

An approach for estimating activity distribution surfaces

Frederike Vogel* Álvaro Méndez Civieta[†] Ying Wei[†] Keith M Diaz[‡]
Jeff Goldsmith[†]

Abstract

A collection of functional quantiles offers a comprehensive portrayal of an individual’s distinct physical activity patterns, characterizing the distributions of activity at specific time points over a defined duration. However, the potential to present such a collection as a unified, complete activity surface for each individual remains unexplored in current literature. In this paper, we rectify this shortcoming by proposing functional quantile surface estimation (FQSE), a methodology designed to calculate fully parameterized activity surfaces at the participant level. We embed practical and theoretical requirements by imposing constraints that guarantee non-negativity and non-crossing properties in the direction of the quantile order, respectively. We achieve this by employing a dimensionality reduction algorithm that enforces non-negativity throughout and incorporates a parametric spline-based score structure that is monotonic across different quantile levels. We assess the proposed methodology in terms of the precision and stability of the estimated quantile surfaces. Our findings indicate that our estimations are not only more rational but also more accurate compared to alternative approaches.

1 Introduction

1.1 Motivation

Wearable devices are an attractive tool for data collection due to their ability to measure a broad range of biosignals relevant to the study of health, often at high temporal resolution and low cost. Accelerometers, for example, provide continuous monitoring of physical activity behaviors in free-living settings, and thereby can be used to characterize 24-hour activity profiles. Modern accelerometers collect 30 or more measurements per second, which can be combined to produce epoch-level summaries of activity like the Euclidean Norm Minus One (ENMO) and the Monitor-Independent Movement Summary (MIMS) (e.g., Bai et al. 2014). In a common analytic approach, epoch-level measurements are categorized as sedentary (≤ 1.5 metabolic equivalents or METS), light (1.5 to 3 METS), moderate (3 to 6 METS), and vigorous (≥ 6 METS); the total time spent in these intensity categories are then associated with health outcomes. However, emerging evidence suggests that both the timing and intensity of physical activity behaviors are important in a range of settings, including all-cause mortality and outcomes related to Alzheimer’s disease (e.g., Smirnova et al. 2020; Ghosal et al. 2022). These results underscore the need for nuanced biomarkers

*Department of Statistics and Data Science, Helmut-Schmidt-University, Hamburg, Germany

[†]Department of Biostatistics, Columbia University, New York, USA

[‡]Department of Medicine, Columbia University, New York, USA

that rigorously summarize engagement in physical activity behaviors across intensity levels while preserving the informative temporal structure of measurements made by wearable devices.

Functional Principal Components Analysis (FPCA) is an established framework for modeling participant-level mean curves (Crainiceanu et al. 2024) but is unable to capture other distributional features. Recently, Méndez Civieta et al. (2024) combined ideas from quantile regression (e.g., Koenker and Bassett 1978) and FPCA to estimate participant-level quantile curves at fixed quantile levels through a method called Functional Quantile Principal Components Analysis (FQPCA). As in FPCA, the approach borrows information across participants to estimate population-level temporal patterns and then estimates scores that reflect the contribution of these patterns to participant-level quantile curves. When applied to accelerometer data, this approach enabled the quantification of participant-level 10th, 50th, and 90th quantile curves, expanding on the more traditional focus on expected value curves. However, the approach by Méndez Civieta et al. (2024) can only be applied at discrete quantile levels. Monotonicity over estimated quantile curves is not enforced, resulting in crossing quantile curve estimates that are theoretically impossible and difficult to interpret. In the context of physical activity data, the lack of a non-negativity constraint allows estimates that are below zero and therefore unreasonable. Lastly, the need to estimate each quantile level separately precludes careful inspection of the full time-by-distribution surface to understand activity profiles in their entirety.

We introduce Functional Quantile Surface Estimation (FQSE) to address current methodological gaps and provide a detailed, complete quantification of participant-level activity from wearable devices. Our approach constructs time-by-distribution surfaces at the participant level to capture physical activity behaviors across a range of intensities. Figure 1 illustrates the goals of our ap-

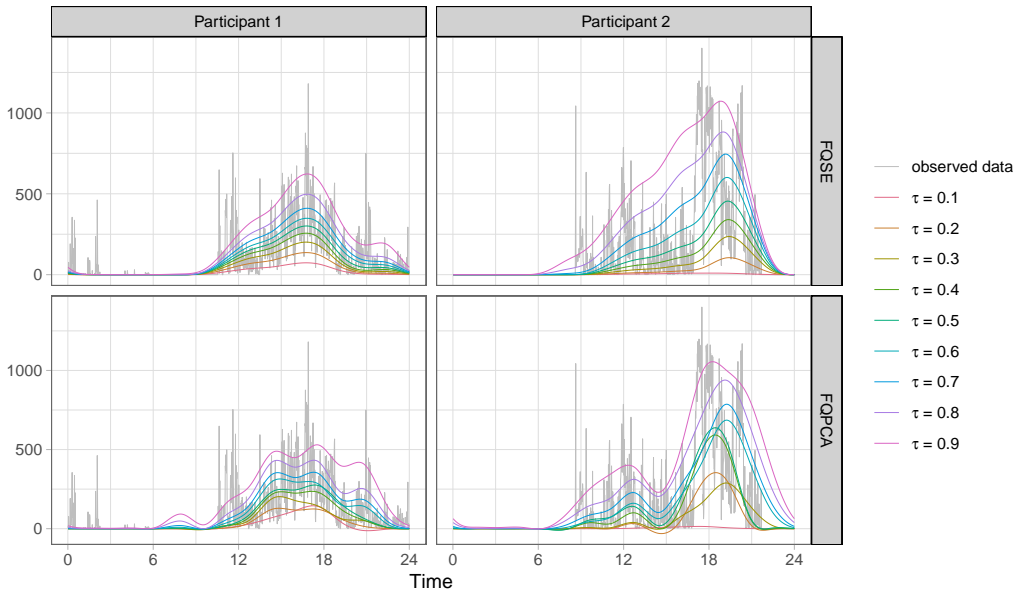


Figure 1: Example estimates by proposed FQSE (top row) in comparison to standard FQPCA (bottom row).

proach. The left and right columns show observed activity data for two participants who have similar low-intensity activity levels but different moderate- and vigorous-intensity physical activity behaviors. In the top row, we overlay the time-by-intensity surfaces produced by our FQSE approach, and in the bottom row, we show results obtained using FQPCA. Although results are

presented at the same discrete collection of quantiles, we emphasize that our methods allow the evaluation of the full time-by-intensity surface while the method in Méndez Civieta et al. (2024) has to be re-applied at each level. Moreover, our methods address the issues discussed above: the quantile surfaces obtained by our approach are non-negative and monotonically increasing in the direction of the quantile level, while the results from existing methods show both negative values and crossing quantile curves. Lastly, we note that our approach is flexible enough to capture distinctive participant-level behaviors.

At the participant level, data consist of single measurements of activity intensity in each epoch. Our perspective is that the observed data point in an epoch represents an individual draw from the participant- and time-dependent distribution of possible activity intensities, and our goal is to reconstruct these distributions in a way that enforces scientifically meaningful constraints. To accomplish this, we construct participant-level time-by-quantile surfaces that are bivariate functions of time and quantile level which can be evaluated for any combination of these. We assume that temporal patterns describing differences across participants are non-negative and the same for all quantile levels, and that participant-level scores on these patterns are monotonically increasing in quantile level. This combination ensures that the participant-level surface estimates are non-crossing and non-negative, and provides a parametric representation of the surface that can be evaluated at any desired quantile level.

Before providing details of the FQSE model construction, we will provide an overview of previous work in the different areas that inform our work.

