



Functional Data Analyses of Gait Data Measured Using In-Shoe Sensors

Jihui Lee¹ · Gen Li² · William F. Christensen³ · Gavin Collins³ · Matthew Seeley⁴ · Anton E. Bowden⁵ · David T. Fullwood⁵ · Jeff Goldsmith²

Received: 1 December 2017 / Revised: 3 November 2018 / Accepted: 1 December 2018
© International Chinese Statistical Association 2018

Abstract

In studies of gait, continuous measurement of force exerted by the ground on a body, or ground reaction force (GRF), provides valuable insights into biomechanics, locomotion, and the possible presence of pathology. However, gold-standard measurement of GRF requires a costly in-lab observation obtained with sophisticated equipment and computer systems. Recently, in-shoe sensors have been pursued as a relatively inexpensive alternative to in-lab measurement. In this study, we explore the properties of continuous in-shoe sensor recordings using a functional data analysis approach. Our case study is based on measurements of three healthy subjects, with more than 300 stances (defined as the period between the foot striking and lifting from the ground) per subject. The sensor data show both phase and amplitude variabilities; we separate these sources via curve registration. We examine the correlation of phase shifts across sensors within a stance to evaluate the pattern of phase variability shared across sensors. Using the registered curves, we explore possible associations between in-shoe sensor recordings and GRF measurements to evaluate the in-shoe sensor recordings as a possible surrogate for in-lab GRF measurements.

Keywords Gait analysis · Ground reaction force · Functional data · Curve registration

✉ Jihui Lee
jil2043@med.cornell.edu

¹ Department of Healthcare Policy and Research, Weill Cornell Medicine, New York, NY 10065, USA

² Department of Biostatistics, Columbia University, New York, NY 10032, USA

³ Department of Statistics, Brigham Young University, Provo, UT 84602, USA

⁴ Department of Exercise Sciences, Brigham Young University, Provo, UT 84602, USA

⁵ Department of Mechanical Engineering, Brigham Young University, Provo, UT 84602, USA

1 Introduction

1.1 Motivation

In order for humans to walk or run, the ground must exert a force on the bottom of the shoe, or foot, that results in acceleration of the human center of mass. This force is called the ground reaction force (GRF). The resultant GRF vector is often resolved into three orthogonal components: a vertical component and two horizontal components often called anterior–posterior (fore–aft) and medial–lateral (side-to-side). Various characteristics of the vertical component of walking and running GRF are measured, because they are associated with common musculoskeletal impairments and disease. For example, characteristics of the vertical GRF during walking and running are associated with knee joint health, including knee osteoarthritis onset and progression [13,32,37]. Vertical GRF is also measured in order to discern progression of a musculoskeletal disease, like knee osteoarthritis, or the effectiveness of a clinical intervention designed to slow disease progression [24].

Currently, the accurate measurement of walking and running GRF requires a subject to walk across a force platform that is either embedded in a laboratory floor or moveable surface (e.g., a force-sensing treadmill). Commercial force platforms and force-sensing treadmills are expensive; further, such instruments are restricted to laboratory environments and require extensive human resources (i.e., expertise) to manage. These challenges prohibit some researchers and most clinicians from making accurate measures of walking and running GRF. In addition, real-world (i.e., out of the laboratory) measures of GRF are currently difficult or impossible to obtain.

These issues have motivated a development of mobile force sensors that can be used to measure GRF outside of a traditional motional analysis laboratory. Novel piezo-responsive foam sensors placed in athletic shoes have been recently developed to accurately estimate walking 3D GRF outside of the laboratory [28]. The electromechanical behaviors of these sensors have been validated for use in various large-strain applications [2,14]. The strain-induced voltage is measured by attaching a conductive material embedded in the foam to a voltage-sensing system, which correlates to the force of impact [21]. By embedding the foam into a shoe sole, it has been shown that the voltage response generated during gait accurately correlates to 3D GRF [28].

