j,j'} n_{ij}n_{ij'}$. Hence it is impractical and computationally inefficient to fit a regression model at once. Therefore, we proposed a two-step procedure to estimate these

covariance estimators, combining the MoM estimators and regression strategy. Equation (8) can be rewritten as,

$$E(X_{ijk}(s)X_{ij'k'}(t)) = \begin{cases} K_{U_0}(s,t) + 2T_{ij}K_{U_{01}}(s,t) + T_{ij}^2K_{U_1}(s,t) \\ \quad + K_V(s,t) + \delta_{kk'}K_W(s,t), & \text{if } j = j' \\ K_{U_0}(s,t) + T_{ij}K_{U_{01}}(s,t) + T_{ij'}K_{U_{01}}(t,s) \\ \quad + T_{ij}T_{ij'}K_{U_1}(s,t), & \text{otherwise} \end{cases} \tag{9}$$

We first estimate the day-level covariance function K_W using a MoM estimator,

$$\widehat{K}_W = \frac{1}{\sum_{i,j} n_{ij}(n_{ij}-1)} \sum_{i,j} \sum_{k,k'} (\boldsymbol{X}_{ijk} - \boldsymbol{X}_{ijk'})(\boldsymbol{X}_{ijk} - \boldsymbol{X}_{ijk'})^T \tag{10}$$

Substituting K_W with the empirical estimator $\widehat{K}_W$ and subtracting this term from both sides of the first line of Equation (9), the day-to-day variation is eliminated from the total variation. The remaining proportion of variation therefore only involves variations at subject and visit levels. Denote $\widetilde{\boldsymbol{X}_{ij\cdot}\boldsymbol{X}_{ij'}^T}$, as the resulting residual variance, the estimators of the covariance functions of the U and V processes can be expressed as,

$$E(\widetilde{\boldsymbol{X}_{ij\cdot}\boldsymbol{X}_{ij'\cdot}^T}) = K_{U_0} + T_{ij}K_{U_{01}} + T_{ij'}K_{U_{01}} + T_{ij}T_{ij'}K_{U_1} + \delta_{jj'}K_V \tag{11}$$

By implementing the two-step procedure, we are able to reduce the dimension of the "outcome" variable in the regression modeling and essentially reduce it to the two-level case. The computational feasibility for this two-level case has been proved in Greven et al. [22], and we then regress the product $\widetilde{\boldsymbol{X}_{ij\cdot}\boldsymbol{X}_{ij'}^T}$, on "predictors" (1, T_{ij}, $T_{ij'}$, $T_{ij}T_{ij'}$, $\delta_{jj'}$) and get the estimated covariance estimators ($\widehat{K}_{U_0}$, $\widehat{K}_{U_{01}}$, $\widehat{K}_{U_1}$, $\widehat{K}_V$).

5.2 | Estimation of Eigenfunctions and Scores

With the estimated covariance functions from the two-step procedure, $\widehat{K}_U = \begin{pmatrix} \widehat{K}_{U_0} & \widehat{K}_{U_{01}} \\ \widehat{K}_{U_{01}} & \widehat{K}_{U_1} \end{pmatrix}$, $\widehat{K}_V$, and $\widehat{K}_W$ in the previous section, using the spectral decomposition, we can easily estimate the eigenvalues λ_l^U, λ_m^V, and λ_r^W, and eigenfunctions $\phi_l^U(t)$, $\phi_m^V(t)$, and $\phi_l^W(t)$, $t \in \mathscr{D}$, at grid points D, that is, $\widehat{K}_U = \sum_{l=1}^{2D} \hat{\lambda}_l \hat{\phi}_l^U (\hat{\phi}_l^U)^T$, $\widehat{K}_V = \sum_{m=1}^{D} \hat{\lambda}_m \hat{\phi}_r^V (\hat{\phi}_r^V)^T$ and $\widehat{K}_W = \sum_{r=1}^{D} \hat{\lambda}_r \hat{\phi}_r^W (\hat{\phi}_r^W)^T$. The eigenfunctions, $\phi_l^U = \{\phi_l^{U_0}(t), \phi_l^{U_1}(t), t \in \mathscr{D}\}$ form an orthonormal basis in $L^2[0,1] \times L^2[0,1]$, and ϕ_m^V and ϕ_r^W are orthogonal vectors in $L^2[0,1]$. If the time variable T_{ij} is standardized to have zero mean and unit variance, that is, $E(T_{ij}) = 0$ and $Var(T_{ij}) = 1$, the variation in $X_{ijk}(t)$ can be decomposed additively and expressed with respect to the estimated eigenvalues, $\int_{\mathscr{D}} var(X_{ijk}(t))dt = \sum_l \lambda_l^U + \sum_m \lambda_m^V + \sum_r \lambda_r^W$. This result has been proved in a two-

level longitudinal FPCA setting by Greven et al. [22] and we extend it to three-level scenario (Section S1). We usually retain a finite numbers of eigenfunctions of subject (N_U), visit (N_V), and day (N_W) levels for further analysis. The number of eigenfunctions retained, N_U, N_V, N_W, is based on a prespecified percentage of explained variation. In our application, we set the threshold to be 95% of explained variation.

For fixed N_U, N_V, and N_W, it is evident that Model (7) is a three-level linear mixed model. Therefore, the principal component scores $\hat{\xi}_{il}$, $\hat{\zeta}_{ijm}$, and $\hat{\eta}_{ijkr}$ can be obtained via the best linear unbiased prediction (BLUP). Let $\boldsymbol{X}_i = vec\{\boldsymbol{X}_{i11}, \ldots, \boldsymbol{X}_{i1n_{i1}}, \ldots, \boldsymbol{X}_{ij1}, \ldots \boldsymbol{X}_{ijn_{ij}}\}$ be a vector with stacked functional inputs for subject i with length $D \times I_i$ and $\boldsymbol{\beta}_i = (\xi_{i1}, \ldots, \xi_{iN_U}, \ldots, \zeta_{i11}, \ldots, \zeta_{i1N_V}, \ldots, \zeta_{in_i1}, \ldots, \hat{\zeta}_{in_iN_V}, \eta_{i111}, \ldots, \eta_{i11N_W}, \ldots, \eta_{i1n_in_ij1}, \ldots, \eta_{in_in_ijN_W})$ be the vector of scores to be estimated. The BLUP for $\boldsymbol{\beta}_i$ is given as,

$$\hat{\boldsymbol{\beta}}_i = (\boldsymbol{Z}_i'\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{X}_i \tag{12}$$

where $\boldsymbol{Z}_i = [\mathbf{1}_{I_i} \otimes \boldsymbol{\Phi}^{U_0} + \boldsymbol{T}_i \otimes \boldsymbol{\Phi}^{U_1} \mid \boldsymbol{I}_{n_i} \otimes (\mathbf{1}_{n_{ij}} \otimes \boldsymbol{\Phi}^{V}) \mid \boldsymbol{I}_{I_i} \otimes \boldsymbol{\Phi}^{W}]$, $\boldsymbol{T}_i = (T_{ij}\delta_{jh})_{j,h=1,\ldots,n_i}$, $\boldsymbol{\Phi}^{U_0} = \{\phi_l^{U_0}(t)\}_{t\in\mathcal{D}, l=1,\ldots,N_U}$, $\boldsymbol{\Phi}^{U_1} = \{\phi_l^{U_1}(t)\}_{t\in\mathcal{D}, l=1,\ldots,N_U}$, $\boldsymbol{\Phi}^{V} = \{\phi_l^{V}(t)\}_{t\in\mathcal{D}, l=1,\ldots,N_V}$, $\boldsymbol{\Phi}^{W} = \{\phi_l^{W}(t)\}_{t\in\mathcal{D}, l=1,\ldots,N_W}$, $\boldsymbol{I}$, is the identity matrix, and $\otimes$ denotes the Kronecker product of matrices. δ_{jh} denotes the matrix of Kronecker deltas with $\delta_{jh} = 1$ if $j = h$ and $\delta_{jh} = 0$ otherwise.

