



Functional Regression Models in Human Movement Biomechanics

Methodology and Applications

Edward Gunning

JHU Biostatistics Wearable and Implantable Technology Group

September 13th 2024

Background

Winter (1979) defines **human movement biomechanics** as

“...the inter-discipline that describes, analyses, and assesses human movement”

Background

Winter (1979) defines **human movement biomechanics** as

“...the inter-discipline that describes, analyses, and assesses human movement”

- Data collected on displacement, velocity and acceleration (**kinematics**) and forces (**kinetics**).

Background

Winter (1979) defines **human movement biomechanics** as

“...the inter-discipline that describes, analyses, and assesses human movement”

- Data collected on displacement, velocity and acceleration (**kinematics**) and forces (**kinetics**).
- Continuously measured throughout a movement → smooth time-varying curves, very suitable for FDA.

Background

Winter (1979) defines **human movement biomechanics** as

“...the inter-discipline that describes, analyses, and assesses human movement”

- Data collected on displacement, velocity and acceleration (**kinematics**) and forces (**kinetics**).
- Continuously measured throughout a movement → smooth time-varying curves, very suitable for FDA.

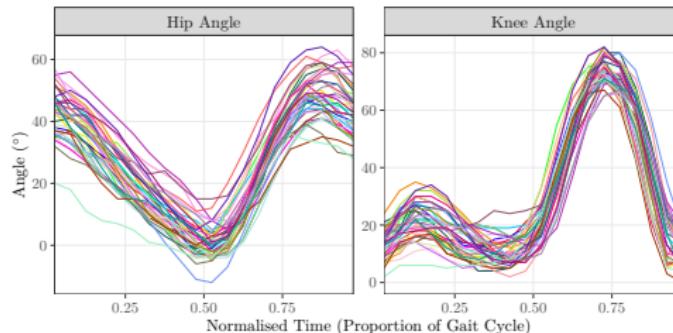


Figure: The childrens' gait dataset (Rice and Silverman, 1991).

Background

Winter (1979) defines **human movement biomechanics** as

“...the inter-discipline that describes, analyses, and assesses human movement”

- Data collected on displacement, velocity and acceleration (**kinematics**) and forces (**kinetics**).
- Continuously measured throughout a movement → smooth time-varying curves, very suitable for FDA.

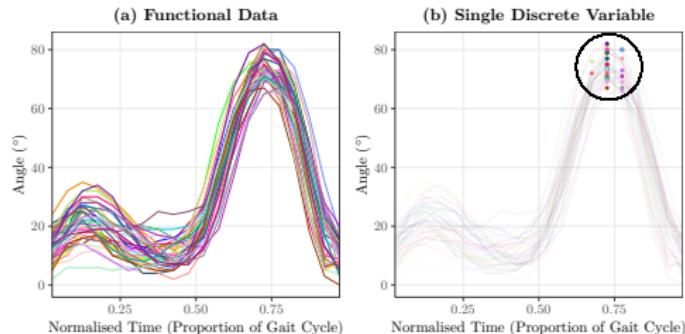


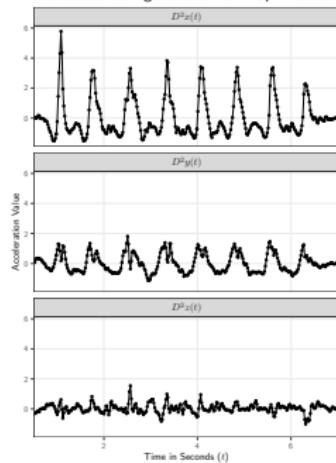
Figure: Data reduction of the knee angle curve to a single discrete variable (peak knee flexion).

Second-Generation Functional Data in Biomechanics

Volume



My Own Motion Data from Apple
Collected During a Short Warm-Up Exercise

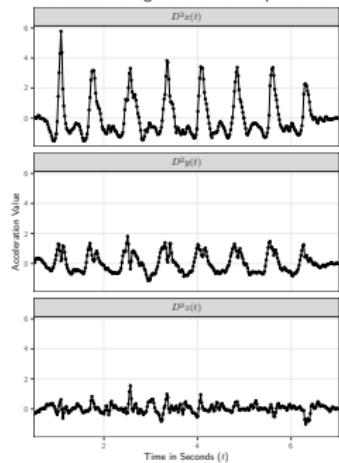


Second-Generation Functional Data in Biomechanics

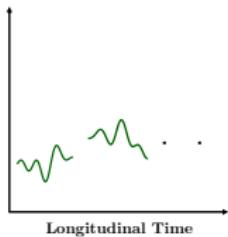
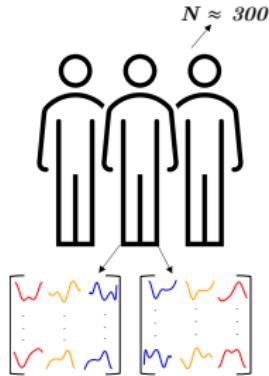
Volume



My Own Motion Data from Apple
Collected During a Short Warm-Up Exercise



Complexity

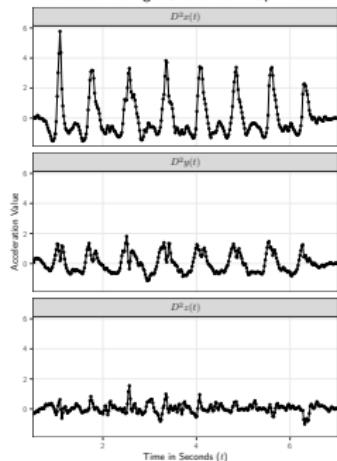


Second-Generation Functional Data in Biomechanics

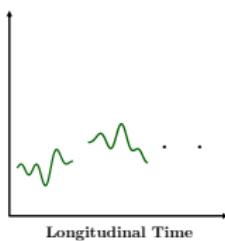
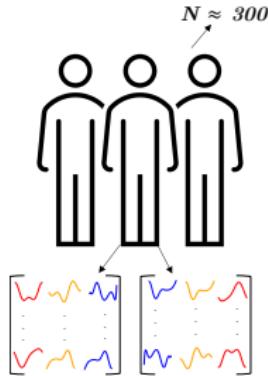
Volume



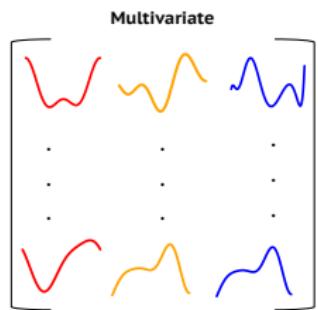
My Own Motion Data from Apple
Collected During a Short Warm-Up Exercise



Complexity



Variety



Human movement as a system of multiple related parts → **multivariate** and **dynamical systems** approaches to analysis.