1.2 Review of Literature

Our work builds on the FQPCA algorithm by Méndez Civieta et al. (2024), which poses a model for participant-level quantile curves that mimics the truncated Karhunen-Loève (KL) expansion commonly used to model expected value curves in FPCA. The model includes a quantile-specific population-level intercept function, quantile-specific principal components, and participant-level scores. While many conventional FPCA methods use an eigendecomposition of a (smoothed) covariance surface (e.g., Yao et al. 2005), FQPCA employs a variation of the probabilistic principal components analysis framework introduced by Tipping and Bishop (1999). That is, FQPCA estimates the quantile-specific parameters in the expansion – the intercept, functional quantile principal components, and scores – through an iterative algorithm that bypasses the need to estimate and decompose a covariance matrix.

We add constraints to FQPCA first by utilizing non-negative matrix factorization (NMF); see Lee and Seung (1999) for an introduction and Wang and Y.-J. Zhang (2013) for a comprehensive review. In NMF, data are decomposed into a product of two non-negative matrices, one of which captures shared features among participants and the other of which represents the contribution of each shared feature to a specific observation. The non-negativity constraints imposed by NMF are intuitive in many settings and often produce more interpretable results than other dimension reduction methods. When analyzing wearable device data, for example, Backenroth et al. (2020) combined NMF methods and a functional data approach to estimate participant-level expected value curves; the approach yielded population-level patterns that were more interpretable than those obtained through (generalized) FPCA. While the conditions guaranteeing that NMF solutions are unique are strict and may not be met in practice (Huang et al. 2014), in many cases the results are numerically stable, accurate in simulations, and robust to random starting conditions.

Next, we take steps to ensure that participant-specific estimated surfaces are monotonically increasing across quantile levels; that is, we require that the participant-level quantile curves do not overlap. Methods to produce non-crossing quantiles have been extensively studied in the non-functional data settings, and popular approaches can be roughly divided into three categories. In post-estimation rearrangement, quantiles are first estimated via a standard procedure resulting in crossing curves, and then are transformed to follow a monotone structure (e.g., Chernozhukov et al. 2010). Next, in constrained stepwise estimation, inequality constraints are established to guarantee monotonicity in the quantile direction (e.g., Bondell et al. 2010). In a simultaneous estimation framework Schnabel and Eilers (2013) proposed to estimate multiple non-crossing quantile curves jointly using scalar covariate and outcome data across participants to produce a so-called quantile sheet. This sheet is a bivariate function of the (single, scalar) covariate and the quantile order; it is estimated using a penalization strategy that ensures the sheet is monotonically increasing in the direction of the quantile order.

There are no existing methods in functional data analysis to rigorously estimate participant-specific time-by-intensity surfaces that are non-negative and non-crossing across quantile levels. We have previously noted the work by Méndez Civieta et al. (2024), which only estimates individual quantile curves but not surfaces and does not include non-negativity constraints for data on physical activity behaviors. In order to obtain time-by-distribution surfaces that could be used as predictors of scalar outcomes, Ghosal et al. (2022) estimated quantile levels in a moving window and using data from multiple days, but did not pose a model for time-by-intensity surfaces or borrow information across participants. Staicu et al. (2012) addressed within-function correlation and temporally and spatially varying skewness, but their quantile model used a parametric approach that estimated time-varying shape parameters with the main purpose of characterizing the dependence of functional processes. Other methods that combine functional data analysis and quantile modeling focus on inherently different problems, mostly situated within regression contexts. These include scalar-on-function regression models (e.g., Cardot et al. 2005), semiparametric approaches for functional linear models (e.g., Lu et al. 2014), and function-on-scalar regression models (e.g., Brockhaus et al. 2015) that focus on covariate effects.

The rest of the paper is organized as follows. The next section covers the model structure and the estimation procedure, including a conceptual perspective and the computational estimation algorithm. The following simulation section evaluates the new method and compares it to existing approaches using simulations. We then apply the algorithm to accelerometer data gathered as part of the National Health and Nutrition Examination Survey (NHANES) data set (Troiano et al. 2008). Lastly, results and further research questions are discussed.

2 Models

2.1 Model structure

Let $X_i(t)$ be a function observed over $t \in \mathcal{T}$, where \mathcal{T} is a bounded closed interval. For participants $1 \leq i \leq I$, we will model the relationship between observed functions and underlying quantile surfaces through

$$X_i(t) = Q_i(t, \tau) + u_i(t, \tau),$$

where the smooth bivariate surface $Q_i(t, \tau)$ depends on t and the quantile level $\tau \in [0, 1]$, and $u_i(t, \tau)$ is the time- and quantile-specific idiosyncratic error.

For each fixed t , $Q_i(t, \cdot)$ characterizes the complete participant-specific distribution at that time point. The observed value $X_i(t)$ is a single draw from the time- and participant-specific distribution described by $Q_i(t, \cdot)$, which we will model without making parametric assumptions. Meanwhile, for each fixed τ , $Q_i(\cdot, \tau)$ captures the diurnal patterns at a certain quantile level. Put differently, for a fixed τ , $Q_i(\cdot, \tau)$ is the participant-specific function over t that describes diurnal activity patterns at a specific intensity level. This framing poses observed data $X_i(t)$ as arising from a smooth latent process with additional noise; unlike most models in FDA, however, the latent smooth curve is quantile-dependent rather than an expected value.

Quantile surfaces $Q_i(t, \tau)$ will be estimated using a combination of population- and participant-level components. Specifically, let

$$Q_i(t, \tau) = \sum_{k=1}^{K_\phi} \lambda_{ik}(\tau) \phi_k(t) =: \boldsymbol{\lambda}_i(\tau)' \boldsymbol{\phi}(t), \quad (1)$$

where $\boldsymbol{\phi}(t) = [\phi_1(t), \dots, \phi_{K_\phi}(t)]'$ is a population-level collection of K_ϕ non-negative loadings and $\boldsymbol{\lambda}_i(\tau) = [\lambda_{i1}(\tau), \dots, \lambda_{iK_\phi}(\tau)]'$ is the corresponding participant-level collection of K_ϕ non-negative and monotonically increasing score functions. We note that the loadings $\phi_k(t)$ depend only on t and are fixed across quantile levels; these are understood to collect the features shared among participants. The score functions $\lambda_i(\tau)$ depend only on τ and do not vary with respect to t , and represent the contribution of the loadings $\phi_k(t)$ to a specific participant i . The loadings $\phi_k(t)$ and score functions $\lambda_{ik}(\tau)$ are assumed to be smooth curves in the spaces $L_2\mathcal{T}$ and $L_2[0, 1]$ of square-integrable functions, respectively. Although this assumption is in general not sufficient to guarantee that $Q_i(t, \tau)$ is in $L_2\mathcal{T} \times [0, 1]$, we will model loadings and scores in a way that the resulting surface is indeed smooth.

Model parameters can be estimated by minimizing the cumulative quantile loss over both time t and quantile τ , i.e., by choosing loadings and scores so that the objective function

$$M(\boldsymbol{\Lambda}(\tau), \boldsymbol{\phi}(t)) = \frac{1}{I} \sum_{i=1}^I \int_0^1 \int_{\mathcal{T}} \rho_\tau(X_i(t) - \boldsymbol{\lambda}_i(\tau)' \boldsymbol{\phi}(t)) dt d\tau \quad (2)$$

is minimized. Here, $\rho_\tau(u) = u(\tau - \mathbf{1}(u < 0))$ denotes the quantile regression check loss function or tilted absolute value function as introduced by Koenker and Bassett (1978), $\mathbf{1}(\cdot)$ is the indicator function, and $\boldsymbol{\Lambda}(\tau) = \{\boldsymbol{\lambda}_1(\tau), \dots, \boldsymbol{\lambda}_I(\tau)\}$ the collection of score functions for all participants and any $\tau \in (0, 1)$.