5.3 | Comparing Different Models

The proposed mLFPCA model was motivated by prospective studies with functional predictors such as the MENU study [5, 6], which is a longitudinal study in which the functional PA measurements were collected daily for each visit (baseline, follow-up) for each participant, thus encompassing a three-level nested and longitudinal structure. To further illustrate the need for deriving our method, we explored the relationship between our three-level method and previous approaches with the aim of gaining deeper insights. For instance, with the three-level longitudinal data (Equation 5), one may consider ignoring the random functional slope process $U_{i1}(t)$ and applying MFPCA [21] (Equation 4). Alternatively, assuming that data are well aligned, the two-level longitudinal FPCA (LFPCA-lv2) provided in Greven et al. [22] (Equation 3), can be fitted on the mean values of daily measurement curves at each visit ($W_{ijk}(t)$ is ignored). While it is intuitive that either of these simplifications can cause a loss of information, we aimed to demonstrate their concrete effects.

Suppose that we fit the data with a MFPCA model, and decompose the total variance with the technique of symmetric sum MoM estimators from Shou et al. [21]. Rewriting the covariance functions as $E\{X_{ijk}(s) - X_{i'j'k'}(s)\}\{X_{ijk}(t) - X_{i'j'k'}(t)\}^T$, we obtain the following decomposed form,

$$\begin{cases} 2K_W(s,t), & \text{if } i=i', j=j', k\neq k' \\ 2(T_{ij}K_{U_{01}}(s,t)+T_{ij'}K_{U_{01}}(t,s)+T_{ij}T_{ij'}K_{U_1}(s,t)+K_V(s,t) \\ \quad +K_W(s,t)), & \text{if } i=i', j\neq j' \\ 2(K_{U_0}(s,t)+T_{ij}K_{U_{01}}(s,t)+T_{ij'}K_{U_{01}}(t,s)+T_{ij}T_{ij'}K_{U_1}(s,t) \\ \quad +K_V(s,t)+K_W(s,t)), & \text{if } i\neq i' \end{cases} \tag{13}$$

Let $K_{V_U}(s,t)=T_{ij}K_{U_{01}}(s,t)+T_{ij'}K_{U_{01}}(t,s)+T_{ij}T_{ij'}K_{U_1}(s,t)+K_V(s,t)$, which combines variation from random slope auto-covariance K_{U_1}, cross-covariance $K_{U_{01}}$ and subject-visit specific covariance K_V. As a result, the total covariance will only be decomposed into three parts, the K_{U_0}, K_{V_U}, and K_W. Therefore, if the model is misspecified as a MFPCA model which ignores the slope process, it is expected to witness an inflation of variation at visit-level, while the rest of the variation at subject and day-level will not be changed. Given that the total amount of variation is a fixed number, the proportions explained by processes at subject and day levels will be underestimated.

On the other hand, if we take the mean of the observed (day-level) curves at each visit, that is, let $\overline{X}_{ij}=\frac{1}{n_{ij}}\sum_{k=1}^{n_{ij}}X_{ijk}$, we then have,

$$\begin{aligned} \overline{X}_{ij}(t) &= \mu(t)+U_{i0}(t)+U_{i1}(t)T_{ij}+V_{ij}(t)+\frac{1}{n_{ij}}\sum_{k=1}^{n_{ij}}W_{ijk}(t) \\ Var(\overline{X}_{ij}) &= K_{U_0}+T_{ij}K_{U_{01}}+T_{ij}K_{U_{01}}+T_{ij}^2K_{U_1}+K_V \\ &\quad +\frac{1}{n_{ij}^2}\sum_{k,k'}Cov(W_{ijk},W_{ijk'}) \end{aligned} \tag{14}$$

The total variation is increased with an additional term $\frac{1}{n_{ij}^2}\sum_{k,k'}Cov(W_{ijk},W_{ijk'})$, compared with Equation (11). As a result, the estimated explained variance within both K_U and K_V will be increased by an approximately similar amount, and correspondingly, the proportions of the respective explained variations increase. However, the relative variation, that is, the ratios between eigenvalues of $\widehat{K}_U$ and $\widehat{K}_V$, remain essentially fixed. In fact, taking the average can be considered as a form of smoothing, especially when the processes $\widehat{W}_{ijk}$ are random errors or only explain a small proportion of the total variation. Therefore, a two-level model may be applicable if the day-to-day variation in the data is ignorable or the major research questions are primarily focused on subject and visit levels.

The above insights for MFPCA and LFPCA models are further explored and demonstrated with our simulation studies in Section 7.

6 | Regression Model

An important next step is to evaluate associations between the multilevel predictors and longitudinal outcomes, for example, PA curves and metabolic health as in the MENU study. Traditional multivariate linear models can be extended to scalar-on-function regression models to explore the associations between scalar outcomes and functional predictors, namely,

$$E(Y_i) = \alpha_0 + \int_{\mathscr{D}} U_i(t)\beta_U(t)dt \tag{15}$$

where $Y_i \in \mathscr{R}$ is a scalar outcome, α_0 is the regression intercept, $U_i(t)$ is a functional predictor, $\beta_U(t)$ is the corresponding functional regression coefficient. With normally-distributed outcomes, a functional principal components regression (FPCR) model [29] reconstructs $U_i(t)$ using the estimated scores and eigenfunctions from the FPCA model (Section 3), $U_i(t) = \sum_{l=1}^{N_U} \xi_{il}\phi_l^U(t)$. Gertheiss et al. [24] extended the method to a longitudinal FPCR setup, where the outcome Y_{ij} was recorded for each subject i at visit j and includes both subject-level functional predictors $U_i(t)$ and visit-level functional predictors $V_{ij}(t)$. The longitudinal model has the form,

$$E(Y_{ij}) = \alpha_0 + b_i + \int_{\mathscr{D}} \beta_U(t)U_i(t, T_{ij})dt + \int_{\mathscr{D}} \beta_V(t)V_{ij}(t)dt \tag{16}$$

where b_i is a subject-specific random effect. We assume $b_i \sim N(0, \tau^2)$ with Y_{ij} conditionally independent given b_i. Here, $U_i(t, T_{ij}) = U_{i0}(t) + U_{i1}(t)T_{ij}$; $U_i(t, T_{ij})$ represents the between-subject variation and $V_{ij}(t)$ represents the within-subject variation over the domain function $\mathscr{D}$. The $\beta_U(t)$ and $\beta_V(t)$ are smooth coefficient functions for processes $U_i(t, T_{ij})$ and $V_{ij}(t)$, respectively. Longitudinal FPCR regression uses the eigenfunction decomposition of the functional predictors so that $U_{i0}(t) = \sum_l^{N_U} \xi_{il}\phi_l^{U_0}(t)$, $U_{i1}(t) = \sum_l^{N_U} \xi_{il}\phi_l^{U_1}(t)$ and $V_{ij}(t) = \sum_m^{N_V} \zeta_{ijm}\phi_m^v(t)$ as described in Section 5.2.

Both standard FPCR and longitudinal FPCR yield smooth coefficient functions, which have a useful interpretation over time t. These coefficient functions in Equations (15) and (16) can be estimated using penalized spline methods via the R package `mgcv` [30, 31]. In particular, since the coefficient functions are expressed via spline bases, they do not explicitly depend on the number of principal components selected [24].

7 | Simulation Study

In this section, simulation studies were implemented to explore the properties of the methods provided in Section 4. In addition to testing the robustness of our proposed methods, another goal is to explore how these methods perform when the model is misspecified under varying simulation settings. Specifically, we performed simulation studies in both

unbalanced cross-sectional (one-level outcome and two-level predictor) and longitudinal (two-level outcome and three-level predictor) setups. Our motivation for considering these two setups is to explicitly evaluate the impact of ignoring the multilevel structure of the data versus ignoring the longitudinal structure. For instance, taking the average at visit level of a two-level predictor could result in simultaneous loss of information in both multilevel and longitudinal structure. However, for a three-level predictor, the averaging process performed on day-level measures still retains the longitudinal structure of the input data, but is expected to lose the three-level structure. We performed a series of simulations studies to validate these assumptions in cross-sectional and longitudinal setups.