Compare and Contrast with WIP Work

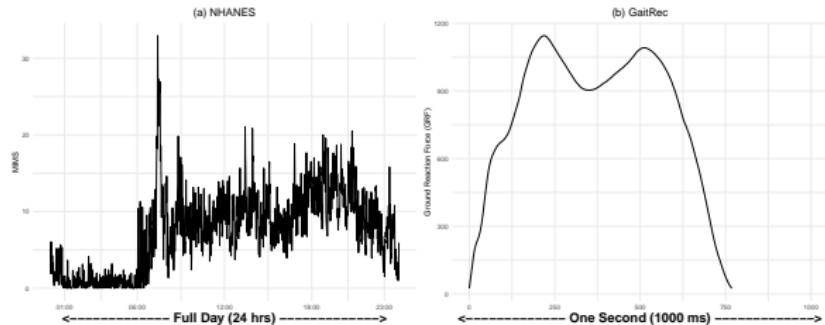


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Compare and Contrast with WIP Work

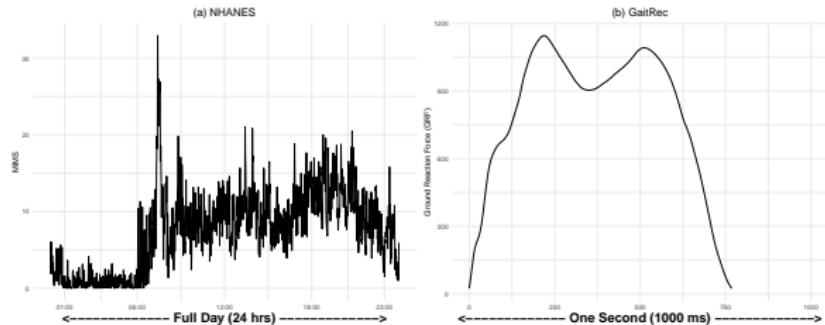


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Differences:

Compare and Contrast with WIP Work

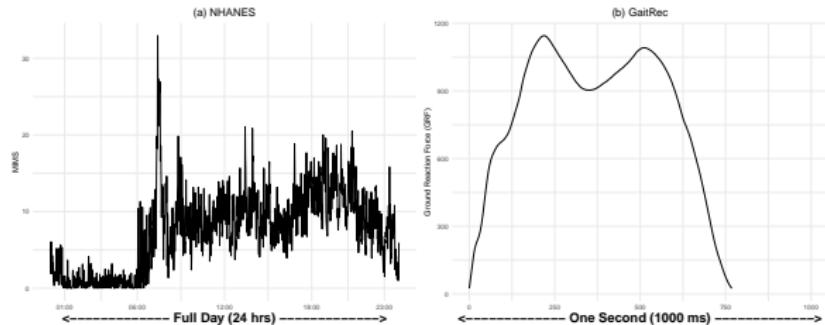


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Differences:

- Movement on macro vs. micro scales.

Compare and Contrast with WIP Work

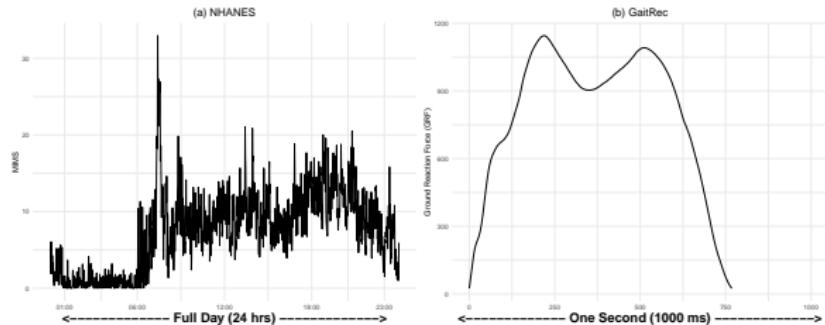


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Differences:

- Movement on macro vs. micro scales.
- Smoothness/ SNR ratios (within and between).

Compare and Contrast with WIP Work

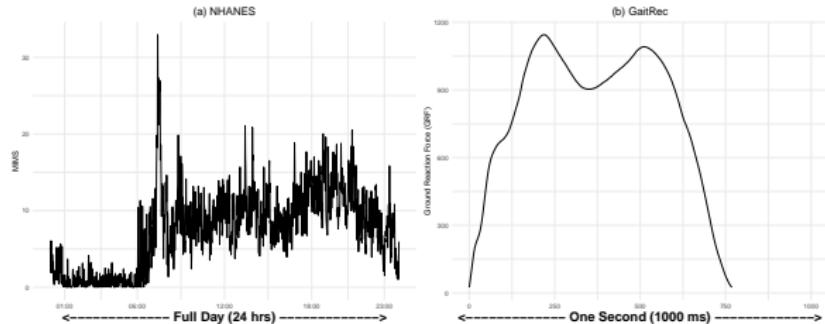


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Differences:

- Movement on macro vs. micro scales.
- Smoothness/ SNR ratios (within and between).

Similarities:

Compare and Contrast with WIP Work

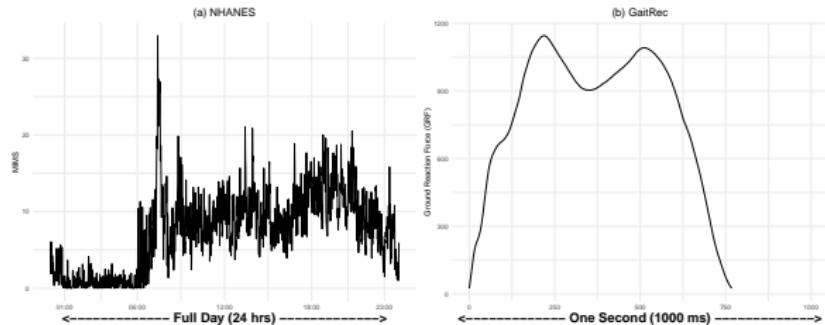


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Differences:

- Movement on macro vs. micro scales.
- Smoothness/ SNR ratios (within and between).

Similarities:

- Complex temporal features (scalar summaries insufficient).

Compare and Contrast with WIP Work

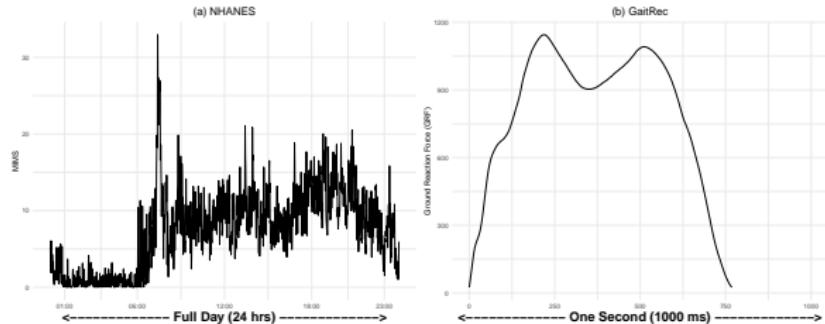


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Differences:

- Movement on macro vs. micro scales.
- Smoothness/ SNR ratios (within and between).