There are several potential advantages of using in-shoe sensor as a surrogate of GRF measurement. While the GRF is only recordable in controlled indoor laboratories, in-shoe sensors can be simply incorporated into the sole of a pair of shoes and deployed anywhere to analyze one's gait in various settings. Furthermore, Seliktar et al [33] pointed out a chance of distortion in one's gait pattern when asked to walk on the force plates. Distortions that interfere with detecting a true gait pattern are less likely to happen when using the in-shoe sensors placed on a regular pair of shoes. Thus, it is important to explore the properties of in-shoe sensor data to understand if in-shoe recordings can be a viable alternative of GRF measurement.

Figure 1 shows recorded values from both in-shoe sensors and gold-standard GRF measurement obtained via a force-sensing treadmill, during the ground contact phase of running (i.e., between heel strike and toe-off, called a “stance”) for five consecutive stances in a healthy individual. Recordings for both in-shoe sensor and vertical GRF

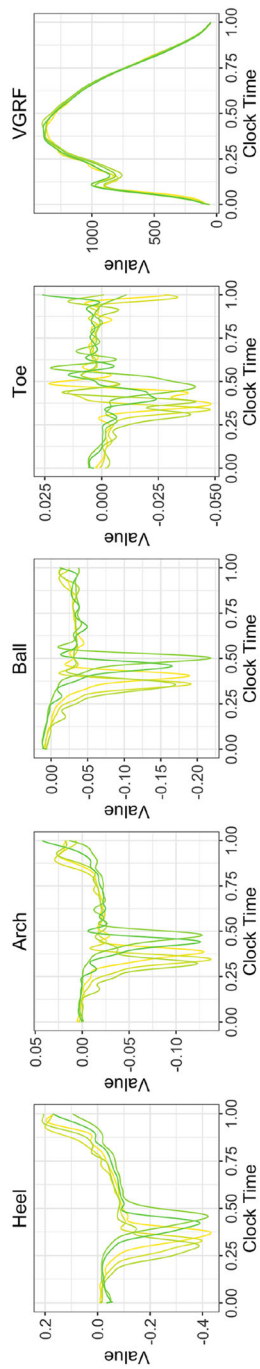


Fig. 1 In-shoe sensor and GRF data: In-shoe sensor recordings and vertical GRF (VGRF) of a healthy subject's 5 stances. Data from in-shoe sensors at four locations are shown in the first four panels. The fifth panel shows the measurement of vertical GRF. Each curve represents a stance

(VGRF) measurements are shown over time for each stance. In the first four panels, the x -axis corresponds to the standardized time frame and the y -axis corresponds to in-shoe sensor readings. The sensor is measuring a triboelectric effect as the embedded nanofillers rub against the base polymer in the foam. This effect may be amplified or concentrated by the voids in the foam which can function as a short duration capacitor that stores voltage and then discharges it. Larger forces produce larger displacements in the foam, which correspond to higher magnitude electrical response; in turn, the responses create more negative values in the sensor readings. The sensor has its lowest (most negative) values when reacting to the largest forces.

The values observed from the sensors reflect the subject's gait pattern. According to Fig. 1, for example, each stance in in-shoe sensor data contains a large dip, after which the value increases and flattens until the completion of the stance. Recordings within a sensor share a common structure that is misaligned across stances: the exact timing of major features depends on stance. Furthermore, the magnitude of the common feature differs between stances. These observations relate to *phase* and *amplitude* variabilities, respectively; phase variability relates to shifts in time, while amplitude variability relates to the change in the magnitude of measurements.

The VGRF measurement shown in the fifth panel of Fig. 1 is more stable than in-shoe sensor recordings, and does not show significant time shifts across stances. Indeed, it has been documented that healthy runners are very consistent in their stances [1,15]. Thus, the phase variability in in-shoe sensor data across stances in Fig. 1 is not expected, and the misalignment arises in the recording process rather than reflecting actual variability across stances. Removing phase variability from the in-shoe sensor data is a necessary step if these measurements are going to be used as surrogate measures of VGRF. In addition, analyzing the sensor data without proper understanding of phase variability may lead us to draw misleading conclusions regarding the amplitude variability of common stance features [29]. Our goal, then, is to explore the elimination of phase variability of in-shoe sensor data without altering the values taken by the curves, so that differences in amplitude variability can be evaluated as alternative measures of the VGRF in studies of walking pattern and gait.