For simulation studies, we compared the performance in both functional models and regression models. The normalized errors (relative bias) between the estimated and true eigenvalues and principal component scores were used as the evaluation criteria for functional modeling. As for regression results, we computed the observed mean squared errors (MSE). For each simulation setting, we generated $M = 100$ replicates with $N = 100$ subjects. For illustration purpose, we only present simulation design and results of three-level functional inputs; details of two-level functional data can be found in Supplementary Material. The corresponding R code for our proposed method and other models used in simulation studies is available at https://github.com/wendylin23/MixedFPCA.

The simulation studies were based on three-level settings and we assumed a fixed numbers of visits $n_i = 3$ and days $n_{ij} = 3$ for each subject. (The time variable $T_{ij}$is generated by standardizing the visits, that is, $T_{ij} = \frac{j - \frac{1}{n_i}\sum j}{sd(j)}$, to have unit variance.) The functional curves $X_{ijk}(t)$ with length of $D = 600$ were generated according to the mLFPCA (Model (5)) and the true model was set as,

$$\begin{cases} y_{ij} = b_i + \int \beta_U(t) U_i(t) dt + \int \beta_V(t) V_{ij}(t) dt + \epsilon_{ij}, \\ X_{ijk}(t) = \sum_{l=1}^{N_U} \xi_{il} \phi_l^{(U_0)}(t) + \sum_{l=1}^{N_U} T_{ij} \xi_{il} \phi_l^{(U_1)}(t) + \sum_{m=1}^{N_V} \zeta_{ijm} \phi_m^{(V)}(t) \\ \qquad + \sum_{r=1}^{N_W} \eta_{ijkr} \phi_r^W(t), t \in \mathcal{D} \\ \xi_{il} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \lambda_l^U), \quad \zeta_{ijm} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \lambda_m^V), \quad \eta_{ijkr} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \lambda_r^W) \end{cases} \tag{17}$$

where the number of eigenfunctions is set as $N_U = N_V = N_W = 4$. The eigenfunctions bases can be orthonormal sine/cosine basis (F-basis) and Legendre polynomials basis (L-basis). It is noted that the sine/cosine basis is orthogonal with each other but it is correlated with the Legendre polynomials basis. In addition, we also considered the cases where white noise (i.e., random error) curves replaced the subject-level slope or visit-level curves, to mimic the situations when the longitudinal or between-visit variability is small. These random error curves (E-basis) were generated via $e_{ij}(t) \sim N(0, \sigma_e^2)$, $\sigma_e = 0.3$, which was used to replace the random slope or visit-specific term in Equation (17). In the following sections, we used

abbreviations to represent the combination of different basis. For instance, “FFFF” refers to the combination of all four orthogonal Fourier basis in U_0, U_1, V, and W processes. In this study, we simulated data based on five types of basis combinations, including

a. FFFF: $\phi_l^{(U_0)}(t)$, $\phi_l^{(U_1)}(t)$, $\phi_m^{(V)}(t)$, $\phi_r^W(t) \sim A\sin(\omega t + \psi)$, orthogonal Fourier basis.

b. FFFL: $\phi_l^{(U_0)}(t)$, $\phi_l^{(U_1)}(t)$, $\phi_m^{(V)}(t) \sim A\sin(\omega t + \psi)$ orthogonal Fourier basis, $\phi_r^W(t) \sim Bt^r$, $r = 0, 1, \ldots$ orthogonal polynomial basis.

c. FLFF: $\phi_l^{(U_0)}(t)$, $\phi_m^{(V)}(t)$, $\phi_r^W(t) \sim A\sin(\omega t + \psi)$, orthogonal Fourier basis, $\phi_l^{(U_1)}(t) \sim Bt^l$, $l = 0, 1 \ldots$ orthogonal polynomial basis.

d. FFFE: $\phi_l^{(U_0)}(t)$, $\phi_l^{(U_1)}(t)$, $\phi_m^{(V)}(t) \sim A\sin(\omega t + \psi)$, orthogonal Fourier basis, $\phi_r^W(t) \sim N(0, \sigma_e^2)$ random errors.

e. FEFF: $\phi_l^{(U_0)}(t)$, $\phi_m^{(V)}(t)$, $\phi_r^W(t) \sim A\sin(\omega t + \psi)$, orthogonal Fourier basis, $\phi_l^{(U_1)}(t) \sim N(0, \sigma_e^2)$ random errors.

where A, $B \in \mathbb{R}$ and $\omega = k_1\pi$, $\psi = k_2\pi$, k_1, $k_2 \in \mathbb{N}$. Corresponding eigenvalues were to be $\lambda_l^U = \lambda_l^V = \lambda_l^W = 0.5^{l-1}$, $l = 1, 2, 3, 4$.

The random error ϵ_{ij} in the regression models were assumed to be normal with variance $\sigma^2 = 2$. b_i is a random intercept process and follows $b_i \sim N(0, 1)$. Following the regression simulation settings in Gertheiss et al. [24], we also used a Gamma density to simulate the true coefficient functions $\beta_U(t)$ and $\beta_V(t)$. For each of the 100 simulated datasets, we implemented the multilevel longitudinal FPCA (Model (5), mLFPCA) and multilevel FPCA (Model (4), MFPCA) on the three-level simulated functional data $X_{ijk}(t)$, and the two-level longitudinal FPCA model (Model (3), LFPCA-lv2) on the day-averaged functional inputs $\overline{X}_{ij\cdot}(t) = \frac{1}{n_{ij}}\sum_j X_{ijk}(t)$. The eigenfunctions, eigenvalues, scores and predicted functional trajectories $\widehat{U}_i(t)$ and $\widehat{V}_{ij}(t)$ were estimated from each model. Furthermore, we considered cases where amounts of variation explained in the visit- and day-level varied, that is, we assumed that the true eigenvalues can vary among levels. Additional tests, such as unbalanced design with missing visits and unequal eigenvalues, were also implemented. To simplify exposition, we focus here on the results using equal eigenvalues at each level, assuming no missing data. Additional results can be found in Supplementary Material.

In Table 1, we show the results of normalized errors (relative bias) between the estimated and true eigenvalues for subject-level process $(\widehat{\lambda}_l^U - \lambda_l^U) / \lambda_l^U$ and visit-level process $(\widehat{\lambda}_m^V - \lambda_m^V) / \lambda_m^V$, respectively, based on 100 replicates from the five simulation scenarios (a–e). At subject level, all three methods provide similarly unbiased estimates of the eigenvalues. But at visit level, the proposed mLFPCA model provides the least biased estimates of the eigenvalues in all five scenarios. Meanwhile, the two-level longitudinal model (LFPCA-lv2) performs consistently better than the three-level FPCA (MFPCA) model, except for the last scenario, where only random error terms are added to the slope-level. This finding conforms with our theoretical intuition in Section 5.3, where we showed that when the model fitting ignores the random slope process and is misspecified as a MFPCA model,

estimated variation at visit level is inflated and the consequences are reflected by these overestimated visit-level eigenvalues. On the other hand, when the LFPCA-lv2 is fitted to the day-averaged data, because the relative explained variation at the subject and visit levels are maintained, the estimation biases are hence much lower. To further elucidate these results, we provide (Table 2) the proportions of explained variance by different levels of the first two principal components, comparing results from the three fitted models with the true setting of the first simulation scenario ((a) FFFF). As expected, these Table 2 further validate our derivation in Equation (13) and (14), showing the impact of misspecification at different levels.

The last column in Table 1 presents MSE results $\frac{1}{N}\sum_i \frac{1}{n_i}\sum_j (y_{ij} - \hat{y}_{ij})^2$ from the regression fitting and the mLFPCA models generally have the best prediction performance. Compared with MFPCA models, the two-level longitudinal FPCA models still perform better in the first four scenarios. Combined with the similar findings in eigenvalues, we conclude that in data applications, misspecifying a three-level longitudinal model with the form of a two-level longitudinal structure (LFPCA-lv2) may be more optimal than misspecifying it as three-level FPCA models (MFPCA). However, it is important to note that all simulated data in our study are well-aligned with a same starting point and time range, which may also increase the relative robustness of the averaging procedure.

As an interesting parenthetical remark, the last two scenarios (d: FFFE and e: FEFF) can be considered as special cases of model misspecification, and can also provide insights on when the simpler MFPCA or LFPCA-lv2 models will perform as well as our proposed mLFPCA model. Specifically, in setups where the proportion of explained variation is small for the random slope processes (Scenario e) or day-specific processes (Scenario d), the mLFPCA model essentially reduces to a three-level MFPCA model (i.e., ignoring the slope process) or a two-level longitudinal (LFPCA-lv2) model (i.e., averaging over days).