Similarities:

- Complex temporal features (scalar summaries insufficient).
- Structured dependence (e.g., multilevel, longitudinal).

Compare and Contrast with WIP Work

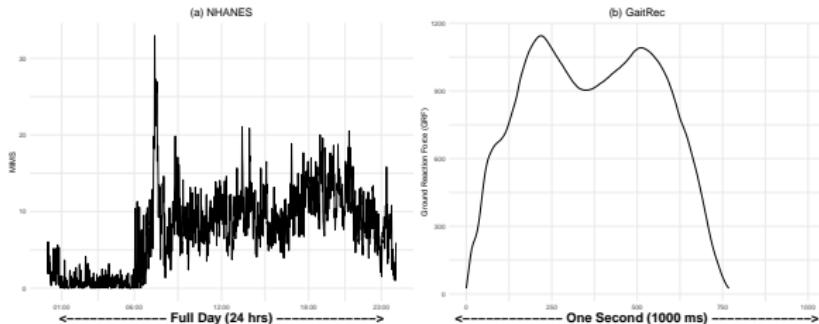


Figure: (a) NHANES (Crainiceanu, Goldsmith, et al., 2024); (b) GaitRec (Horsak et al., 2020).

Differences:

- Movement on macro vs. micro scales.
- Smoothness/ SNR ratios (within and between).

Similarities:

- Complex temporal features (scalar summaries insufficient).
- Structured dependence (e.g., multilevel, longitudinal).

→ We need principled and efficient statistical modelling approaches to extract information from rich, complex and structured functional data.

GaitRec (Horsak et al., 2020)

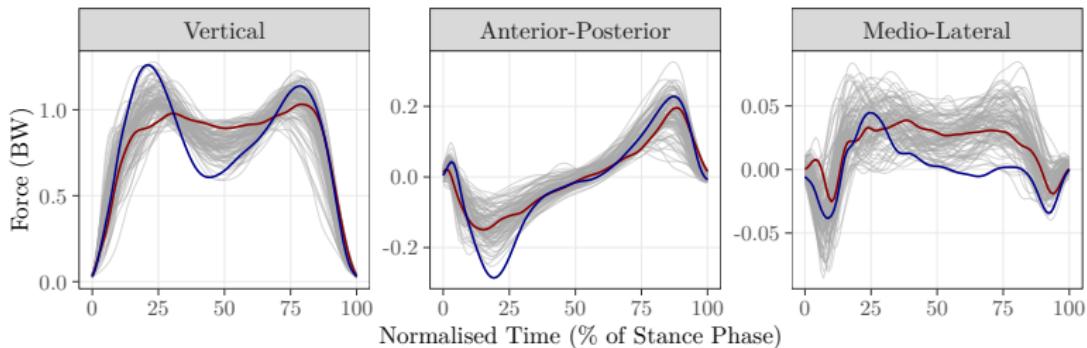


Figure: A sample of observations from the GaitRec dataset (Horsak et al., 2020).

GaitRec (Horsak et al., 2020)

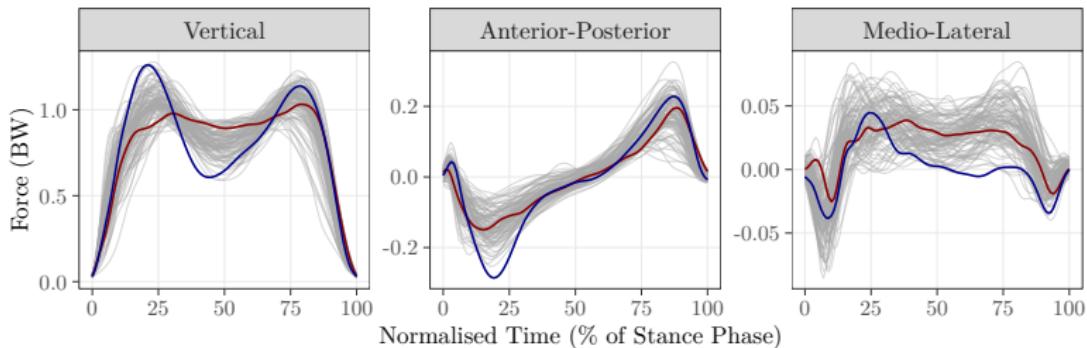


Figure: A sample of observations from the GaitRec dataset (Horsak et al., 2020).

- Ground reaction force and centre of pressure during walking (kinetics).

GaitRec (Horsak et al., 2020)

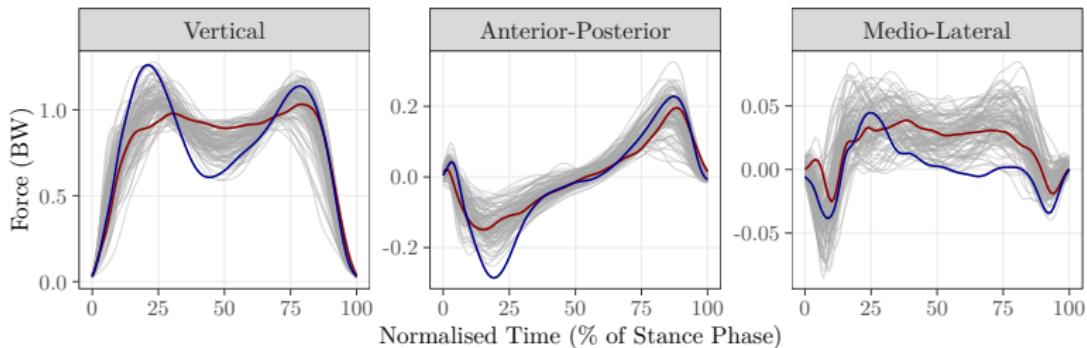


Figure: A sample of observations from the GaitRec dataset (Horsak et al., 2020).

- Ground reaction force and centre of pressure during walking (kinetics).
- 2,295 individuals measured under different conditions (75,732 trials in total).