Our model specification, which estimates the bivariate surface over t and τ through two collections of univariate functions, facilitates a straightforward sequential estimation strategy. First, the non-negative loadings $\phi_k(t)$ are estimated by fixing the quantile level at a pre-specified value; $\boldsymbol{\phi}(t)$ does not depend on τ , and estimates obtained for a specific value can then be used to construct the surface in (1). For this step, described in Section 2.2, we use a quantile-specific model that is similar to the one in Méndez Civieta et al. (2024) but which incorporates non-negativity constraints and thus requires a new estimation strategy. Second, using techniques described in Section 2.3, we estimate participant-level score functions $\boldsymbol{\lambda}_i(\tau)$ given the loading functions $\boldsymbol{\phi}(t)$ obtained in the first step. We again constrain non-negativity and further enforce monotonicity across quantiles to ensure the resulting quantile surfaces have the desired properties.

2.2 Estimating non-negative loadings at a pre-specified quantile level

We first set a reference quantile level τ_0 in order to estimate the population-level loadings $\phi(t)$. In principle, because the loadings are assumed to be constant across all values of τ , any choice of τ_0 is valid; in practice, we recommend $\tau_0 = 0.5$, i.e., the median, and encourage the use of sensitivity analyses to examine robustness to this choice. By focusing on a single quantile, the objective function in (2) is simplified as we do not require score functions evaluated across the entire quantile domain.

Denote by $\mathbf{t}_i = (t_{i1}, \dots, t_{iJ_i})$ the vector of length J_i that is the discrete time grid over which the trajectory $X_i(t)$ is observed in practice. Correspondingly, denote by $\mathbf{X}_i(\mathbf{t}_i)$ the vector of length J_i where each element is the evaluation of the function $X_i(t)$ over the elements of \mathbf{t}_i . Let $\phi(\mathbf{t}_i) \in \mathbb{R}^{K_\phi \times J_i}$ be the matrix of loadings $\phi(t)$ evaluated over the grid \mathbf{t}_i . We employ a basis representation for the population-level loadings. Specifically, let $\psi(\mathbf{t}_i) \in \mathbb{R}_+^{K_\psi \times J_i}$ be the evaluation of spline basis functions $\psi(t) = [\psi_1(t), \dots, \psi_{K_\psi}(t)]'$ over the observation grid \mathbf{t}_i . We use cubic B-spline basis functions for which $\psi_k(t) \geq 0$ for all k and t . By constraining each coefficient vector $\mathbf{b}_{k,\phi} \in \mathbb{R}_+^{K_\psi}$ corresponding to factor $\phi_k(t)$ to be non-negative, we ensure that the estimated factors are non-negative as well.

Notationally, define $\mathbf{B}_\phi \in \mathbb{R}_+^{K_\phi \times K_\psi}$ as the spline coefficient matrix with column k containing $\mathbf{b}_{k,\phi}$ so that

$$\phi(\mathbf{t}_i) = \mathbf{B}_\phi \psi(\mathbf{t}_i).$$

Let $\text{vec}(\mathbf{B}_\phi) \in \mathbb{R}^{K_\phi K_\psi}$ be the vector created by stacking rows of the matrix \mathbf{B}_ϕ and $[\boldsymbol{\lambda}_{i,\tau_0} \otimes \psi(t_{ij})]'$ the vector of length $K_\phi K_\psi$ that is the participant-specific Kronecker product with subentries $\lambda_{ij}(\tau_0) \psi(t_{ij})'$. We estimate spline coefficients \mathbf{B}_ϕ leveraging an iterative algorithm that optimizes the objective function

$$M(\boldsymbol{\Lambda}_{\tau_0}, \mathbf{B}_\phi) = \frac{1}{I} \sum_{i=1}^I \sum_{j=1}^{J_i} \rho_{\tau_0} (X_i(t_{ij}) - \boldsymbol{\lambda}'_{i,\tau_0} \mathbf{B}_\phi \psi(t_{ij})). \quad (3)$$

This objective function differs from that in (2) through the use of a single reference quantile τ_0 and corresponding score vectors $\boldsymbol{\Lambda}_{\tau_0} := (\boldsymbol{\lambda}_{1,\tau_0}, \dots, \boldsymbol{\lambda}_{I,\tau_0})$, as well as the use of a spline expansion for loading functions. Our iterative procedure alternates between the estimation of spline coefficients \mathbf{B}_ϕ and of the score matrix $\boldsymbol{\Lambda}_{\tau_0}$:

- Given current values of spline coefficients $\hat{\mathbf{B}}_\phi$ we estimate scores $\boldsymbol{\lambda}_i$ using

$$\hat{\boldsymbol{\lambda}}_{i,\tau_0} = \arg \min_{\boldsymbol{\lambda}_i} \frac{1}{J_i} \sum_{j=1}^{J_i} \rho_{\tau_0} (X_i(t_{ij}) - \boldsymbol{\lambda}_{i,\tau_0}' \hat{\mathbf{B}}_\phi \psi(t_{ij})), \quad \text{subject to: } \boldsymbol{\lambda}_{i,\tau_0} \geq 0 \forall i.$$

- Given the current score estimates $\hat{\boldsymbol{\Lambda}}_{\tau_0}$, we estimate spline coefficients using

$$\text{vec}(\hat{\mathbf{B}}_\phi) = \arg \min_{\mathbf{B}_\phi} \frac{1}{I} \sum_{i=1}^I \frac{1}{J_i} \sum_{j=1}^{J_i} \rho_{\tau_0} (X_i(t_{ij}) - [\hat{\boldsymbol{\lambda}}_{i,\tau_0} \otimes \psi(t_{ij})] \text{vec}(\mathbf{B}_\phi))$$

subject to : $\text{vec}(\mathbf{B}_\phi) \geq 0$.

Both estimation steps are standard quantile regressions with additional non-negativity constraints. The initial values for spline coefficients $\hat{\mathbf{B}}_\phi$ are taken to be random non-negative values

drawn from a folded Normal distribution. After each iteration, the objective value (3) is measured, and the algorithm is repeated until the difference between successive objective values falls below a pre-defined convergence threshold. In our implementations, we use the absolute threshold of 0.0001, but note that a measure of relative change in quantile error is a plausible alternative. The selection of tuning parameters K_ϕ and K_ψ is described in section 2.4.

To promote numeric stability, we set

$$\int_0^1 \phi_j(t)^2 dt = 1$$

as an additional constraint, which is enforced by rescaling spline coefficients $\hat{\mathbf{B}}_\phi$ at each iteration. Unlike FPCA methods, the estimated loading functions are not constrained to be orthonormal. Like NMF and similar methods, formal requirements for uniqueness and identifiability of the estimated loadings may not be realistic in practice. That said, our empirical results suggest that loadings can be stably and reliably estimated in a range of settings. We also emphasize that our primary interest is in estimating quantile surfaces, and loading functions obtained through this algorithm are useful for that purpose even without formal results for uniqueness and identifiability.

2.3 Estimation of score functions given non-negative loadings

After obtaining estimates of the non-negative loadings $\hat{\phi}(t_i) = \hat{\mathbf{B}}_\phi \psi(t_i)$ shared across quantile levels, we now estimate the collection of participant-level score functions $\lambda_i(\tau)$ in the specification of quantile surfaces $Q_i(t, \tau)$ in (1). The score functions must be monotonically increasing and non-negative to ensure that quantile surfaces have the desired structure, and are estimated separately for each participant.