Finally, we note that the PC scores are unbiasedly estimated by all the methods, as expected (Section S2A). We also provide additional simulation results exploring the impacts of missing data in Section S2B.

8 | Data Application

We provide a brief background on the rationale and protocol of the MENU Study, before delving into the data application. The prevalence of obesity in the US has been steadily increasing over the last 20 years with recent age-adjusted estimates indicating that 42.4% of US adults are obese [32, 33]. Obesity can be associated with serious health risks [34]. For instance, compared with persons with normal weight, overweight or obese persons are more vulnerable to glucoregulatory dysfunction and dyslipidemia, major risk factors for cardiovascular disease and other comorbidities [35-37]. In addition, overweight status and obesity increase the risk of end-stage renal disease and many types of cancer [38, 39]. Since weight gain occurs when energy expenditure (EE) remains low while dietary consumption levels are high, certain amounts of PA for increasing EE are commonly considered as part of treatment plans for achieving weight-loss in obese individuals [40]. Based on these principles, the MENU trial, conducted under the auspices of the NIH-funded

Trans-disciplinary Research on Energetics and Cancer (TREC) Center at UCSD from 2011 to 2017, recruited $n = 245$ overweight women to a 12-month behavioral (diet and PA) weight-loss intervention [5, 6].

Each MENU participant was randomly assigned to one of three diet groups: a lower fat (20% of energy) and higher carbohydrate (65% of energy) diet; a lower carbohydrate (45% energy) and higher monounsaturated fat (35% energy) diet; or a walnut-rich (35% fat) and lower carbohydrate (45%) diet. All participants received the same PA intervention, with a prescribed goal of engaging in at least 60 min/day of purposeful exercise at a moderate level of intensity. Hence for the current investigation, which is focused on PA change, we treat the MENU Study as a longitudinal cohort study. There were three study-related clinic visits at baseline, 6 and 12 months [5, 6]. PA was measured using a triaxial accelerometer device, the GT3X+ Actigraph (Acti-Graph LLC, Pensacola FL). Participants were instructed to wear the devices for 7 days during waking hours and measurements of health outcomes were collected at each visit. Actigraph data were collected at high-resolution of 30 Hz, then processed into per minute PA counts [7]. More specifically, triaxial activity count vector outputs (AC_x, AC_y, AC_z) from these devices were summarized as magnitudes $\sqrt{AC_x^2 + AC_y^2 + AC_z^2}$, which are commonly referred to as magnitude counts and are related to intensity of the activity [7].

The goal of the current work was to utilize the longitudinal accelerometer-based PA to implement three-level functional data methods, and evaluate associations with longitudinal health outcomes. For this purpose, to ensure consistent data availability across participants, we extracted daily PA counts on three random days, including weekdays and weekends, for each participant at each visit. Sensitivity analysis were performed to show that the selected three-day data were representative of whole-week measures, explaining similar amount of variation at day-level. Figure 1 presents an example of three-day activity records for one participant at each visit. Since the starting time and duration time of device wearing are not constant across participants and days, we realigned daily records, so that all participants had a “common” starting time of device wear denoted as “0” on the x-axis. We kept the first 600 min (x-axis) from the records with at least 10 h of device wear (per standard protocols). The figure of daily PA magnitude based on clock time can be found in Section S3. Each data point (y-axis) represents minute-wise PA activity magnitude.

For exploring the association between PA and overweight/obese status, we considered several related health outcomes, including BMI, insulin levels and homeostatic model assessment (HOMA). The BMI, computed as weight in kilograms divided by height in meters squared (kg/m^2), is commonly used to identify overweight/obese status if BMI >25.0 [33]. In addition, high-levels of insulin have been proven to be associated with lifestyle-dependent obesity risk factors [41]. The HOMA index is a measure of insulin resistance and is computed as fasting insulin (microU/L) × fasting glucose (nmol/L)/22.5. Therefore, lower values of each outcome indicate better metabolic health. Brief descriptive statistics on demographics and health outcomes of MENU participants are provided in Section S3.

We first fitted the proposed three-level longitudinal FPCA (mLFPCA) model on the daily PA counts data for all subjects at each visit. Figure 2 presents the first three estimated

principal components for the random intercept, random slope, visit-specific and day-specific processes by columns. The top row of the figure provides the first principal component at each level. It shows that the red curve, which represents adding (a multiple of) the principal component to the mean, is always higher than the mean (black) curve in each figure. Specifically, a high score on this component at the subject-level (Level 1) indicates that a participant is on average more physically active, and has a higher increase across visits, compared to a participant with a low score. Similarly, a high score at the visit-level (Level 2) indicates that the participant has higher activity on that visit compared to her PA at other visits. Interestingly, the first principal component at the day-level emphasizes higher (or lower) activity during the first 200 min. Overall, the first components of the subject-, visit-, and day-level processes explain around 77% of total variation. The other components display more oscillatory patterns. For instance, the second Level 1 intercept component (middle plot in first column) is negative (i.e., below the mean) for the first 200 min but becomes positive (i.e., above the mean) for the last 200 min, meaning subjects with positive scores in this component tend to be less physically active earlier but more active later in their day. Since the Level 2 and Level 3 eigenfunctions represent the random visit- and day-specific functional deviation respectively, each component can be interpreted correspondingly.

Figure 3 provides an example of the daily raw, smoothed and model-recovered PA curves at each visit of one participant. As is evident, the model-recovered curves mirror closely the (smoothed) observed PA data, which further illustrates the robustness and applicability of our proposed model. Finally, in order to explain 95% of the variance, we retained $N_U = 10$, $N_V = 5$ and $N_W = 13$ principal components at the three levels, respectively.

Next, we aimed to fit a regression model to evaluate associations between functional PA inputs and longitudinal health outcomes. For this, as described in Section 6, the estimated scores and principal components were used to reconstruct the subject-level process U_i and visit-level process V_{ij} for each subject i and visit j. We then fitted the longitudinal FPCR model on each health outcome, respectively, with the reconstructed curves $\widehat{U}_i$ and $\widehat{V}_{ij}$, as denoted in Equation (16). The model also included a random intercept and additional covariates including age, ethnicity, smoking status and $\mathbf{1}(\text{visit} > 1)$. Figure 4 gives the estimated coefficient functions for log(Insulin), BMI and HOMA levels, with 95% pointwise confidence intervals, which are Bayesian component-wise variable width intervals constructed from the estimated Bayesian posterior covariance matrix [42]. The coefficients at a given time-point (on the x-axis) are considered significant if the 95% confidence limits at that time do not cross the reference horizontal line at $y = 0$. The coefficient function for the subject-level (Level 1) process of BMI was negative for the first 400 min, suggesting that on average, a subject with higher levels of PA had lower BMI compared to a subject with lower PA levels, especially in the first 200 min, during which the 95% confidence interval excluded the null value of zero. The subject-level coefficient curves for log(Insulin) and HOMA were also negative during the first 400+ min, suggesting possible beneficial associations with PA, although we cannot rule out the null hypothesis for these markers since the 95% confidence bands included zero. Notably, the estimated coefficient curves at visit-level (Level 2) of these two markers were negative at earlier times (with the

95% intervals excluding the zero value), suggesting that increasing PA earlier in the day from one visit to the next was associated with lower insulin and HOMA. Thus, although inference for the regression patterns at subject and visit level varied slightly among the three health outcomes, we observed negative associations between PA and all outcomes, and most interestingly, we found that engaging in PA earlier in the day seemed beneficial for mitigating overweight/obese risk factors.

9 | Discussion

In this work, we proposed a mLFPCA approach and compared its performance with other functional principal component models that have been previously applied on multilevel data, by means of both simulation study and real data application. Specifically, the proposed model was designed to fit data from a longitudinal study that has three-level functional inputs. It includes a two-step estimation procedure and eigen-expansion based methods to capture and decompose the covariance structures of the observed functional (PA) curves. The association between the functional predictor (PA) and overweight/obesity related health outcomes was then examined via functional regression approaches. In addition, a wide range of simulation studies were performed to validate and compare model performances.