GaitRec (Horsak et al., 2020)

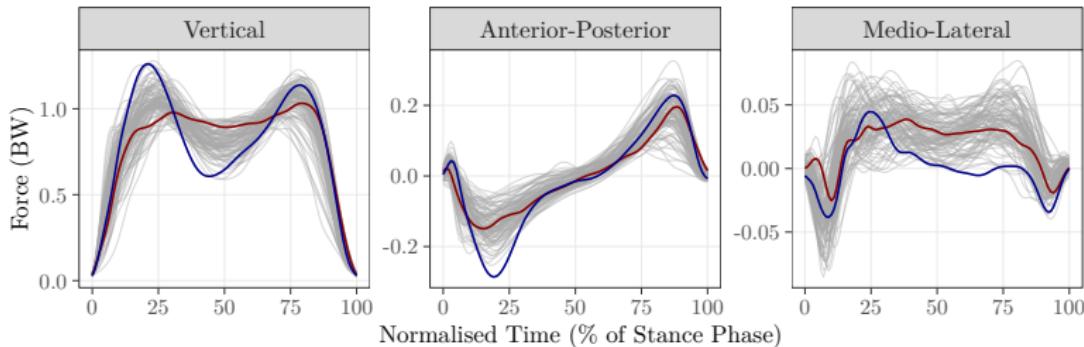
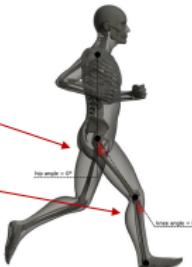
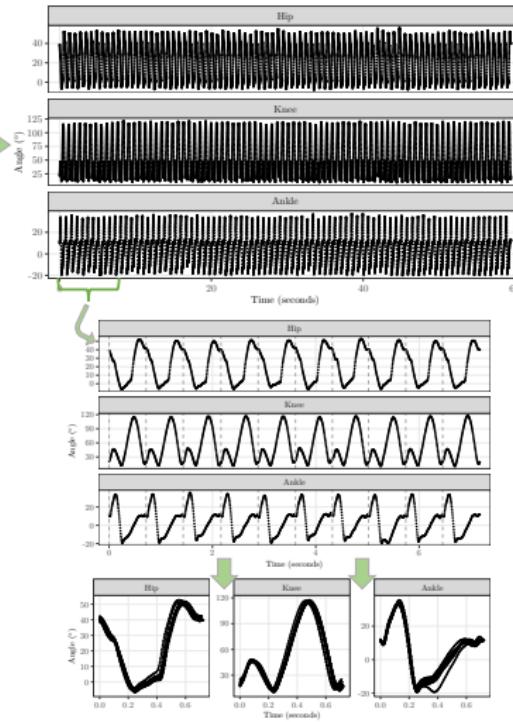


Figure: A sample of observations from the GaitRec dataset (Horsak et al., 2020).

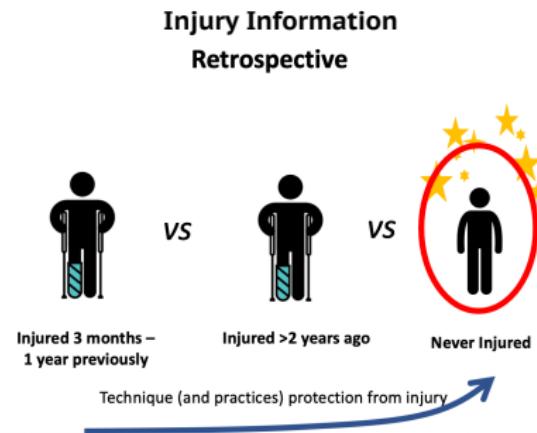
- Ground reaction force and centre of pressure during walking (kinetics).
- 2,295 individuals measured under different conditions (75,732 trials in total).
- Healthy controls and individuals with impairments in the hip, knee, ankle and calcaneus.

Motivating Dataset

Motivating Dataset

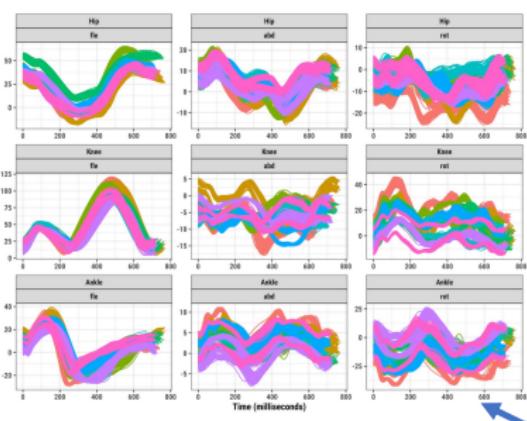


Motivating Dataset

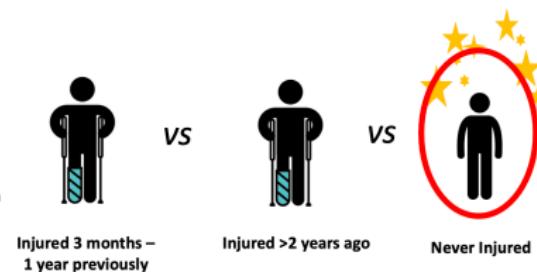


Motivating Dataset

Multivariate Functional Data



Injury Information
Retrospective



Covariates Affecting Running Technique



adjust for

Initial Approach (Gunning et al., 2024b)

- “Start simple” → focus on average bilateral hip and knee flexion angles.
- Regress the average bivariate functions on scalar covariates, e.g., injury status, sex, speed ...

Initial Approach (Gunning et al., 2024b)

- “Start simple” → focus on average bilateral hip and knee flexion angles.
- Regress the average bivariate functions on scalar covariates, e.g., injury status, sex, speed ...

$$\begin{array}{c} \text{Y}(t) \\ \left[\begin{array}{cc} \text{green wavy line} & \text{green wavy line} \\ \text{green wavy line} & \text{green wavy line} \\ \vdots & \vdots \\ \vdots & \vdots \\ \text{blue wavy line} & \text{blue wavy line} \\ \text{blue wavy line} & \text{blue wavy line} \end{array} \right] \end{array} = \begin{array}{c} \text{X} \\ \left[\begin{array}{cc} 1 & \textcolor{red}{1.5} \\ 1 & \textcolor{red}{0} \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & \textcolor{red}{1.2} \\ 1 & \textcolor{red}{0} \end{array} \right] \end{array} + \begin{array}{c} \text{B}(t) \\ \left[\begin{array}{cc} \text{green wavy line} & \text{green wavy line} \\ \text{red wavy line} & \text{red wavy line} \end{array} \right] \end{array} + \begin{array}{c} \text{Z} \\ \left[\begin{array}{ccc} \textcolor{green}{1} & \cdots & 0 \\ \textcolor{green}{1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ 0 & \cdots & 1 \end{array} \right] \end{array} + \begin{array}{c} \text{U}(t) \\ \left[\begin{array}{cc} \text{green wavy line} & \text{green wavy line} \\ \vdots & \vdots \\ \text{blue wavy line} & \text{blue wavy line} \end{array} \right] \end{array} + \begin{array}{c} \text{E}(t) \\ \left[\begin{array}{cc} \text{green wavy line} & \text{green wavy line} \\ \text{blue wavy line} & \text{blue wavy line} \\ \vdots & \vdots \\ \vdots & \vdots \\ \text{blue wavy line} & \text{blue wavy line} \\ \text{blue wavy line} & \text{blue wavy line} \end{array} \right] \end{array}$$

Initial Approach (Gunning et al., 2024b)

Initial Approach (Gunning et al., 2024b)

- Developed a simple, flexible and scalable multivariate functional mixed effects modelling approach based on the basis modelling approach of Morris and Carroll (2006).