Estimation requires that score functions $\lambda_{ik}(\tau)$ are observed over a discrete grid of τ -values $\boldsymbol{\tau} := (\tau_1, \dots, \tau_G)$. Unlike the grids t_i , which are determined by the sampling design of observed data, the grid $\boldsymbol{\tau}$ is selected by the user to balance estimation accuracy and computational feasibility. Monotonicity in score functions will be achieved by using I-Splines (i.e., integrated B-Splines), denoted by $\boldsymbol{\Theta}(\boldsymbol{\tau}) = [\Theta_1(\boldsymbol{\tau}), \dots, \Theta_{K_\Theta}(\boldsymbol{\tau})]'$, with corresponding discrete measurements $\boldsymbol{\Theta}(\boldsymbol{\tau}) \in \mathbb{R}_+^{K_\Theta \times G}$. When combined with non-negative coefficients $\mathbf{B}_{\lambda_i} \in \mathbb{R}_+^{K_\phi \times K_\Theta}$, the resulting score functions are non-negative and monotonically increasing by design. After implementing this spline expansion for each of the K_ϕ participant-specific score functions $\lambda_i(\boldsymbol{\tau})$, we have the compact expression

$$\lambda_i(\boldsymbol{\tau}) = \mathbf{B}_{\lambda_i} \boldsymbol{\Theta}(\boldsymbol{\tau}).$$

Combining this expansion for participant-level score functions and the estimates of population-level loading functions obtained in Section 2.2 yields the estimation problem

$$\hat{\mathbf{B}}_{\lambda_i} = \arg \min \frac{1}{J_i} \sum_{j=1}^{J_i} \sum_{g=1}^G \rho_{\tau_g}(X_i(t_{ij}) - (\mathbf{B}_{\lambda_i} \boldsymbol{\Theta}(\boldsymbol{\tau}_g))' \hat{\phi}(t_{ij})), \quad \text{subject to: } \mathbf{B}_{\lambda_i} \geq 0.$$

In contrast to the estimation of spline coefficients for the population-level loadings, this is not a simple quantile regression. The overall optimization function is instead a sum across quantile-specific local optimization functions determined by the grid $\boldsymbol{\tau}$. However, as a sum of standard convex quantile regressions, the proposed optimization function is itself convex. The grid $\boldsymbol{\tau}$ is required to be dense enough to allow score functions $\lambda_i(\boldsymbol{\tau})$ to be modeled with sufficient flexibility,

and K_Θ should be chosen to ensure that participant-specific variability can be flexibly modeled. Recommendations for these tuning parameters are provided in section 2.4.

After obtaining the estimated spline coefficients for participant-level score functions, we can produce complete quantile surfaces using

$$\hat{Q}_i(t, \tau) = \hat{\lambda}_i(\tau)' \hat{\phi}_i(t) = (\hat{\mathbf{B}}_{\lambda_i} \Theta(\tau))' \hat{\mathbf{B}}_\phi \psi_i(t).$$

Again, given the structure of the spline bases and the constraints on spline coefficients, these surfaces are non-negative and monotonically increasing in τ at every time point t . We also note that the surfaces can be evaluated at any combination of t and τ without any additional estimation or model fitting.

2.4 Computational aspects and tuning parameter settings

Our estimation algorithm was implemented using R. The estimation of the loading functions at the reference level τ_0 , as outlined in Section 2.2, uses the `quantreg` package (Koenker 2022), which allows direct incorporation of the necessary non-negativity constraints in the quantile loss function. Because estimation of the score functions sums loss functions over the grid τ as discussed in Section 2.3, we could not leverage standard `quantreg` tools. Instead, we use R's built-in `optim` method, which also allows the incorporation of non-negativity constraints. Because estimation is done separately for each participant, it is possible to parallelize this step. Our methods are stored in a publicly available repository.

The collection of tuning parameters consists of the number of factors K_ψ ; the degrees of freedom K_ψ in the B-spline basis used to model the factors; the degrees of freedom K_Θ in the I-spline basis used to model score functions; and the grid τ . Our analyses of the motivating data suggest that results are consistent across a range of choices for K_Θ and τ because the underlying score functions $\lambda_i(\tau)$ are generally smooth. In particular, setting K_Θ to define a reasonably rich basis expansion (i.e., $K_\Theta = 10$ or 15) and choosing a dense grid τ that roughly spans the quantile axis (i.e., $\tau = [0.05, 0.1, \dots, 0.9, 0.95]$) yielded reasonable estimates without inducing excessive computational demands. We encourage the use of sensitivity analyses when selecting these parameters in other settings.

Smoothness in the direction of time t depends on the number of factors K_ψ and the degrees of freedom K_ψ used to model the factors; choosing these parameters is somewhat more difficult than selecting parameters related to the fit over τ . First, we note that the smoothness-inducing penalties often used in functional data analysis were numerically unstable and computationally inefficient. Next, because loading functions are non-orthogonal and therefore non-sequential, we are unable to directly use a percent variance explained threshold. Instead, we tuned these parameters using a 5-fold cross-validation procedure. Each fold selects four-fifths of the observed time points of each curve as training data, treating the remainder as the testing set. Loadings and score functions are estimated as described in 2.2 and 2.3 for a specific quantile τ_0 using the training data, and the fit is evaluated by calculating the quantile error $\int_0^{24} \rho_\tau \left(X_i(t) - \hat{Q}_i(t, \tau) \right) dt$ on the remaining testing points. Averaging across participants and folds then yields a single error value for each K_ψ and K_ϕ . In our applications, we observed that a small selection/range of $K_\psi \in \{10, 15, 20, 25\}$ appears to be enough to capture the main proportions of variability in the data. Furthermore, while we again recommend setting $\tau_0 = 0.5$, the choice of τ_0 did not affect results significantly.

As we aim to retain enough components to get reasonable fits but also stay parsimonious in our estimates, we propose to then create scree-like plots that depict the percentage of quantile error explained when using a specific $K_\phi \geq 1$, compared to an intercept-only model, i.e.,

$$Q_i(t, \tau) = \phi_0(t)$$

where $\phi_0(t)$ is a non-negative loading function of time t and estimated analogously to the estimation procedure described in 2.2 with scores being equal to 1 throughout. This procedure allows us to assess how much quantile error is explained sequentially until it plateaus, allowing for an appropriate setting of K_ϕ that does not necessarily results in the lowest overall quantile error but balances the accuracy of fits and parsimony. Corresponding visualizations for our simulations and the real data example are presented in supplementary Figures 5 and 6.

3 Simulations

We aim to assess the estimation accuracy of the proposed FQSE approach on simulated datasets that are designed to mimic key features of activity profiles observed in practice. We also compare the performance of our novel approach to that of the existing FQPCA method Méndez Civieta et al. (2024).

3.1 Simulation Design

Observed activity trajectories in simulated datasets are generated to reflect typical activity patterns. Levels are lower at the beginning and end of the domain (i.e. in the "early morning" and "nighttime" periods of the "day"). Higher activity levels are observed in two peaks, centered in the mid-morning and mid-afternoon, which are separated by a mid-day dip. Lastly, activity levels differ across participants.

We construct latent quantile surfaces that underlie simulated trajectories $X_i(t)$, $i = 1, \dots, I$, by setting

$$Q_i(t, \tau) = \boldsymbol{\lambda}_i(\tau)' \boldsymbol{\phi}(t).$$

The non-negative loading functions $\boldsymbol{\phi}(t)$ and corresponding score functions $\boldsymbol{\lambda}_i(\tau)$, described below, generate data with the desired diurnal activity patterns. Given these, we sample data by setting the time grid $\mathbf{t} = (t_1, \dots, t_J)$, common across participants, to consist of equally-spaced time points over $[0, 24]$. For each participant and time point t_j , we sample a single observation from the time-specific distribution by sampling a quantile τ from a uniform distribution over $[0, 1]$, evaluating the participant-specific score functions at that quantile, and combining with population-level loadings:

$$X_i(t_j) = \boldsymbol{\lambda}_i(U_i(t_j))' \boldsymbol{\phi}(t_j) \quad U_i(t_j) \sim \mathcal{U}[0, 1].$$

We create the two peaks intended to mimic diurnal activity patterns by setting $\boldsymbol{\phi}(t) = [\phi_1(t), \phi_2(t)]$, where $\phi_1(t) = P_1(t) \cos(t/3)^2$ controls a peak in morning activity, and $\phi_2(t) = P_2(t) \sin(0.6(t/3) - 4.7)^2$ affects afternoon activity. In defining these functions, P_1 is a skewed normal density function with scale 4.5, location 13.5, and $\alpha = -3$, and P_2 is a skewed normal density function with scale 6, location 10.5 and $\alpha = 5$. We then generate score functions to induce variability across simulated participants. We let $\lambda_{i1}(\tau) = 15S(\tau, \alpha_{i1}, \beta_{i2})$ and $\lambda_{i2}(\tau) = 9S(\tau, \alpha_{i2}, \beta_{i2})$, where

$S(x, a, b) = 1/(1 + e^{-a(x-b)})$ is the sigmoid function; $\alpha_{i1}, \alpha_{i2} \sim \mathcal{U}[0, 30]$ are factors affecting the sigmoidal slope; and $\beta_{i1}, \beta_{i2} \sim \mathcal{U}[0, 2]$ are sigmoidal shifting effects regulating the presence of low-level activity across samples. Note that the range of the sigmoid function S is $[0, 1]$, no matter how a, b are chosen.