Our proposed model can be regarded as a natural extension of previous methodology on multilevel and longitudinal FPCA [20-22]. To demonstrate the necessity of such an extension, we provided both theoretical illustration in Section 5.3 and simulations in Section 7. On the one hand, compared with previous implementations that are common in health behavior research, which used averaging to reduce the number of nested levels (e.g., averaging over days at each visit), our method retained the three-level longitudinal design structure, and thus can fully extract the variation information included in all nested levels and provide solid inference. On the other hand, though our proposed method consistently yields better performance, under certain scenarios simpler models may be acceptable and perform adequately depending on the study aims, and when they accurately reflect the pertinent information contained in the data. For instance, if the day-to-day variation in a three-level dataset explains a relatively small proportion of total variability, a two-level longitudinal FPCA model may be applicable for inputs averaged over days. Importantly, our simulation studies illustrate different misspecification effects in cross-sectional and longitudinal setups. We found that in the setting with cross-sectional outcomes, even with longitudinal functional inputs (Section S2C), the simpler misspecified multilevel FPCA models (which ignore the slope term but retain the multilevel structure in the functional inputs) had superior performance in terms of estimation and prediction, compared to single-level FPCA which reduced the levels of the functional inputs by averaging. However, in the setting with longitudinal outcomes, averaging the functional inputs (over the third-level) had superior performance compared to multilevel FPCA (which ignored the longitudinal component; Table 1). Thus, depending on the structure of the outcome data, misspecifying the longitudinal functional component appears to strongly influence results. We believe that these results can guide researchers in how to choose simpler approaches, should they wish to do so. Of course, preserving the full data structure, that is, all levels of the functional data, using the proposed mLFPCA model performs best, and although this method appears to be

more complex, the simulation studies and application indicate that the computation is, in fact, fairly efficient.

We also implemented the mLFPCA model in a data application. Our analysis of the MENU study, revealed a negative association between temporal patterns of PA and overweight/obese related health outcomes, that is, PA accumulated earlier in the day was related to healthier status. However, we need to exercise caution in making causal claims about directions of associations. Further work on how timing of PA most impacts health could be useful when designing intervention trials and informing public health recommendations.

Future studies could further extend our approach to functional data in longitudinal studies with more than three levels, such as studies in which the variation between morning versus evening PA are of interest. In fact, the structural FPCA proposed by Shou et al. [21] provided a general estimation procedure for data with any number of levels, but they do not explicitly consider longitudinal designs. By combining their approach and ours, we expect that these methods could be extended to multilevel (with >3) levels) longitudinal data, and we aim to pursue a similar approach in future work. Also, in our data application, the functional PA predictors are assessed at three levels (subject-specific, visit-specific, and day-specific) whereas the outcomes are only assessed at two levels (subject-specific and visit-specific). Our methods are also applicable when predictors and outcomes are both assessed at the same levels, for example, 3 levels. We will implement this extension in future work as relevant clinical questions arise. In addition, with functional regression models where predictors are measured more frequently than outcomes, we could consider other summary metrics that can incorporate information from the higher levels in the predictors. For instance, Steele et al. [43] proposed multilevel structural equation models for longitudinal data, but the implementation involved with functional data needs further exploration, which we aim to address in future studies.

In summary, in this work, we developed an analytic approach for three-level longitudinal functional data, as are common in health behavior (e.g., diet, sleep, PA) studies. Implementing this new model posed several challenges. Firstly, the data exhibited an unbalanced structure, with more levels in the predictor than in the outcome. Despite this, we were able to construct a model that retained most of the information from the original dataset. Secondly, high-dimensional data with multiple levels can present computational and storage issues. Our proposed two-step algorithm efficiently handled these challenges. Lastly, although numerous studies and algorithms have been developed in the FPCA field, the issue of model misspecification has not been adequately addressed. To address this gap, we utilized analytic calculations and comprehensive simulation studies, and evaluated the performance of the proposed model under different scenarios. We compared our model to potentially simpler models, with albeit missspecified structure. We provided guidance regarding when the simpler models may be appropriate, which is crucial for real-world data applications. We applied our method to data from a longitudinal study on obesity measures and PA assessed via accelerometry and obtained meaningful results. Importantly, our approach can be applied to other applications with densely sampled data for example, continuous glucose monitoring, heart rate monitoring and so forth. We believe that this work could add to the body of computational methods for analyzing the full spectrum of data

from wearable sensors, which are becoming more and more common in public health and biomedical applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

CLR and LN were partially supported by funding from the National Cancer Institute (U54 CA155435-01); CD and LN were partially supported by funding from the National Institute on Aging (PO1AG052352) and from the National Heart, Lung and Blood Institute (R01HL130483); JZ and LN were partially supported by funding from the National Institute of Diabetes and Digestive and Kidney Disease (R01DK114945). Support for high performance computing was provided by National Science Foundation grants 2100237 and 2120019. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or National Science Foundation.

Data Availability Statement

Software in the form of R code is available on Github (https://github.com/wendylin23/MixedFPCA). Data that support the findings in this paper are available upon request.

References

1. Troiano RP, McClain JJ, Brychta RJ, and Chen KY, “Evolution of Accelerometer Methods for Physical Activity Research,” British Journal of Sports Medicine 48, no. 13 (2014): 1019–1023. [PubMed: 24782483]
2. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, and Tremblay MS, “Physical Activity of Canadian Adults: Accelerometer Results From the 2007 to 2009 Canadian Health Measures Survey,” Health Reports 22, no. 1 (2011): 7.
3. Luke A, Dugas LR, Durazo-Arvizu RA, Cao G, and Cooper RS, “Assessing Physical Activity and Its Relationship to Cardiovascular Risk Factors: NHANES 2003–2006,” BMC Public Health 11, no. 1 (2011): 1–11. [PubMed: 21199570]
4. Trost SG, Pate RR, Freedson PS, Sallis JF, and Taylor WC, “Using Objective Physical Activity Measures With Youth: How Many Days of Monitoring Are Needed?,” Medicine & Science in Sports & Exercise 32, no. 2 (2000): 426. [PubMed: 10694127]
5. Le T, Flatt SW, Natarajan L, et al. , “Effects of Diet Composition and Insulin Resistance Status on Plasma Lipid Levels in a Weight Loss Intervention in Women,” Journal of the American Heart Association 5, no. 1 (2016): e002771. [PubMed: 26811166]
6. Rock CL, Flatt SW, Pakiz B, et al. , “Effects of Diet Composition on Weight Loss, Metabolic Factors and Biomarkers in a 1-Year Weight Loss Intervention in Obese Women Examined by Baseline Insulin Resistance Status,” Metabolism 65, no. 11 (2016): 1605–1613. [PubMed: 27733248]
7. Bassett DR, “Device-Based Monitoring in Physical Activity and Public Health Research,” Physiological Measurement 33, no. 11 (2012): 1769.
8. Freedson P, Pober D, Janz KF, et al. , “Calibration of Accelerometer Output for Children,” Medicine and Science in Sports and Exercise 37, no. 11 (2005): S523. [PubMed: 16294115]
9. Troiano RP, Berrigan D, Dodd KW, et al. , “Physical Activity in the United States Measured by Accelerometer,” Medicine and Science in Sports and Exercise 40, no. 1 (2008): 181. [PubMed: 18091006]
10. Tudor-Locke C, Brashear MM, Johnson WD, and Katzmarzyk PT, “Accelerometer Profiles of Physical Activity and Inactivity in Normal Weight, Overweight, and Obese US Men and Women,” International Journal of Behavioral Nutrition and Physical Activity 7, no. 1 (2010): 1–11. [PubMed: 20145731]