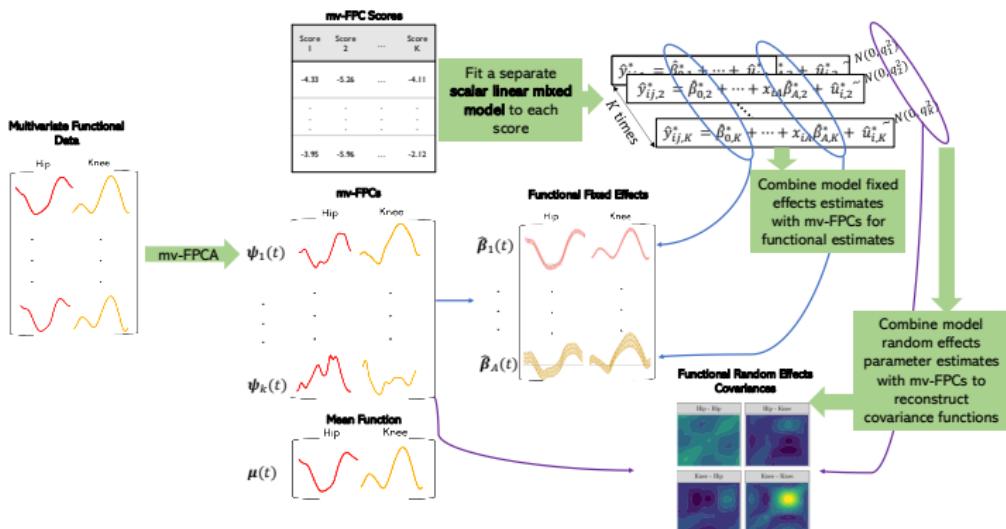
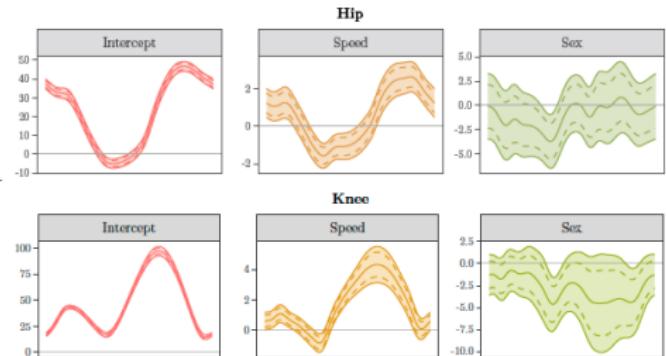


Figure: Flowchart of methodology.

Initial Approach (Gunning et al., 2024b)

Estimation and Inference

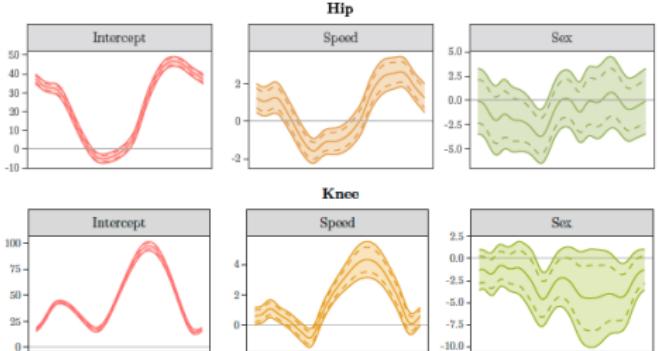
- Bayesian approach of Morris and Carroll (2006) uses MCMC for estimation and inference of functional parameters.
- We adapt resampling and simulation ideas from Crainiceanu, Staicu, et al. (2012) to implement in a frequentist setting using lme4.



Initial Approach (Gunning et al., 2024b)

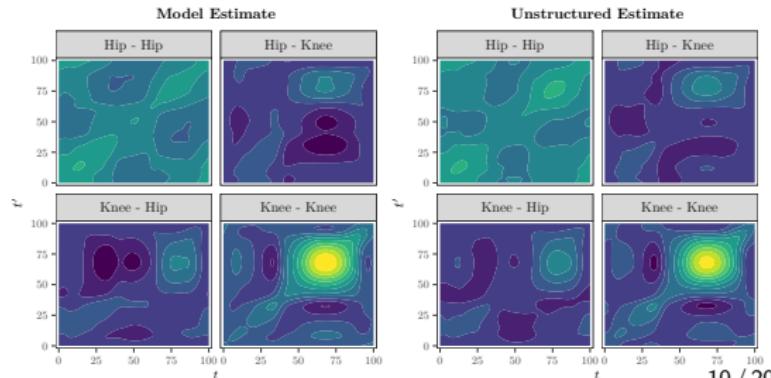
Estimation and Inference

- Bayesian approach of Morris and Carroll (2006) uses MCMC for estimation and inference of functional parameters.
- We adapt resampling and simulation ideas from Crainiceanu, Staicu, et al. (2012) to implement in a frequentist setting using lme4.



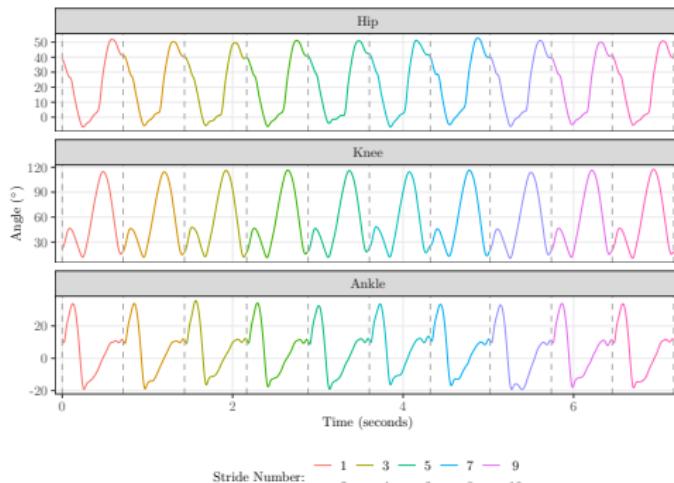
Covariance Analysis

- Modelling each basis coefficient separately makes potentially limiting assumptions on the form of the functional random effects and error covariances.
- We propose to check this by using an extension of multilevel FPCA (Di et al., 2009) to calculate unstructured estimates to compare with.



Summary and Limitations of Initial Approach

- Simple methodological approach combines existing methodologies, e.g., Morris and Carroll (2006), Di et al. (2009), and Crainiceanu, Staicu, et al. (2012).
- Scientific results consistent with and expand upon existing biomechanical literature (e.g., based on scalar values).
- **However, working with an average stride is potentially wasteful.** For just a one-minute run, this is a data reduction of almost 100 : 1.



Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Solution: Develop a structured functional model for all of the strides:

- Hip, knee and ankle → **multivariate**.
- Multiple strides measured bilaterally for each subject → **multilevel**.
- Strides have a natural time ordering → **longitudinal**.

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Our multivariate multilevel longitudinal functional model is

$$\mathbf{y}_{ijl}(t) = \boldsymbol{\beta}_0(t, T_{ijl}) + \sum_{a=1}^A x_{ija} \boldsymbol{\beta}_a(t) + \mathbf{u}_i(t, T_{ijl}) + \mathbf{v}_{ij}(t, T_{ijl}) + \boldsymbol{\varepsilon}_{ijl}(t),$$

where

- $\mathbf{y}_{ijl}(t) = \left(y_{ijl}^{(\text{hip})}(t), y_{ijl}^{(\text{knee})}(t), y_{ijl}^{(\text{ankle})}(t) \right)^\top$,
- $t \in [0, 100]\text{(\%)}$ ("functional time") and $T \in [0, 1]$ ("longitudinal time"), and
- $l = 1, \dots, n_{ij}, j \in \{\text{left, right}\}$, and $i = 1, \dots, N$.

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Our multivariate multilevel longitudinal functional model is

$$\mathbf{y}_{ijl}(t) = \underbrace{\beta_0(t, T_{ijl})}_{\text{Intercept}} + \sum_{a=1}^A x_{ija}\beta_a(t) + \mathbf{u}_i(t, T_{ijl}) + \mathbf{v}_{ij}(t, T_{ijl}) + \varepsilon_{ijl}(t),$$

where

- $\mathbf{y}_{ijl}(t) = \left(y_{ijl}^{(\text{hip})}(t), y_{ijl}^{(\text{knee})}(t), y_{ijl}^{(\text{ankle})}(t) \right)^\top$,
- $t \in [0, 100]\text{(%)}$ ("functional time") and $T \in [0, 1]$ ("longitudinal time"), and
- $l = 1, \dots, n_{ij}, j \in \{\text{left, right}\}$, and $i = 1, \dots, N$.

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Our multivariate multilevel longitudinal functional model is

$$\mathbf{y}_{ijl}(t) = \beta_0(t, T_{ijl}) + \underbrace{\sum_{a=1}^A x_{ija} \beta_a(t)}_{\text{Fixed Effects}} + \mathbf{u}_i(t, T_{ijl}) + \mathbf{v}_{ij}(t, T_{ijl}) + \varepsilon_{ijl}(t),$$

where

- $\mathbf{y}_{ijl}(t) = (y_{ijl}^{(\text{hip})}(t), y_{ijl}^{(\text{knee})}(t), y_{ijl}^{(\text{ankle})}(t))^{\top}$,
- $t \in [0, 100]\text{(\%)}$ (“functional time”) and $T \in [0, 1]$ (“longitudinal time”), and
- $l = 1, \dots, n_{ij}, j \in \{\text{left, right}\}$, and $i = 1, \dots, N$.

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Our multivariate multilevel longitudinal functional model is

$$\mathbf{y}_{ijl}(t) = \beta_0(t, T_{ijl}) + \sum_{a=1}^A x_{ija} \beta_a(t) + \underbrace{\mathbf{u}_i(t, T_{ijl})}_{\text{Subject}} + \mathbf{v}_{ij}(t, T_{ijl}) + \varepsilon_{ijl}(t),$$

where

- $\mathbf{y}_{ijl}(t) = \left(y_{ijl}^{(\text{hip})}(t), y_{ijl}^{(\text{knee})}(t), y_{ijl}^{(\text{ankle})}(t) \right)^\top$,
- $t \in [0, 100]\%$ ("functional time") and $T \in [0, 1]$ ("longitudinal time"), and
- $l = 1, \dots, n_{ij}, j \in \{\text{left, right}\}$, and $i = 1, \dots, N$.

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Our multivariate multilevel longitudinal functional model is

$$\mathbf{y}_{ijl}(t) = \beta_0(t, T_{ijl}) + \sum_{a=1}^A x_{ija} \beta_a(t) + \mathbf{u}_i(t, T_{ijl}) + \underbrace{\mathbf{v}_{ij}(t, T_{ijl})}_{\text{Subject \& Side}} + \varepsilon_{ijl}(t),$$

where

- $\mathbf{y}_{ijl}(t) = \left(y_{ijl}^{(\text{hip})}(t), y_{ijl}^{(\text{knee})}(t), y_{ijl}^{(\text{ankle})}(t) \right)^\top$,
- $t \in [0, 100]\%$ ("functional time") and $T \in [0, 1]$ ("longitudinal time"), and
- $l = 1, \dots, n_{ij}, j \in \{\text{left, right}\}$, and $i = 1, \dots, N$.

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Our multivariate multilevel longitudinal functional model is

$$\mathbf{y}_{ijl}(t) = \beta_0(t, T_{ijl}) + \sum_{a=1}^A x_{ija} \beta_a(t) + \mathbf{u}_i(t, T_{ijl}) + \mathbf{v}_{ij}(t, T_{ijl}) + \underbrace{\varepsilon_{ijl}(t)}_{\text{Error}},$$

where

- $\mathbf{y}_{ijl}(t) = \left(y_{ijl}^{(\text{hip})}(t), y_{ijl}^{(\text{knee})}(t), y_{ijl}^{(\text{ankle})}(t) \right)^\top$,
- $t \in [0, 100]\text{(%)}$ ("functional time") and $T \in [0, 1]$ ("longitudinal time"), and
- $l = 1, \dots, n_{ij}$, $j \in \{\text{left, right}\}$, and $i = 1, \dots, N$.

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Again, take a basis modelling approach¹ by representing

$$\mathbf{y}_{ijl}(t) = \sum_{k=1}^K y_{ijl,k}^* \psi_k(t),$$

where $y_{ijl,k}^*$ are scalar basis coefficients and $\psi_k(t)$ are multivariate basis functions in the *functional direction* (we choose mv-FPCs).

Reduces the problem to fitting a series of K mixed models of the form:

$$y_{ijl,k}^* = \beta_{0,k}^*(T_{ijl}) + \sum_{a=1}^A x_{ia} \beta_{a,k}^* + u_{i,k}^*(T_{ijl}) + v_{ij,k}^*(T_{ijl}) + \varepsilon_{ijl,k}^*,$$

which is a standard multilevel functional model (Di et al., 2009) in the *longitudinal direction*.

¹ Similar idea proposed in univariate two-level case by Park and Staicu (2015).