Temporal realignment of curves is referred to as *curve registration* in the functional data analysis literature. Specifically, curve registration shifts, stretches, and compresses the observations in time so that major features are aligned across curves. In this process, clock (originally observed) time is converted into the system (common across curves) time via *time warping functions*. Let t^* and t denote clock and system time, respectively. A warping function $h : [0, 1] \rightarrow [0, 1]$ represents the functional relationship between clock time and system time through $t^* = h(t)$. The warping functions are monotone increasing with $h(0) = 0$ and $h(1) = 1$. The principal challenge in registration, then, is the estimation of warping functions. Curve registration is often necessary before applying additional statistical methods to smooth curves, and warping functions themselves can contain useful information about observed curves. In this article, we use a curve registration to understand the phase and amplitude variabilities arising in the in-shoe sensor data.

We investigate the variability present in the in-shoe sensor data recorded for three representative healthy subjects, with more than 300 stances per subject. We emphasize the importance of understanding the source of variability and the utility of adequately

addressing phase variability. Based on the hypothesis that time shifts across sensors within the same stance may be similar, we examine the correlation of warping functions using a permutation test. We further examine the amplitude variability of the in-shoe sensor recordings after registration, particularly with relation to GRF, by employing function-on-function regression models. All analyses are conducted separately for each subject due to unique running patterns.

1.2 Literature Review

In studies of gait, curve registration is valuable but underutilized in reducing intra-subject variability. In studies of joint mechanical power data, Sadeghi et al. [29,30] emphasized the importance of curve registration as a preprocessing step. Both studies obtained joint mechanical power data for the right lower limb of healthy subjects using a 3D video-based system. Sadeghi et al. [29] applied a straightforward registration method to align salient features of the observed curves before comparing key power bursts; Sadeghi et al. [30] implemented more flexible registration method. These studies found that realigning the observed curves reduces the temporal variability induced by external sources and instrumental issue and facilitates the focus on sources of variability that reflect meaningful biomechanical features. More recently, curve registration was used in gait studies on healthy subjects [3,12] and for comparing healthy individuals to stroke patients [38].

In functional data analysis, curve registration was introduced to identify a shared structural pattern in a sample of curves and to understand individual realizations of the shared pattern [16,18,25]. There have been a variety of approaches to and applications of curve registration. For an in-depth history of curve registration, see Marron et al. [19,20].

A simple approach to curve registration, referred to as landmark registration, locates important features of the observed curves by hand or using an automated process and realigns them using piecewise linear warping functions [7,16]; landmark registration was implemented in Sadeghi et al. [29]. Although landmark registration is simple and straightforward to implement, it can be difficult and time consuming to determine landmark locations, and the performance of this approach may be poor in the area away from the landmarks.

More flexible methods for curve registration have been introduced. In principle, they estimate nonlinear, monotone warping functions that map the system (warped) time into observed clock time. [34] proposes a method that does not require landmarks by considering uniform shifts in time to realign the observed curves. Ramsay [26] and Ramsay and Li [25] propose an iterative algorithm with two steps: first, estimate the cross-sectional mean of the registered curves using current warping function estimates, and second, update warping function estimates to minimize distance to the shared mean. This approach was used by Sadeghi et al. [30]. Building on this framework, functional principal component analysis (FPCA) has been widely used to model the common structure shared by registered curves [5,6,17].