Figure 2 illustrates some key features of our data generation process. The top row shows example curves taken from a single simulated data with $I = 50$ curves observed over a grid of length $J = 100$. There is a relatively narrow peak corresponding to initial (morning) activity and a wider peak corresponding to afternoon/evening activity. There is variability across participants in the magnitude of each peak as well as in the distribution of observed data under each peak. There are low levels of activity in the early morning and late night for each participant. The second row displays the true underlying quantiles and illustrates the sigmoidal structure used to induce differences across subjects. For instance, note that subjects S3 and S5 have similar 90th quantiles but the lower quantiles are more compressed for S5 than for S3; these differences are reflected in the observed data.

We create 100 datasets for each combination $I = \{50, 100, 200\}$ and grid lengths $J = \{100, 200\}$ in order to investigate the performance of our proposed method. For each dataset, we deploy the methods described in 2.2 to estimate quantile surfaces using tuning parameters informed by the cross-validation procedure for the illustrative data set; information on this tuning parameter selection is stated below. We compare our approach to the methods of Méndez Civieta et al. (2024) in terms of estimation accuracy over a discrete grid of quantile levels, with accuracy evaluated in terms of mean integrated squared error (MISE). For each participant i and quantile level τ , the integrated squared error computes $\int_0^{24} [Q_i(t, \tau) - \hat{Q}_i(t, \tau)]^2 dt$. The MISE is taken across all subjects in each dataset.

We choose the tuning parameters K_ϕ, K_ψ, K_Θ and the discrete grid τ following the recommendations in section 2.4. Determining a plausible number of loading spline coefficients K_ψ involves leveraging a 5-fold cross-validation procedure on a single simulated dataset for each combination I, J . We vary the number of factors and use the quantile error, see section 2.4, as a goodness of fit metric on the test data to select the optimal number of spline coefficients among the choices of $K_\psi \in \{10, 15, 20, 25\}$ by comparing mean errors across samples for the median quantile, i.e., $\tau = 0.5$. A robustness check confirmed that setting $K_\psi \geq 15$ and $K_\phi \geq 2$ effectively captures the majority of the activity variability across samples, compare Figure 5 in the supplement. After ensuring that the accuracy and stability of the final estimates are comparable for all these settings, we select $K_\psi = 15$ and $K_\phi = 3$ for the subsequent analyses below. We then tested different settings for $K_\Theta \in \{10, \dots, 15\}$. As the choice does not affect the accuracy of our results, we opt for $K_\Theta = 11$, representing a reasonably high number of score spline coefficients, and $\tau = 0.05, 0.1, \dots, 0.95$. For the comparative results provided by the method by Méndez Civieta et al. (2024), we chose tuning parameters according to the cross-validation routine suggested in their paper which is an analog to ours.

3.2 Simulation results

FQSE and FQPCA estimates corresponding to the illustrative samples in the top row of Figure 2 can be found in the third and fourth rows in that Figure, respectively. To facilitate visual inspection and comparison, we limit the number of displayed quantile levels to $\tau \in \{0.1, \dots, 0.9\}$. Heuristically speaking, our method yields visually reasonable results that closely align with the true participant-

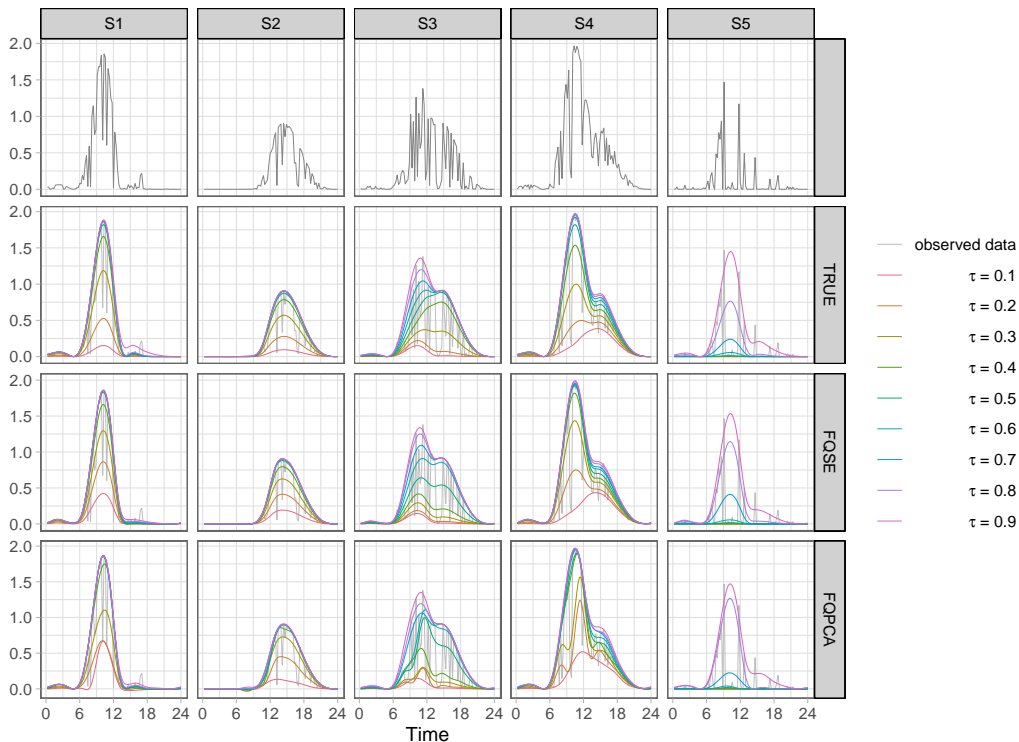


Figure 2: Example estimated curves with true quantiles.

specific quantile surfaces shown in the second row. As anticipated, FQSE yields quantile curves that are both non-negative and non-crossing, which can improve estimation accuracy especially for quantiles near 0. While FQPCA is able to capture the main features, the estimated quantile curves appear to be less accurate in general; the estimated curves can also be negative (samples S1 and S2) or crossing (S1, S3, S4, and S5).

Figure 3 shows the distribution of MISE across 100 simulated data sets for each combination of sample sizes and grid lengths. As in Figure 2, we report MISEs for quantile levels $\tau \in \{0.1, \dots, 0.9\}$.

These results, which are consistent with those shown in Figure 2, suggest that FQSE reliably outperforms FQPCA in terms of estimation accuracy. We attribute this improvement in performance to the borrowing of information across quantiles when estimating scores, as well as to the non-negativity and non-cross constraints. For both methods, estimation errors are lowest for small and large quantile levels due to the sigmoidal structure of the score functions. Increases in the sample size and the grid length improve estimation accuracy for both methods. We also note that the error distributions are skewed and contain a number of outliers; this is somewhat more pronounced for FQPCA than for FQSE, as well as for smaller sample sizes and grid lengths.