11. Matthews CE, Chen KY, Freedson PS, et al. , "Amount of Time Spent in Sedentary Behaviors in the United States, 2003–2004," American Journal of Epidemiology 167, no. 7 (2008): 875–881. [PubMed: 18303006]
12. Neuenschwander M, Barbaresko J, Pischke CR, et al. , "Intake of Dietary Fats and Fatty Acids and the Incidence of Type 2 Diabetes: A Systematic Review and Dose-Response Meta-Analysis of Prospective Observational Studies," PLoS Medicine 17, no. 12 (2020): e1003347. [PubMed: 33264277]
13. Pettee Gabriel K, Sidney S, Jacobs DR Jr., et al. , "Ten-Year Changes in Accelerometer-Based Physical Activity and Sedentary Time during Midlife: The CARDIA Study," American Journal of Epidemiology 187, no. 10 (2018): 2145–2150. [PubMed: 29893772]
14. Bürgi F, Meyer U, Granacher U, et al. , "Relationship of Physical Activity With Motor Skills, Aerobic Fitness and Body Fat in Preschool Children: A Cross-Sectional and Longitudinal Study (Ballabeina)," International Journal of Obesity 35, no. 7 (2011): 937–944. [PubMed: 21448128]
15. Ramsay J and Silverman B, Functional Data Analysis (New York, NY: Springer, 2005).
16. Wang JL, Chiou JM, and Müller HG, "Functional Data Analysis," Annual Review of Statistics and Its Application 3 (2016): 257–295.
17. Shang HL, "A Survey of Functional Principal Component Analysis," AStA Advances in Statistical Analysis 98 (2014): 121–142.
18. Yao F and Lee TC, "Penalized Spline Models for Functional Principal Component Analysis," Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68, no. 1 (2006): 3–25.
19. Zhang X and Wang JL, "From Sparse to Dense Functional Data and Beyond," Annals of Statistics 44, no. 5 (2016): 2281–2321.
20. Di CZ, Crainiceanu CM, Caffo BS, and Punjabi NM, "Multilevel Functional Principal Component Analysis," Annals of Applied Statistics 3, no. 1 (2009): 458. [PubMed: 20221415]
21. Shou H, Zipunnikov V, Crainiceanu CM, and Greven S, "Structured Functional Principal Component Analysis," Biometrics 71, no. 1 (2015): 247–257. [PubMed: 25327216]
22. Greven S, Crainiceanu C, Caffo B, and Reich D, "Longitudinal Functional Principal Component Analysis," Electronic Journal of Statistics 4 (2010): 1022–1054, 10.1214/10-EJS575. [PubMed: 21743825]
23. Goldsmith J, Crainiceanu CM, Caffo B, and Reich D, "Longitudinal Penalized Functional Regression for Cognitive Outcomes on Neuronal Tract Measurements," Journal of the Royal Statistical Society: Series C (Applied Statistics) 61, no. 3 (2012): 453–469. [PubMed: 22679339]
24. Gertheiss J, Goldsmith J, Crainiceanu C, and Greven S, "Longitudinal Scalar-on-Functions Regression with Application to Tractography Data," Biostatistics 14, no. 3 (2013): 447–461. [PubMed: 23292804]
25. Xu SY, Nelson S, Kerr J, et al. , "Modeling Temporal Variation in Physical Activity Using Functional Principal Components Analysis," Statistics in Biosciences 11, no. 2 (2019): 403–421.
26. Lin W, Zou J, Di C, Sears DD, Rock CL, and Natarajan L, "Longitudinal Associations between Timing of Physical Activity Accumulation and Health: Application of Functional Data Methods," Statistics in Biosciences 15, no. 2 (2023): 309–329. [PubMed: 37383028]
27. Xvi MJ, "Functions of Positive and Negative Type, and Their Connection the Theory of Integral Equations," Philosophical Transactions of the Royal Society of London: Series A (Containing Papers of a Mathematical or Physical Character) 209, no. 441–458 (1909): 415–446.
28. Koch GG, "Some Further Remarks Concerning "A General Approach to the Estimation of Variance Components"," Technometrics 10, no. 3 (1968): 551–558.
29. Reiss PT and Ogden RT, "Functional Principal Component Regression and Functional Partial Least Squares," Journal of the American Statistical Association 102, no. 479 (2007): 984–996.
30. Wood SN, Generalized Additive Models: An Introduction With R (Boca Raton, FL: Chapman and Hall/CRC Press, 2006).
31. Wood SN, "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models," Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73, no. 1 (2011): 3–36.
32. Hales CM, Fryar CD, Carroll MD, Freedman DS, Aoki Y, and Ogden CL, "Differences in Obesity Prevalence by Demographic Characteristics and Urbanization Level Among Adults in the United

States, 2013–2016," Journal of the American Medical Association 319, no. 23 (2018): 2419–2429. [PubMed: 29922829]

33. Fryar CD, Carroll MD, and Ogden CL, "Prevalence of Overweight, Obesity, and Severe Obesity Among Children and Adolescents Aged 2–19 Years: United States, 1963–1965 Through 2015–2016" (2018).
34. Members EP, Jensen MD, Ryan DH, et al. , "Executive Summary: Guidelines (2013) for the Management of Overweight and Obesity in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Obesity Society Published by the Obesity Society and American College of Cardiology/American Heart Association Task Force on Practice Guidelines," Based on a Systematic Review From the the Obesity Expert Panel 22, no. S2 (2013): S5–S39.
35. Nguyen NT, Magno CP, Lane KT, Hinojosa MW, and Lane JS, "Association of Hypertension, Diabetes, Dyslipidemia, and Metabolic Syndrome With Obesity: Findings from the National Health and Nutrition Examination Survey, 1999 to 2004," Journal of the American College of Surgeons 207, no. 6 (2008): 928–934. [PubMed: 19183541]
36. Eckel RH, Jakicic JM, Ard JD, et al. , "2013 AHA/ACC Guideline on Lifestyle Management to Reduce Cardiovascular Risk: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines," Journal of the American College of Cardiology 63, no. 25 Part B (2014): 2960–2984. [PubMed: 24239922]
37. Fox CS, Golden SH, Anderson C, et al. , "Update on Prevention of Cardiovascular Disease in Adults with Type 2 Diabetes Mellitus in Light of Recent Evidence: A Scientific Statement from the American Heart Association and the American Diabetes Association," Circulation 132, no. 8 (2015): 691–718. [PubMed: 26246173]
38. Cy H, McCulloch CE, Iribarren C, Darbinian J, and Go AS, "Body Mass Index and Risk for End-Stage Renal Disease," Annals of Internal Medicine 144, no. 1 (2006): 21–28. [PubMed: 16389251]
39. Kushi LH, Doyle C, McCullough M, et al. , "American Cancer Society Guidelines on Nutrition and Physical Activity for Cancer Prevention: Reducing the Risk of Cancer with Healthy Food Choices and Physical Activity," CA: A Cancer Journal for Clinicians 62, no. 1 (2012): 30–67. [PubMed: 22237782]
40. Haskell WL, Lee IM, Pate RR, et al. , "Physical Activity and Public Health: Updated Recommendation for Adults from the American College of Sports Medicine and the American Heart Association," Circulation 116, no. 9 (2007): 1081. [PubMed: 17671237]
41. Kolb H, Stumvoll M, Kramer W, Kempf K, and Martin S, "Insulin Translates Unfavourable Lifestyle Into Obesity," BMC Medicine 16, no. 1 (2018): 1–10.
42. Marra G and Wood SN, "Coverage Properties of Confidence Intervals for Generalized Additive Model Components," Scandinavian Journal of Statistics 39, no. 1 (2012): 53–74.
43. Steele F, Clarke P, Leckie G, Allan J, and Johnston D, "Multilevel Structural Equation Models for Longitudinal Data Where Predictors are Measured More Frequently Than Outcomes: An Application to the Effects of Stress on the Cognitive Function of Nurses," Journal of the Royal Statistical Society: Series A (Statistics in Society) 180, no. 1 (2017): 263–283.