Srivastava et al. [36] suggests a metric-based framework for registering elastic curves. The method, like other iterative algorithms for curve registration, alternates

between two steps until convergence. In this approach, the Fisher–Rao Riemannian metric and square-root velocity function (SRVF) are used to quantify the distance between two curves. The Fisher–Rao metric is a widely used tool to compare the shape of curves; for example, the work of Peter and Rangarajan [23] represents landmark-based shapes using a Gaussian mixture model and computes geodesic distances between two shapes using the parametric Fisher–Rao metric. Srivastava et al. [35] use the extension of nonparametric Fisher–Rao metric directly on the space of functions, and its parametrization invariance is used to separate the phase and amplitude variabilities. The SRVF maps the Fisher–Rao metric to Euclidean space, which enables the comparison of distances in \mathbb{L}^2 space; it greatly simplifies computation in analyzing shapes. This method is extended to generative models in Tucker et al. [39] and the analysis of shape of elastic curves in Euclidean space in Srivastava et al. [35]; it has also been applied to proteomics data and spike train data [40,41].

The remainder of this paper is organized as follows. Section 2 introduces the gait data and describes the questions of interest. Section 3 illustrates the use of curve registration, constructs a permutation test framework for evaluating correlation within stance across sensors, describes a metric that quantifies the amplitude variability before and after curve registration, and introduces functional regression methods to model the association between in-shoe sensor recordings and VGRF curves. We present the results of curve registration, permutation testing, quantification of amplitude variability, and regression model fitting in Sect. 4. We conclude with a discussion of our findings and open areas for future research in Sect. 5.

2 Data

Data were collected in a biomechanics laboratory at Brigham Young University. Subjects were instrumented with a Cosmed K4b2 portable metabolic analyzer (Cosmed K4b2, Cosmed, Rome, Italy) and standardized athletic shoes instrumented with the nanocomposite piezo-responsive foam (NCPF) sensors, accompanying electrical components, and an accelerometer attached to the dorsal aspect of the shoe (Fig. 2). Next, to be able to account for any NCPF sensor drift, subjects completed a 15-min warm-up run at 2.68 m/s. After this warm-up, subjects completed five different trials in a randomized order. Each trial consisted of 4 min of walking or running, at one of the following speeds: 1.34, 2.23, 2.68, 3.13, or 3.58 m/s. A 1-min walk (1.34 m/s) was completed before and after every trial, to be able to characterize any drift in the NCPF sensors. Voltages, recorded via the NCPF sensors and microcontroller (1000 Hz), energy expenditure, recorded via the Cosmed (breath by breath), and accelerations, measured via the shoe accelerometer (16 Hz), were measured throughout the entire collection period (the warm-up period and all five trials), which lasted approximately 50 min.

The data set analyzed here consists of measurements obtained from three healthy female subjects who wear size 8 shoes. The mass and height of the subjects are (52, 60, 60) kg and (161, 165, 169) cm, respectively. Each subject was required to 1) be between the ages of 18 and 30; 2) have no history of lower-extremity injury within the past 6 months; 3) have no history of lower-extremity surgery in their lifetime; 4) be able to



Fig. 2 In-shoe sensor: (Left) NCPF sensor with wires embedded to measure the voltage response during impacts. (Right) Shoe equipped with NCPF foam sensors, microcontroller, and battery. The sensors were embedded under the insole at the heel, arch, ball, and toe

walk and run without pain; and 5) be able to comfortably run consistently for at least 5 continuous km. Each subject ran at 6mph (2.68 m/s) for 4 min and completed 362, 321, and 308 stances, respectively. Consistent with much of the scientific literature in this area, we define “stance” as a period from heel strike (point at which vertical GRF rises above 50 N) to toe-off (point at which VGRF drops below 50 N). From each stance, information such as VGRF and in-shoe sensor measurements at four locations (heel, arch, ball, and toe) has been collected. Completing a stance takes different amount of time; once a stance is extracted, the curves are linearly interpolated to a common domain with 200 discrete time points. This preprocessing formats the data in a way that is suitable for our subsequent analysis, but does not alter stances beyond a simple stretching to a common domain.