4 Analysis of Accelerometer Data from the NHANES

We apply our method to accelerometer data gathered in the National Health and Nutrition Examination Survey (NHANES) (Troiano et al. 2008), in which data are collected on a nationally representative sample of adults and children in the United States. In two waves in 2003-2004 and 2005-2006, participants were asked to wear uniaxial ActiGraph GT1M accelerometers at the hip and the respective minute-level activity step counts were tracked over the course of seven consec-

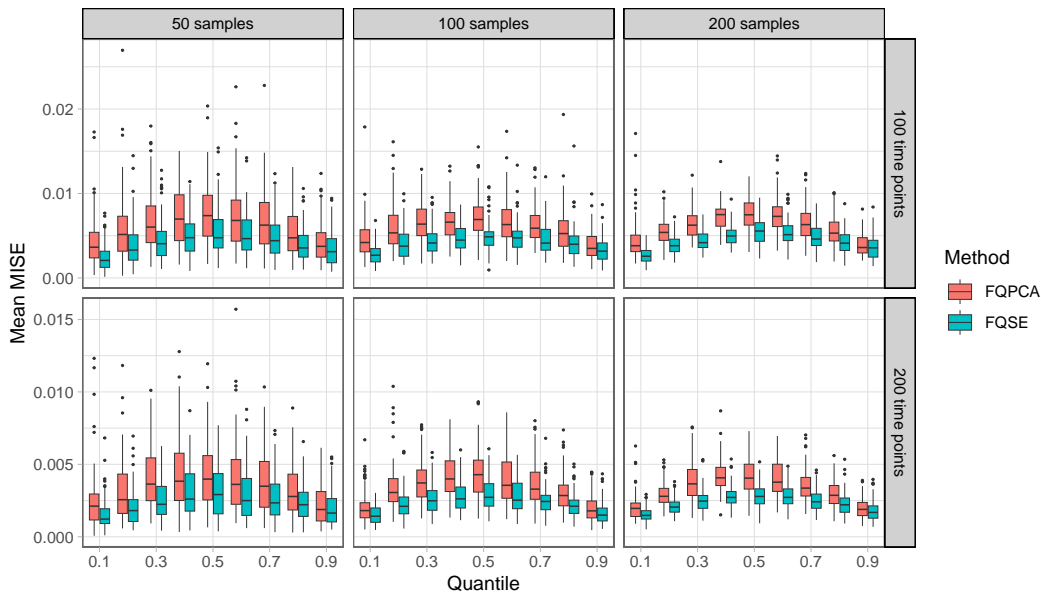


Figure 3: Average MSE across 100 data sets of simulated data with varying numbers of sample sizes and time points. For each data set, the sample-specific MSE was computed and the average taken across samples.

utive days. We access and organize the NHANES data using the R package `rnhanesdata` (Leroux et al. 2018). We summarize them by averaging across days and include only samples with data available across all seven days for consistency throughout the analysis data, resulting in $I = 6849$ participants to consider. The minute-level recordings result in 1440 observations per participant.

To restrain the computational demand, we did not run a full 5-fold cross-validation procedure as in the simulations, but opted for evaluating tuning parameters on a single training and testing split of a reduced data set. Specifically, we selected 500 participants and downsampled their recordings by extracting one sample point per 10-minute interval, performed a rough 80/20% split on the remaining points, and conducted training and testing as before, see section 2.2. This selection of tuning parameters therefore involves a reduced sample size and employs a thinning technique. Upon thorough investigation of the dataset, we determined that the sample size is sufficiently large for accurate assessment of out-of-sample prediction errors and that increasing the sample size is unlikely to alter the chosen tuning parameters. In any case, our results indicated robustness against variations in the parameters. Moreover, the number of time points remains significantly higher than the number of basis functions, K_ψ . This favorable ratio allows us to identify poor fits and achieve satisfactory out-of-sample prediction accuracy.

Inspecting the scree plot of explained quantile error when adding factors, see Figure 6 in the Supplement, revealed, again, that a choice of $K_\psi \geq 10$ for the spline degrees of freedom yielded good estimation results independent of the choice of the number of factors K_ϕ . The scree plot furthermore suggests that a representation using only $K_\phi \geq 4$ factors already explains the vast proportion of overall quantile error ($\geq 90\%$) that cannot be improved much when adding more factors. Therefore, we set $K_\phi = 4$ and $K_\psi = 15$, with the additional benefit that the computational effort is reduced to a few hours, allowing a fast investigation of the most essential data set characteristics.

Figure 4 visualizes the results of our analysis in a way that emphasizes the connection be-

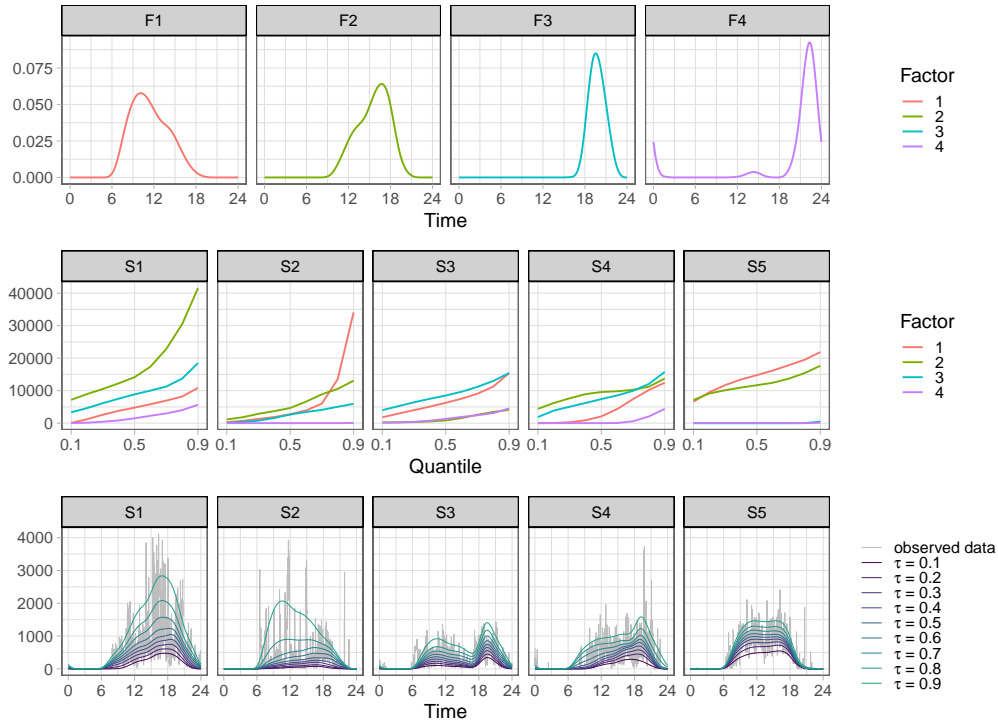


Figure 4: Estimated loadings (top), corresponding scores (middle), and example estimated curves (bottom).

tween our model specification and the estimated quantile surfaces. Estimated loading functions $\phi_k(t)$, which are shared across participants and do not depend on τ , are shown in the top row; participant-specific loading functions $\lambda_k(\tau)$ are shown for five participants in the middle row. The combination of the loading and score functions as in (1) yields estimated quantile surfaces, which are shown for the same collection of participants in the bottom row of Figure 4. A visual comparison of the estimated surfaces and the observed data (shown in grey in the bottom row) suggests that the surfaces capture relevant features. Moreover, an interrogation of the loading and score functions connects these with key characteristics in the data. For participant S2, the score function corresponding to the first factor shows a rapid increase for $\tau > .7$, while all other factor curves remain relatively low. This is reflected in the clear peak for the upper quantiles during the early day, and is consistent with the distribution of observed data in those times. For participant S1, each factor is present to varying degrees. This participant is most active in the late afternoon and early evening, which is reflected in ordering of the score functions (i.e. those associated with factors 2 and 3 have higher magnitudes). Similar observations connecting score functions, reconstructed surfaces, and underlying data distributions can be made for participants S3, S4, and S5, who generally have more moderate activity levels. Our model specification poses bivariate quantile surfaces as the combination of loading functions that are continuous in t and participant-level score functions that are continuous in τ . These continuous functions are shown in the top and middle rows of Figure 4, and can be used to reconstruct the participant-specific quantile surfaces at any combination of t and τ . The display of these surfaces using functions over t at a grid of τ values in the bottom row of Figure 4 is intended to facilitate visual comparisons, and we emphasize that obtaining similar curves using FQPCA would require fitting separate models.