Multivariate Multilevel Longitudinal Functional Model (Gunning et al., 2024a)

Again, we use basis expansions, this time in the longitudinal direction:

$$\beta_{0,k}^*(T) = \sum_{d=1}^D \beta_{0,k,d}^* \xi_d(T), \quad u_{i,k}^*(T) = \sum_{d=1}^D u_{i,k,d}^* \xi_d(T) \quad \text{and} \quad v_{ij,k}^*(T) = \sum_{d=1}^D v_{ij,k,d}^* \xi_d(T).$$

Question: Which basis functions should we use for $u_{i,k}^*(T)$ and $v_{ij,k}^*(T)$?

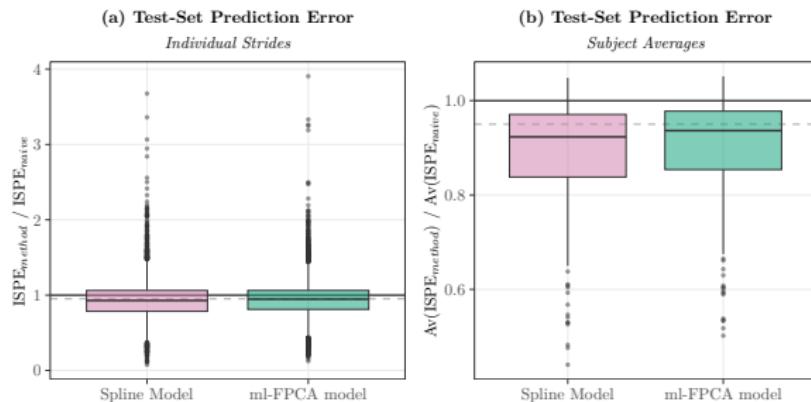


Figure: Representation using pre-specified spline (pink) and empirically-determined ml-FPCA basis (Cui et al., 2022, green) basis functions gives similar results.

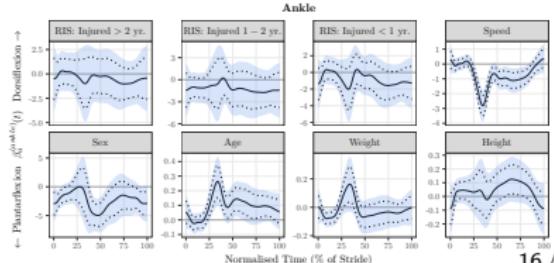
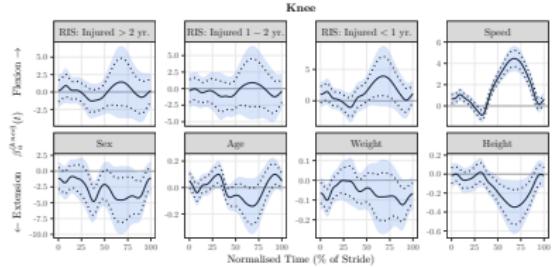
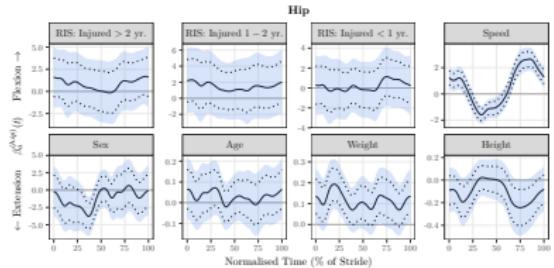
Results: Fixed Effects

- Fixed effects estimates

$$\hat{\beta}_a(t) = \sum_{k=1}^K \hat{\beta}_{a,k}^* \psi_k(t).$$

- Simultaneous bands account for multiple comparisons across t and the hip, knee and ankle. Obtained using bootstrap and simulation approaches^a.
- Scientific results consistent with existing biomechanical knowledge.

^aSee Faraway (1997), Ruppert, Wand, and Carroll (2003), Crainiceanu, Staicu, et al. (2012), and Cui et al. (2022).



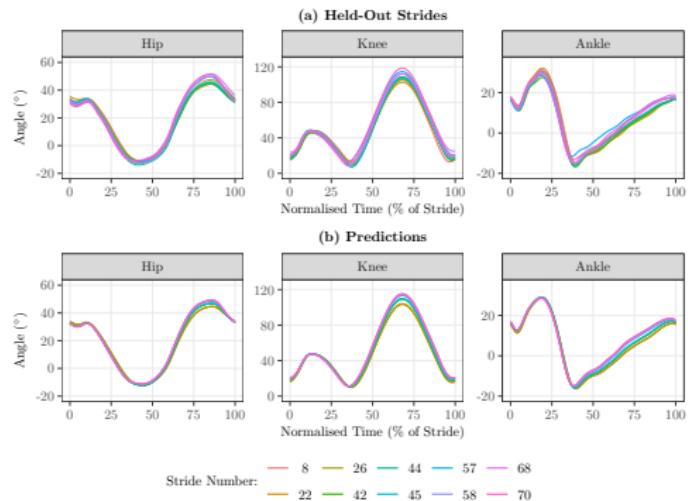
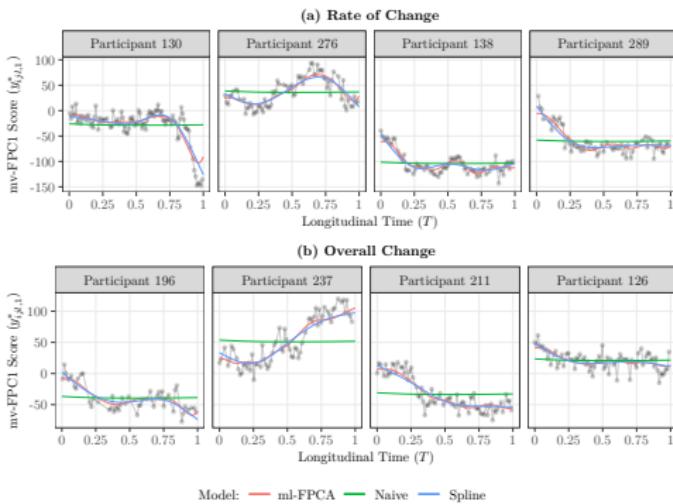
Results: Random Effects

Rates of change of $\hat{\mathbf{u}}_i(t, T)$ and $\hat{\mathbf{v}}_{ij}(t, T)$ w.r.t. T characterise changes over the course of the treadmill run.



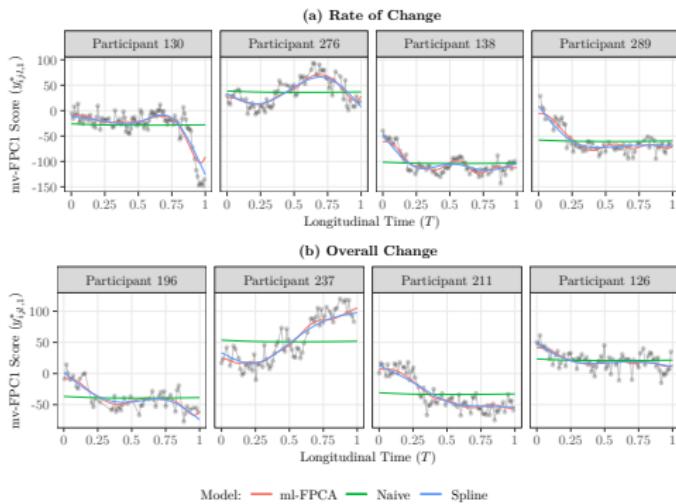
Results: Random Effects

Rates of change of $\hat{\mathbf{u}}_i(t, T)$ and $\hat{\mathbf{v}}_{ij}(t, T)$ w.r.t. T characterise changes over the course of the treadmill run.