The resulting data are shown in Fig. 3. For each in-shoe sensor and subject, a common feature exists across stances but is misaligned in time. In addition to the observed phase variability, there is amplitude variability across individuals and stances. The distinctive patterns in the in-shoe sensor and VGRF data across subjects are due to different running styles. We classify the Subject 1 as a mid-foot striker; this runner is more likely to land near the middle of the foot, rather than the heel, leading to low amplitude variability in the heel sensor and high amplitude variability in the arch and ball sensors. This running pattern also results in the absence of an impact transient, or initial local maximum, in the VGRF. In contrast, the remaining subjects are heel strikers who show greater amplitude variability in the heel sensor, lower amplitude variability in the arch, ball, and toe sensors, and have, to differing degrees, the impact transient in the VGRF.

These distinctive running patterns may have long term implications for biomechanics and the health of joints involved in running [27]. The initial peak or impact transient is an important element of GRF data, as is the slope, or load rate, of the impact transient. A steeper impact transient load rate is thought to be correlated with certain musculoskeletal injury [4], and the absence of this peak for the Subject 1 might suggest that she is at lower risk for certain types of injuries [27].

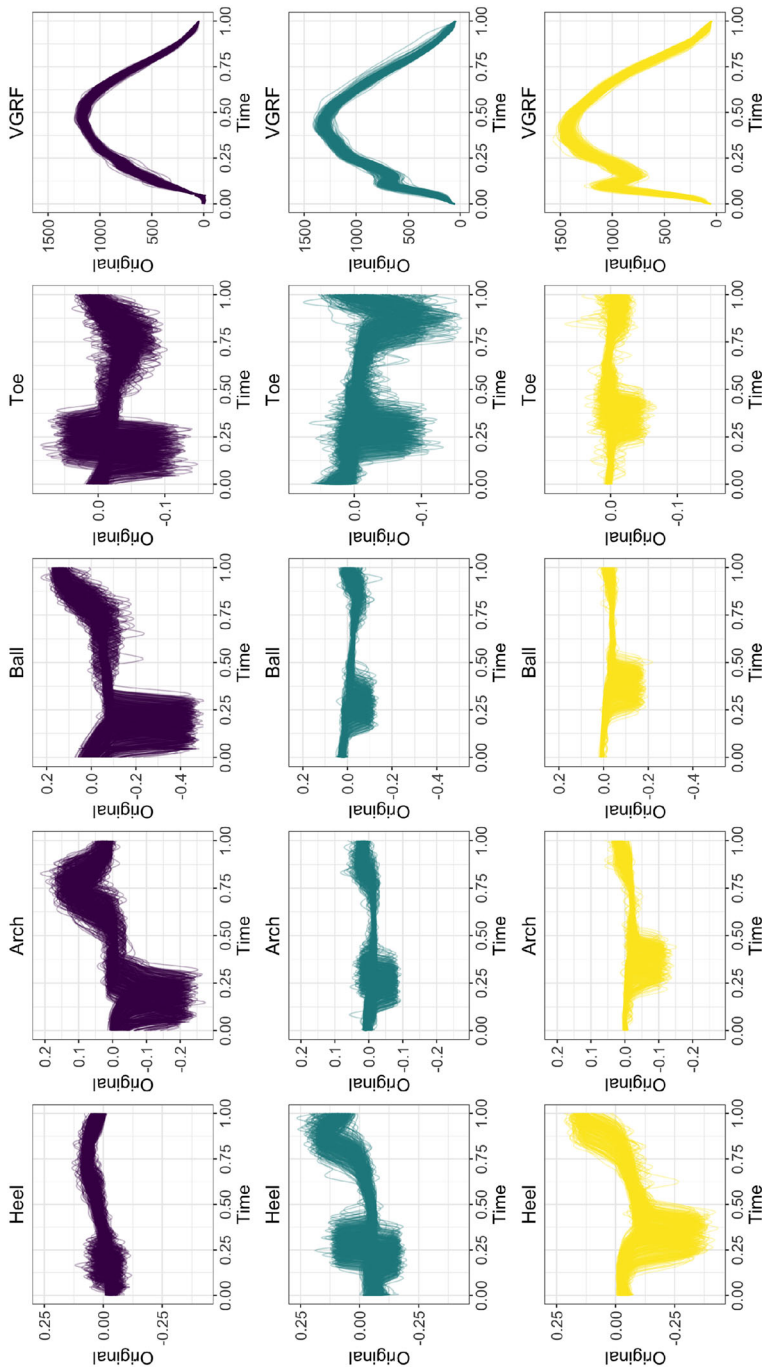


Fig. 3 Original measurements: Observed data for three runners are shown in rows; data from in-shoe sensors at various locations are shown in the first four columns and the direct measure of VGRF is shown in the fifth column. Curves are color coded per subject, and each curve represents a stance

As noted in Sect. 1.1, time shifts in in-shoe sensor data are not expected for healthy subjects. The stance-level force measurements are expected to be relatively consistent, and shifts away from a common structure shared by every curve are an issue of measurement rather than true phase variability; this is emphasized by the consistency of concurrently measured VGRFs during data collection. There are several possible reasons for the presence of phase variability in in-shoe sensor data, including the need for a warming-up duration before sensors output consistent voltages corresponding to consistent impacts and the possibility of misalignment in the sensor recording mechanisms (e.g., inconsistent identification of heel strike). It is also unclear whether the phase variation is similar across sensors within the same stance; that is, whether recordings made by the difference sensors on the same stance are time-shifted in a similar way. For new and developing technologies, such as NCPF sensors, preprocessing the observed data is valuable for understanding how the technology can be used and improved.