In this example, the interpretation of the different factors is straightforward; they each emphasize activity during a specific time period of the day, and were ordered according to the timing of their most prominent peak. The contribution of these factors in individual quantile surfaces is evident in the corresponding score functions which can be interpreted as above. While estimated factors may not always be so easily understood, interpretability is often a strength of NMF-style methods and similar findings have been observed in related settings (Backenroth et al. 2020). Meanwhile, we reiterate that our primary goal is not to identify interpretable factors but to build a model for quantile surfaces by separating population-level and participant-specific components. Our results indicate that this approach allows for a detailed and scientifically compelling quantification of participant-level activity data. This lays the groundwork for future investigations into the nuanced role of the timing and intensity of physical activity behaviors in health.

5 Discussion

FQSE introduces a novel approach for the estimation of latent participant-level quantile surfaces to simultaneously understand the timing and intensity of physical activity behaviors. We use an NMF-like model structure to facilitate the population-level loading functions that capture patterns of variation across time. We additionally model participant-level scores as non-negative and smooth monotonic functions of the quantile level. Simulations support both visual plausibility and quantitative effectiveness, and our analysis of accelerometer data from the NHANES yields detailed participant-level quantification of activity timing and intensity.

Our method leverages a two-step estimation approach: Loading functions are calculated at a fixed quantile, and subsequently score functions are derived given the loading functions. We considered a joint estimation but found the sequential approach to be notably more computationally feasible and effective in practice. We employ a single reference quantile to estimate loading functions, but integrating over the quantile level could contribute to better function and surface estimates. Lastly, we note that translating the stepwise or additive constrained estimation procedure as proposed by Bondell et al. (2010) and Muggeo et al. (2013) to a functional PCA-style framework would circumvent the need to choose a reference quantile level but would not lead to a fully parameterized bivariate surface function.

Nonparametric assumptions could further enhance the flexibility of the model in terms of capturing participant-specific variation. Adopting alternative models for the score functions may also mitigate the need for additional tuning parameters. In general, finding an "optimal" number of loadings in NMF is difficult to determine as opposed to PCA as factors are not estimated sequentially. Here, we recommend choosing K_ψ using cross-validation for a few values of K_ϕ for some fixed quantile level τ . Altering this level did not influence the outcome notably in our analyses. Further discussions on this topic can be found in Wang and Y.-J. Zhang (2013). Additionally, our model requires setting the number K_Θ and the grid τ of quantile levels to consider for the estimation of the score functions. However, we found that the number of K_Θ can be chosen moderately without compromising explained variation, and that the grid τ can be chosen comprehensively without increasing computational time drastically.

While we also found that in our applications, the estimated loadings and scores were consistent across random initializations, future work may incorporate constraints to enforce the uniqueness of solutions. However, as stated by Backenroth et al. (2020), existing approaches can be onerously restrictive. Incorporating roughness penalties may also encourage sparsity. Here, we controlled

the smoothness of estimates by varying the degrees of freedom K_ψ which led to accelerated computational speed. In general, we would argue that a much faster computational time compensates for this limitation, especially if one is interested in investigating the resulting individual activity surfaces as opposed to the loading functions themselves.

Possible extensions of the model may consider pre-registration of the underlying functional data (Wrobel et al. 2019; Marron et al. 2015). Participant-specific schedules are likely to cause temporal misalignments and shifts in activity intensities within data sets. Our method expresses these different impacts of temporal variations in the score functions, but additional registration could lead to a more parsimonious basis representation for amplitude and phase variation. However, it is unclear how large the effect on computational cost is. Our model could also be extended to a multilevel or multivariate perspective (Di et al. 2009; Goldsmith et al. 2015; Happ and Greven 2018; J. Zhang et al. 2021). Analyzing the NHANES data from a multilevel perspective would allow to depict participant-level day-to-day variation or day-level deviations. Data sets that include processes from multiple accelerometers could benefit from assessing the joint variation of the involved elements.

In general, all these modifications aim at creating even more profound representations of functional physical activity data, ultimately enabling us to connect temporal and distributional effects with underlying human conditions.

References

- Backenroth, Daniel, Russell T. Shinohara, Jennifer A. Schrack, and Jeff Goldsmith (2020). Non-Negative Decomposition of Functional Count Data. In: *Biometrics* 76.4, pp. 1273–1284. DOI: 10.1111/biom.13220.
- Bai, Jiawei, Bing He, Haochang Shou, Vadim Zipunnikov, Thomas A. Glass, and Ciprian M. Crainiceanu (2014). Normalization and Extraction of Interpretable Metrics from Raw Accelerometry Data. In: *Biostatistics* 15.1, pp. 102–116. DOI: 10.1093/biostatistics/kxt029.
- Bondell, Howard D., Brian J. Reich, and Huixia Wang (2010). Noncrossing Quantile Regression Curve Estimation. In: *Biometrika* 97.4, pp. 825–838. DOI: 10.1093/biomet/asq048.
- Brockhaus, Sarah, Fabian Scheipl, Torsten Hothorn, and Sonja Greven (2015). The Functional Linear Array Model. In: *Statistical Modelling* 15.3, pp. 279–300. DOI: 10.1177/1471082X14566913.
- Cardot, Hervé, Christophe Crambes, and Pascal Sarda (2005). Quantile Regression When the Covariates Are Functions. In: *Journal of Nonparametric Statistics* 17.7, pp. 841–856. DOI: 10.1080/10485250500303015.
- Chernozhukov, Victor, Iván Fernández-Val, and Alfred Galichon (2010). Quantile and Probability Curves Without Crossing. In: *Econometrica* 78.3, pp. 1093–1125. DOI: 10.3982/ECTA7880.
- Crainiceanu, Ciprian M., Jeff Goldsmith, Andrew Leroux, and Erjia Cui (2024). *Functional Data Analysis with R*. CRC Press. 338 pp. ISBN: 978-1-00-385630-6. Google Books: m1cIEQAAQBAJ.
- Di, Chong-Zhi, Ciprian M. Crainiceanu, Brian S. Caffo, and Naresh M. Punjabi (2009). Multilevel Functional Principal Component Analysis. In: *The Annals of Applied Statistics* 3.1, pp. 458–488. DOI: 10.1214/08-AOAS206SUPP.
- Ghosal, Rahul, Vijay R. Varma, Dmitri Volfson, Jacek Urbanek, Jeffrey M. Hausdorff, Amber Watts, and Vadim Zipunnikov (2022). Scalar on Time-by-Distribution Regression and Its Application for Modelling Associations between Daily-Living Physical Activity and Cognitive Func-