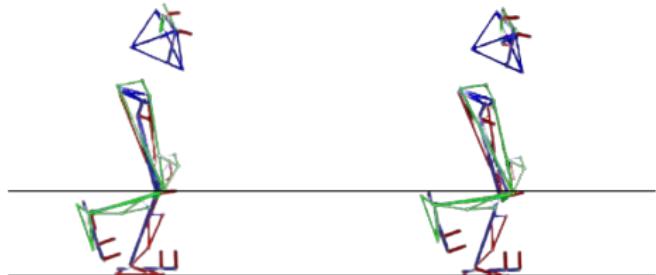


Results: Random Effects

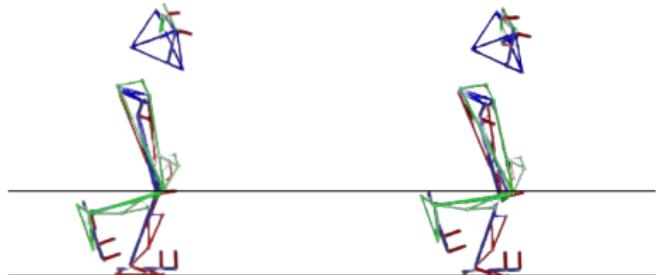
Rates of change of $\hat{\mathbf{u}}_i(t, T)$ and $\hat{\mathbf{v}}_{ij}(t, T)$ w.r.t. T characterise changes over the course of the treadmill run.



(c) Start of Treadmill Run



(d) End of Treadmill Run



Summary

- Simple and flexible approach for modelling streams of smooth multivariate functional data that arise in biomechanics.
- Characterise population-level fixed effects and intra-individual longitudinal changes during the treadmill run – practically meaningful insights.
- Potential to use the model in different contexts and extend it to more complex settings.

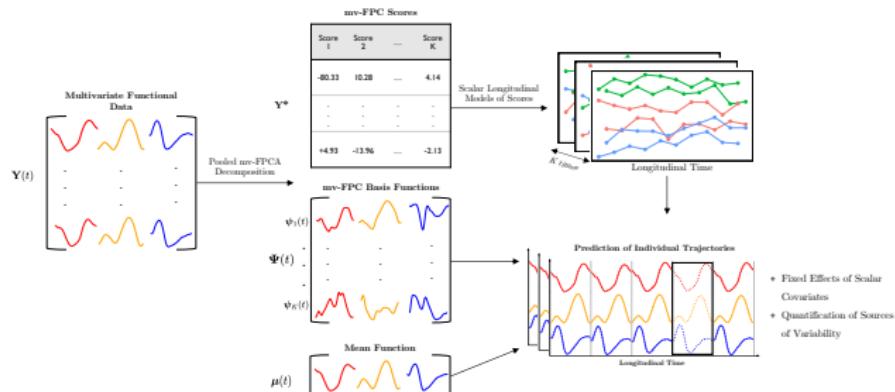


Figure: Summary of the modelling approach.

Thank you!

More Information

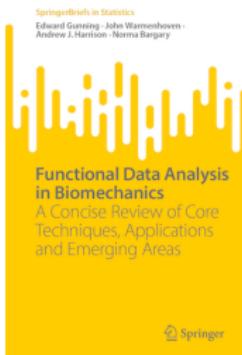
Literature:

- E Gunning et al. (2024b). *Analysing Kinematic Data from Recreational Runners using Functional Data Analysis*. arXiv:2408.08200 [stat]
- E Gunning et al. (2024a). “A Multivariate Multilevel Longitudinal Functional Model for Repeatedly Observed Human Movement Data”. In: arXiv:2408.08481 [stat]

Email: edward.gunning@pennmedicine.upenn.edu

Website: <https://edwardgunning.github.io/>

Upcoming Book:



References I

-  Crainiceanu, CM, J Goldsmith, A Leroux, and E Cui (2024). *Functional data analysis with R*. Chapman & Hall/CRC.
-  Crainiceanu, CM, AM Staicu, S Ray, and N Punjabi (Nov. 2012). "Bootstrap-based inference on the difference in the means of two correlated functional processes". In: *Statistics in Medicine* 31.26, pp. 3223–3240.
-  Cui, E, A Leroux, E Smirnova, and CM Crainiceanu (Jan. 2022). "Fast Univariate Inference for Longitudinal Functional Models". In: *Journal of Computational and Graphical Statistics* 31.1, pp. 219–230.
-  Di, CZ, CM Crainiceanu, BS Caffo, and NM Punjabi (Mar. 2009). "Multilevel functional principal component analysis". In: *The Annals of Applied Statistics* 3.1, pp. 458–488.
-  Faraway, JJ (1997). "Regression Analysis for a Functional Response". In: *Technometrics* 39.3, pp. 254–261.
-  Gunning, E, S Golovkine, AJ Simpkin, A Burke, S Dillon, S Gore, KA Moran, S O'Connor, Whyte, and N Bargary (2024a). "A Multivariate Multilevel Longitudinal Functional Model for Repeatedly Observed Human Movement Data". In.
-  Gunning, E, S Golovkine, AJ Simpkin, A Burke, S Dillon, S Gore, KA Moran, S O'Connor, Whyte, and N Bargary (2024b). *Analysing Kinematic Data from Recreational Runners using Functional Data Analysis*.

References II

-  Horsak, B, D Slijepcevic, AM Raberger, C Schwab, M Worisch, and M Zeppelzauer (May 2020). "GaitRec, a large-scale ground reaction force dataset of healthy and impaired gait". In: *Scientific Data* 7.1, p. 143.
-  Morris, JS and RJ Carroll (Apr. 2006). "Wavelet-based functional mixed models". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 68.2, pp. 179–199.
-  Park, SY and AM Staicu (2015). "Longitudinal functional data analysis". In: *Stat* 4.1, pp. 212–226.
-  Rice, JA and BW Silverman (1991). "Estimating the Mean and Covariance Structure Nonparametrically When the Data are Curves". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 53.1, pp. 233–243.
-  Ruppert, D, MP Wand, and RJ Carroll (2003). *Semiparametric Regression*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press.
-  Winter, DA (1979). *Biomechanics of Human Movement*. Wiley.