3 Methods

Unlike classic statistical methods, where one observation consists of a single scalar value, functional data analysis considers the basic unit of observation to be a smooth curve. The gait data illustrated in Fig. 3 are an example of functional data, since stances recorded by each sensor within each subject are the observations of interest. We introduce notation to denote curves for each subject and conduct curve registration to each sensor and subject separately. Let $y_i^s(t^*)$, $t^* \in [0, 1]$, denote the observed curve for the i -th stance for sensor s , where $i = 1, 2, \dots, I$, and $s = 1$ (Heel), 2 (Arch), 3 (Ball), 4 (Toe). Here, the clock time t^* is curve specific; the goal of curve registration is to estimate warping functions $t^* = h_i^s(t)$ that map the shared system time t to curve-specific clock time t^* .

3.1 Curve Registration

We implement the curve registration proposed by Srivastava et al. [36]. Like other approaches, this method is based on an iterative algorithm that alternates two steps until convergence. In the first step, the mean of registered curves using current warping functions is estimated; this mean is referred to as a template. In the second step, the warping functions are updated to minimize the distance between the curves and the current template, using the square-root velocity function (SRVF) to calculate the distance between two curves. The SRVF represents the Fisher–Rao Riemannian metric, which is a widely used tool to compare the shape of curves in \mathbb{L}^2 , and alleviates the computational complexity of the algorithm.

More concretely, for an absolutely continuous function $y_i^s(t^*)$, $t \in [0, 1]$ with its derivative $y_i^{s'}(t^*)$, define the SRVF $q_i^s : [0, 1] \rightarrow \mathbb{R}$ as

$$q_i^s(t^*) = \text{sign}(y_i^{s'}(t^*))\sqrt{|y_i^{s'}(t^*)|}.$$

With warping function $t^* = h(t)$, the SRVF of the registered curve $y_i^s(h^{-1}(t^*))$ can be written as $q_{i,h}^s(t^*) = q(h^{-1}(t^*))\sqrt{(h^{-1})'(t^*)}$. For any curves $y_i^s(t^*)$ and $y_{i'}^s(t^*)$, $i, i' \in \{1, 2, \dots, I\}$ and a warping function $h(t)$, the distance between SRVFs of registered curves is the same as that between SRVFs of unregistered curves; i.e., $\|q_{i,h}^s(t^*) - q_{i',h}^s(t^*)\| = \|q_i^s(t^*) - q_{i'}^s(t^*)\|$. This distance function in \mathbb{L}^2 is invariant to time warping, and it can be used to define a mean template.