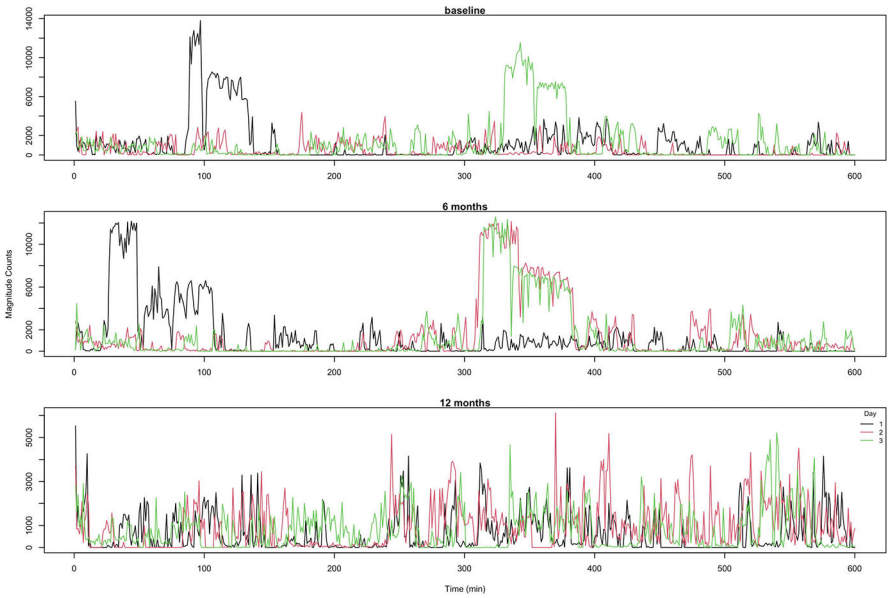

FIGURE 1 |.

An example of the first 600 min (*x*-axis) of daily physical activity (PA) magnitude counts on 3 days (denoted by varying colors) from minute-level accelerometer count data for one subject across three visits (baseline, 6, and 12 months). The daily records were realigned to have a "common" starting time of device wear denoted as "0" on the *x*-axis. Values on *y*-axis provide the vector magnitude of PA measured from accelerometers.

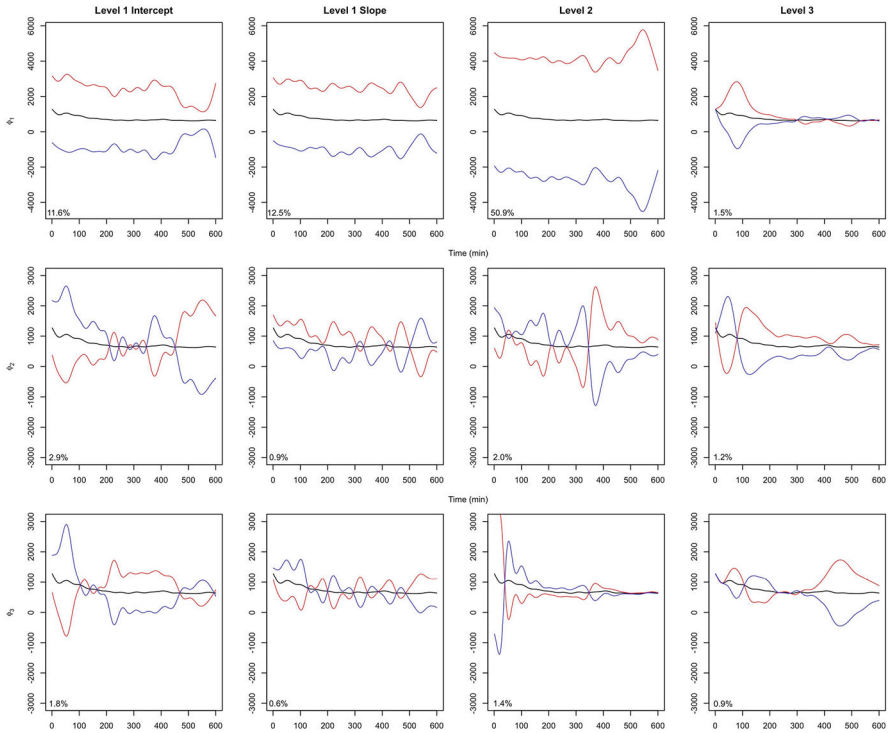

FIGURE 2 |.

MENU study application: The first three estimated principal components (rows) for the random intercept (1st column), random slope (2nd column), visit-specific process (3rd column) and day-specific process (4th column). The *y*-axis of the plots give the overall mean value curve $\mu(t)$ (black) with addition (red) or subtraction (blue) of 2 square root of eigenvalues multiplying first, second or third level principal component curves. The %s in the bottom left of each graph are the percent of variation explained by that component. The *x*-axis is time (min) representing 10 h (600 min) of activity.

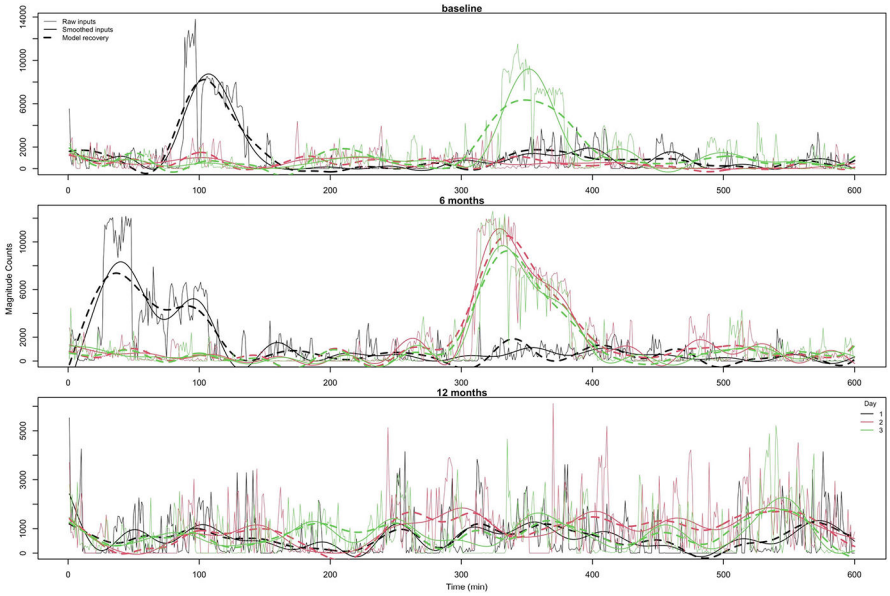

FIGURE 3 |.

An example of PA magnitude counts (*y*-axis) with raw count inputs (thin solid), smoothed curves (thick solid) and model-recovered curves (thick dashed) at baseline (top), 6 months (middle), and 12 months (bottom), during 600 min (*x*-axis) of each day, starting from when the participants began wearing the device. The daily records were realigned to have a "common" starting time of device wear denoted as "0" on the *x*-axis. Different colors of the line represent the day of the measurement.

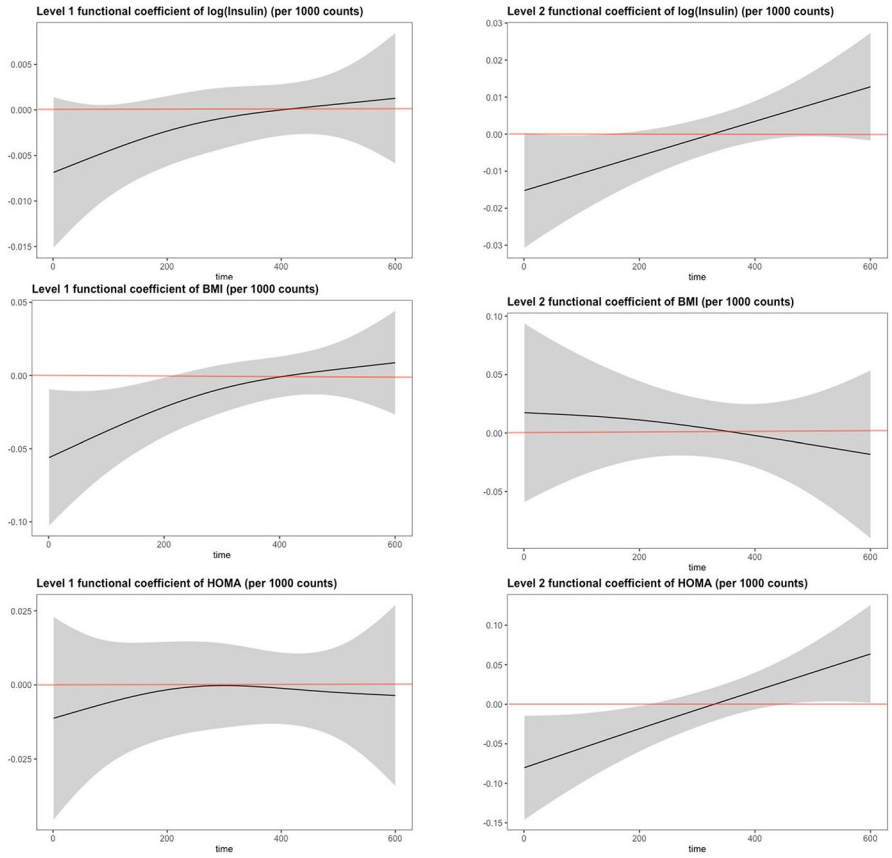

FIGURE 4 |.