- tions in Alzheimer’s Disease. In: *Scientific Reports* 12.1 (1), p. 11558. DOI: 10.1038/s41598-022-15528-5.
- Goldsmith, Jeff, Vadim Zipunnikov, and Jennifer Schrack (2015). Generalized Multilevel Function-on-Scalar Regression and Principal Component Analysis. In: *Biometrics* 71.2, pp. 344–353. DOI: 10.1111/biom.12278.
- Happ, Clara and Sonja Greven (2018). Multivariate Functional Principal Component Analysis for Data Observed on Different (Dimensional) Domains. In: *Journal of the American Statistical Association* 113.522, pp. 649–659. DOI: 10.1080/01621459.2016.1273115.
- Huang, Kejun, Nicholas D. Sidiropoulos, and Ananthram Swami (2014). Non-Negative Matrix Factorization Revisited: Uniqueness and Algorithm for Symmetric Decomposition. In: *IEEE Transactions on Signal Processing* 62.1, pp. 211–224. DOI: 10.1109/TSP.2013.2285514.
- Koenker, Roger (2022). *quantreg: Quantile Regression*. R package version 5.93. URL: <https://CRAN.R-project.org/package=quantreg>.
- Koenker, Roger and Gilbert Bassett (1978). Regression Quantiles. In: *Econometrica* 46.1, pp. 33–50. DOI: 10.2307/1913643. JSTOR: 1913643.
- Lee, Daniel D. and H. Sebastian Seung (1999). Learning the Parts of Objects by Non-Negative Matrix Factorization. In: *Nature* 401.6755 (6755), pp. 788–791. DOI: 10.1038/44565.
- Leroux, Andrew, C Crainiceanu, E Smirnova, and Q Cao (2018). *rnhanesdata: Nhanes accelerometry data pipeline*. In: *R package version 1*.
- Lu, Ying, Jiang Du, and Zhimeng Sun (2014). Functional Partially Linear Quantile Regression Model. In: *Metrika* 77.2, pp. 317–332. DOI: 10.1007/s00184-013-0439-7.
- Marron, J. S., James O. Ramsay, Laura M. Sangalli, and Anuj Srivastava (2015). Functional Data Analysis of Amplitude and Phase Variation. In: *Statistical Science* 30.4, pp. 468–484. JSTOR: 24780816.
- Méndez Civieta, Álvaro, Ying Wei, Keith M. Diaz, and Jeff Goldsmith (2024). “Functional Quantile Principal Component Analysis”. Submitted for Publication.
- Muggeo, Vito M. R., Mariangela Sciandra, Agostino Tomasello, and Sebastiano Calvo (2013). Estimating Growth Charts via Nonparametric Quantile Regression: A Practical Framework with Application in Ecology. In: *Environmental and Ecological Statistics* 20.4, pp. 519–531. DOI: 10.1007/s10651-012-0232-1.
- Schnabel, Sabine K. and Paul H. C. Eilers (2013). Simultaneous Estimation of Quantile Curves Using Quantile Sheets. In: *AStA Advances in Statistical Analysis* 97.1, pp. 77–87. DOI: 10.1007/s10182-012-0198-1.
- Smirnova, Ekaterina, Andrew Leroux, Quy Cao, Lucia Tabacu, Vadim Zipunnikov, Ciprian Crainiceanu, and Jacek K. Urbanek (2020). The Predictive Performance of Objective Measures of Physical Activity Derived from Accelerometry Data for 5-Year All-Cause Mortality in Older Adults: National Health and Nutritional Examination Survey 2003–2006. In: *The Journals of Gerontology: Series A* 75.9, pp. 1779–1785.
- Staicu, Ana-Maria, Ciprian M. Crainiceanu, Daniel S. Reich, and David Ruppert (2012). Modeling Functional Data with Spatially Heterogeneous Shape Characteristics. In: *Biometrics* 68.2, pp. 331–343. DOI: 10.1111/j.1541-0420.2011.01669.x.
- Tipping, Michael E. and Christopher M. Bishop (1999). Probabilistic Principal Component Analysis. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61.3, pp. 611–622. DOI: 10.1111/1467-9868.00196.

- Troiano, Richard P., David Berrigan, Kevin W. Dodd, Louise C. Mâsse, Timothy Tilert, and Margaret McDowell (2008). Physical Activity in the United States Measured by Accelerometer. In: *Medicine & Science in Sports & Exercise* 40.1, pp. 181–188. DOI: 10.1249/mss.0b013e31815a51b3.
- Wang, Yu-Xiong and Yu-Jin Zhang (2013). Nonnegative Matrix Factorization: A Comprehensive Review. In: *IEEE Transactions on Knowledge and Data Engineering* 25.6, pp. 1336–1353. DOI: 10.1109/TKDE.2012.51.
- Wrobel, Julia, Vadim Zipunnikov, Jennifer Schrack, and Jeff Goldsmith (2019). Registration for Exponential Family Functional Data. In: *Biometrics* 75.1, pp. 48–57. DOI: 10.1111/biom.12963.
- Yao, Fang, Hans-Georg Müller, and Jane-Ling Wang (2005). Functional Data Analysis for Sparse Longitudinal Data. In: *Journal of the American Statistical Association* 100.470, pp. 577–590. DOI: 10.1198/016214504000001745.
- Zhang, Jun, Greg J Siegle, Tao Sun, Wendy D’andrea, and Robert T Krafty (2021). Interpretable Principal Component Analysis for Multilevel Multivariate Functional Data. In: *Biostatistics (Oxford, England)* 24.2, pp. 227–243. DOI: 10.1093/biostatistics/kxab018.

Supplement

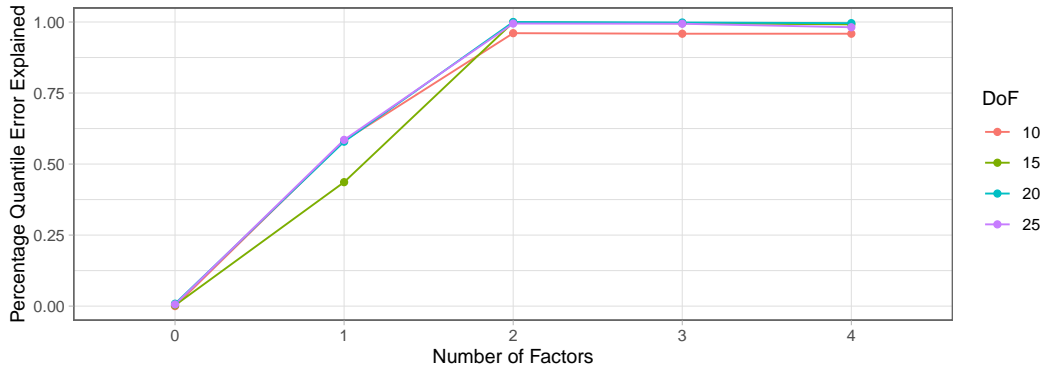


Figure 5: Cross-validated mean quantile error reduction for $\tau = 0.5$ for different numbers of factors. The figure shows the percentage of quantile error explained when increasing the number of factors in reference to the quantile error that occurs when using an intercept-only model (i.e., zero factors). 0% explained error refers to the overall highest occurring error across factors and degrees of freedom $K_\psi \in \{10, 15, 20, 25\}$. Means are taken across samples and folds.

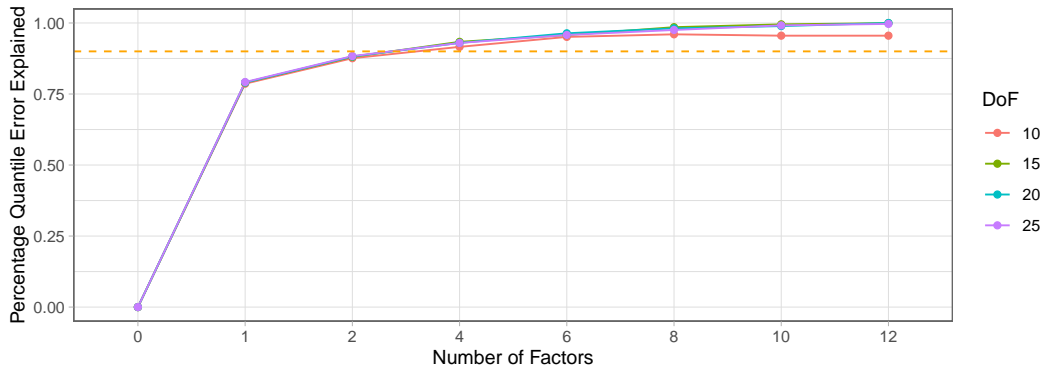


Figure 6: Quantile error for $\tau = 0.5$ for different factors. A subset of $I = 500$ NHANES sample curves was downsampled extracting one value per 10-minute interval resulting in 144 observations per participant. These time points were split into a training (roughly 80%) and a testing portion (roughly 20%). After deploying FQSE on the training curves, the quantile error was calculated on the remaining testing curves, and the average was calculated across participants. The figure shows the percentage of quantile error explained when increasing the number of factors in reference to the quantile error that occurs when using an intercept-only model (i.e., zero factors). 0% explained error refers to the overall highest occurring error across factors and degrees of freedom $K_\psi \in \{10, 15, 20, 25\}$.

Figure 5 and 6 show the quantile error plotted against the varying number of spline degrees of freedom and its robustness with regard to the number of factors for simulated data and a subset of NHANES, respectively.