In the first iteration, the template $\mu(t^*)$ is taken to be the observed $q_i^s(t^*)$ that is closest to the sample mean $\frac{1}{I} \sum_{j=1}^I q_j^s(t^*)$. Then, in the second step of the iterative algorithm, for each observed curve $y_i^s(t^*)$, warping function estimate is updated as $\arg\min_h \|\mu(t^*) - q_{i,h}^s(t^*)\|$. That is, the warping function estimate minimizes the distance of SRVF between the template and registered curves. In subsequent iterations, the template is updated as a mean of SRVFs q_{i,h_i}^s with current warping function estimates. The algorithm iterates the two steps of calibrating the template $\mu(t^*)$ and updating the warping functions until convergence.

Although we use the SRVFs for curve registration and achieve computational efficiency, the q functions are difficult to interpret in themselves. We instead focus on the aligned sensor measurements for subsequent analyses so that we make an interpretable inference regarding the gait measurements.

3.2 Sources of Variability

We conduct exploratory analyses on registered curves to evaluate the utility of curve registration. Following Tucker et al. [39] and Kneip and Ramsay [17], we define the amplitude variability of observed curves $\{y_i^s(t^*), i = 1, 2, \dots, I, t^* \in [0, 1]\}$ for sensor s as

$$V^s = \frac{1}{I-1} \int_0^1 \sum_{i=1}^I \left(y_i^s(t^*) - \frac{1}{I} \sum_{j=1}^I y_j^s(t^*) \right)^2 dt^*.$$

Similarly, we define the amplitude variability after curve registration with time transformation $t^* = h_i^s(t)$ and denote the resulting quantity as V_h^s . This definition quantifies the amplitude variability as the mean integrated sum of squared differences between curves in a sample and their mean. We expect that the amplitude variability before curve registration, V^s , may contain variability attributable to misalignment in time. Therefore, we compare the amplitude variability before and after curve registration to describe the utility of curve registration.

In addition, we conduct functional principal component analysis (FPCA) to understand the patterns observed in amplitude variability after curve registration [8, 42]. FPCA allows a parsimonious representation of registered curves by decomposing the observed functions into a mean, scores, and shared functional principal components:

$$y_i^s(t) = \mu^s(t) + \sum_{k=1}^K c_{ik}^s \phi_k^s(t) + \epsilon_i^s(t) \quad (1)$$

The representation in (1) is based on the Karhunen-Loève representation of the $y_i^s(t)$ in which $\mu(t)$ is the population mean, the ϕ_k^s are population level basis functions obtained through an eigendecomposition of the covariance $\text{Cov}(y_i^s(t), y_i^s(t'))$ with corresponding eigenvalues λ_k^s such that $\lambda_1^s \geq \lambda_2^s \geq \dots$, the subject-specific scores c_{ik}^s with $E(c_{ik}^s) = 0$ and $\text{Var}(c_{ik}^s) = \lambda_k^s$ are uncorrelated random variables, and the $\epsilon_i^s(t)$ are white noise errors.

In our analyses, FPCA is conducted for each sensor separately to understand the directions of variation for each sensor within subjects.

3.3 Phase Variability Across Sensors

We are interested in exploring the similarity of phase shifts across sensors within stance. Similarity in time shifts across sensors within stance may reasonably be expected given the biomechanical process underlying these data; dissimilarity would suggest that phase shifts are not consistent across sensors in the same stance.