MENU study application: Estimated functional coefficients curve with 95% pointwise confidence intervals (shaded grey) when implementing the longitudinal FPCR model on log(Insulin) (top), BMI (middle) and HOMA (bottom), with U (Level 1) and V (Level 2) processes reconstructed from the fitted multilevel longitudinal FPCA (mLFPCA) model as functional predictors, after adjusting for age, ethnicity, smoking status, and visit >1; x-axis is time in minutes.

TABLE 1 |

Simulation results for five simulated scenarios: Relative bias or normalized errors across $M = 100$ simulation samples (Mean (SD))[a] for subject-level eigenvalues $(\widehat{\lambda}_l^U - \lambda_l^U) / \lambda_l^U$, and visit-level eigenvalues $(\widehat{\lambda}_m^V - \lambda_m^V) / \lambda_m^V$ for the four principal components (PC1–PC4), and mean-squared error from regression[b] $(\frac{1}{N}\sum_i \frac{1}{n_i}\sum_j (y_{ij} - \hat{y}_{ij})^2)$, fitted via multilevel longitudinal FPCA (mLFPCA), multilevel FPCA (MFPCA) and two-level longitudinal FPCA (LFPCA-lv2). Five simulation scenarios were based on the combination of different basis (F: Fourier basis, L: Legendre polynomials basis, E: random error curve).

	Subject-level eigenvalue $(\widehat{\boldsymbol{\lambda}}_l^U - \boldsymbol{\lambda}_l^U) / \boldsymbol{\lambda}_l^U$				Visit-level eigenvalue $(\widehat{\boldsymbol{\lambda}}_m^V - \boldsymbol{\lambda}_m^V) / \boldsymbol{\lambda}_m^V$				Regression fitted value
	PC1	**PC2**	**PC3**	**PC4**	**PC1**	**PC2**	**PC3**	**PC4**	**Mean-squared error**
Scenario (a): FFFF									
mLFPCA	0.02 (0.64)	0.02 (0.31)	0.03 (0.16)	0.07 (0.12)	0.19 (0.69)	−0.01 (0.35)	0.004 (0.14)	0.01 (0.07)	0.0002 (0.0003)
MFPCA	0.01 (0.60)	0.04 (0.33)	0.04 (0.17)	0.08 (0.14)	1.03 (0.85)	1.86 (1.44)	0.59 (0.46)	0.79 (0.59)	0.03 (0.02)
LFPCA-lv2	0.00 (0.65)	0.00 (0.31)	0.01 (0.17)	0.03 (0.10)	0.17 (0.68)	−0.04 (0.35)	−0.04 (0.14)	−0.01 (0.07)	0.0041 (0.0007)
Scenario (b): FFFL									
mLFPCA	−0.05 (0.67)	−0.01 (0.31)	0.00 (0.17)	0.06 (0.12)	0.20 (0.73)	−0.08 (0.35)	−0.07 (0.15)	−0.03 (0.07)	0.0021 (0.0003)
MFPCA	−0.03 (0.70)	0.00 (0.35)	0.02 (0.17)	0.04 (0.14)	0.95 (0.82)	1.84 (1.44)	0.54 (0.43)	0.72 (0.55)	0.03 (0.02)
LFPCA-lv2	0.00 (0.67)	0.01 (0.31)	0.02 (0.18)	0.05 (0.11)	0.38 (0.79)	0.20 (0.41)	0.26 (0.24)	0.36 (0.28)	0.02 (0.004)
Scenario (c): FLFF									
mLFPCA	0.02 (0.62)	0.07 (0.38)	−0.03 (0.15)	0.06 (0.12)	0.19 (0.69)	−0.01 (0.35)	0.00 (0.14)	0.01 (0.07)	0.0028 (0.0012)
MFPCA	−0.39 (0.73)	−0.48 (0.46)	−0.15 (0.19)	0.00 (0.10)	0.60 (0.77)	0.79 (0.58)	0.77 (0.56)	0.64 (0.48)	0.09 (0.03)
LFPCA-lv2	0.01 (0.62)	0.06 (0.38)	−0.05 (0.16)	0.04 (0.10)	0.17 (0.68)	−0.04 (0.35)	−0.04 (0.14)	−0.01 (0.07)	0.0045 (0.012)
Scenario (d): FFFE									
mLFPCA	0.03 (0.65)	0.02 (0.32)	0.03 (0.17)	0.07 (0.13)	0.19 (0.72)	−0.02 (0.35)	0.00 (0.14)	0.00 (0.10)	0.02 (0.01)

Stat Med. Author manuscript; available in PMC 2025 May 28.

	Subject-level eigenvalue $(\widehat{\lambda}_l^U - \lambda_l^U) / \lambda_l^U$				Visit-level eigenvalue $(\widehat{\lambda}_m^V - \lambda_m^V) / \lambda_m^V$				Regression fitted value
	PC1	**PC2**	**PC3**	**PC4**	**PC1**	**PC2**	**PC3**	**PC4**	**Mean-squared error**
MFPCA	0.02 (0.61)	0.04 (0.36)	0.03 (0.18)	0.04 (0.14)	1.05 (0.87)	1.86 (1.46)	0.61 (0.48)	0.79 (0.59)	0.07 (0.05)
LFPCA-lv2	0.03 (0.65)	0.01 (0.32)	0.02 (0.17)	0.04 (0.12)	0.30 (0.75)	0.08 (0.36)	0.01 (0.14)	0.01 (0.14)	0.01 (0.001)
Scenario (e): FEFF									
mLFPCA	−0.02 (0.59)	0.02 (0.32)	0.02 (0.15)	0.08 (0.11)	0.19 (0.69)	−0.01 (0.35)	0.00 (0.14)	0.01 (0.07)	0.0002 (0.0003)
MFPCA	−0.02 (0.60)	0.01 (0.33)	0.00 (0.15)	0.0 (0.09)	0.10 (0.51)	0.00 (0.23)	−0.02 (0.10)	−0.01 (0.05)	0.0002 (0.0003)
LFPCA-lv2	−0.04 (0.59)	0.01 (0.32)	0.01 (0.16)	0.05 (0.10)	0.17 (0.68)	−0.04 (0.35)	−0.04 (0.14)	−0.01 (0.07)	0.0038 (0.0007)

[a] The Mean (SD) was computed errors from $M = 100$ simulations.

[b] The mean-squared error from regression was based on $N = 100$ subjects and $n_i = 3$ visits.

TABLE 2 |

Proportions of variance explained in simulation scenario "FFFF" by different levels of the first two principal components ($\phi_l^{U_0}$, $\phi_l^{U_1}$, ϕ_m^V, ϕ_r^W) in (a) simulated model, (b) multilevel longitudinal FPCA (mLFPCA), (c) multilevel FPCA (MFPCA), (d) two-level longitudinal FPCA (LFPCA-lv2), averaged across $M = 100$ simulation samples. The proportions in (a) simulated model were computed based on the simulated data, representing the ground truth.

	# Component	$\boldsymbol{\phi}_l^{U_0}$	$\boldsymbol{\phi}_l^{U_1}$	$\boldsymbol{\phi}_m^V$	$\boldsymbol{\phi}_r^W$
(a) Simulated model	1	0.09	0.09	0.18	0.18
	2	0.04	0.04	0.09	0.09
(b) mLFPCA	1	0.09	0.09	0.18	0.17
	2	0.04	0.04	0.09	0.09
(c) MFPCA	1	0.09	0	0.22	0.17
	2	0.04	0	0.17	0.08
(d) LFPCA-lv2	1	0.13	0.14	0.27	0
	2	0.07	0.07	0.13	0