To assess similarity, we use a permutation test based on a functional analog of Pearson's correlation that considers the relationship between two warping functions after subtracting the identity function (i.e., $h(t) = t$). Namely, we define the correlation between two warping functions from sensor s and s' within the i -th stance as

$$\text{Corr}_i = \frac{\int_0^1 \{h_i^s(t) - t\} \{h_i^{s'}(t) - t\} dt - \int_0^1 \{h_i^s(t) - t\} dt \int_0^1 \{h_i^{s'}(t) - t\} dt}{\sqrt{[\int_0^1 \{h_i^s(t) - t\}^2 dt - [\int_0^1 \{h_i^s(t) - t\} dt]^2][\int_0^1 \{h_i^{s'}(t) - t\}^2 dt - [\int_0^1 \{h_i^{s'}(t) - t\} dt]^2}}}.$$

Terms in the numerator and denominator are based on the expectations and variances that appear in the usual definition of Pearson's correlation. Let $\hat{h}_i^s(t)$ denote the estimated warping function of the i -th stance for sensor s , where $i = 1, 2, \dots, I$ and $s = 1, 2, 3, 4$. The test statistic used to evaluate the similarity across sensors is the mean of correlations within stance Corr_i ; that is, we compute the correlation between two sensors within each stance, and then average across stances.

To determine statistical significance, we randomly permute the stance label of second sensor, recompute within-stance correlations, and average across stances. Test statistic for k -th permuted sample, denoted as T_k , is defined as a mean of correlations in the permuted sample. We repeat the permutation process K times to obtain a null distribution for our test statistic. The p value for permutation test is defined to be

$$p\text{value} = 2 \times \frac{\sum_{k=1}^K \mathbb{1}\{T < T_k\}}{K}.$$

3.4 Association Between In-Shoe Sensors and GRF

We are interested in the relationship between VGRF curves and the aligned in-shoe sensor recordings because any meaningful relationship between VGRF curves and in-shoe sensor recordings may support the utilization of inexpensive in-shoe sensor measurement as a surrogate for GRF measurement.

To understand possible associations, we apply function-on-function regression models using VGRF as an outcome and in-shoe sensor recording as a predictor. Let $\text{VGRF}_i(t)$ be the measure of VGRF from the i -th stance and let $y_i^s(u)$, $u \in [0, 1]$ denote the registered curve of the i -th stance for sensor s . We fit a model with VGRF as a response and the sensor s recording as a predictor:

$$\text{VGRF}_i(t) = \beta_0(t) + \int_0^1 y_i^s(u) \beta^s(u, t) du + \varepsilon_i(t). \quad (2)$$

We fit separate models for each subject using the tensor-product spline approach described in Scheipl et al. [31] and implemented in the `ppfr` function in the `refund` package in R [9]. The bivariate coefficient surface $\beta^s(u, t)$ is smooth over both u and t and relates the predictor measured over u to the response measured t , respectively. In this model, fitted values at time t are obtained fixing t and viewing $\beta^s(u, t)$ as as a univariate coefficient function over u , which is multiplied by the predictor and integrated over u . The bivariate smoothness of $\beta^s(u, t)$ allows the effect of predictor functions to vary over the domain of the response.

4 Results

4.1 Curve Registration

Curve registration is applied separately to each individual and each measurement. Figure 4 shows in-shoe sensor curves after registration, and can be directly compared to Fig. 3. The fifth panel in Fig. 4 shows *unregistered* VGRF measurement; although the VGRF is not realigned, it is presented in the figure to aid visual comparisons.

Curve registration achieves a reduction in phase variability; curves are better aligned, and amplitude variability is more easily understood. For example, the spikes in the early phase of the heel sensor for each subject were largely obscured before registration. In the middle of Subject 1's arch and Subject 2's toe sensors, there is some variability after registration.

Inverse warping functions for such realignments of Subject 1 are presented in Fig. 5, with observed clock times on the x -axis and aligned system times on the y -axis. Roughly speaking, inverse warping functions above the identity line shift an early peak in the observed time and map it to a later system time, while inverse warping functions below the identity warping shift later peaks to an earlier system time. In the arch sensor, the range of system time when clock time is 0.5 is wide compared to the other sensors, which is reasonable given the wide time shifts across stances in the originally observed curves. Visual inspection of inverse warping functions suggests that adjacent sensors, such as the heel and arch or the ball and toe, are similar for this subject, which may imply some association in time shifts for these sensors. The similarity in warping functions across sensors within a subject motivates the pairwise comparison of warping functions across sensors presented in Sect. 4.3.

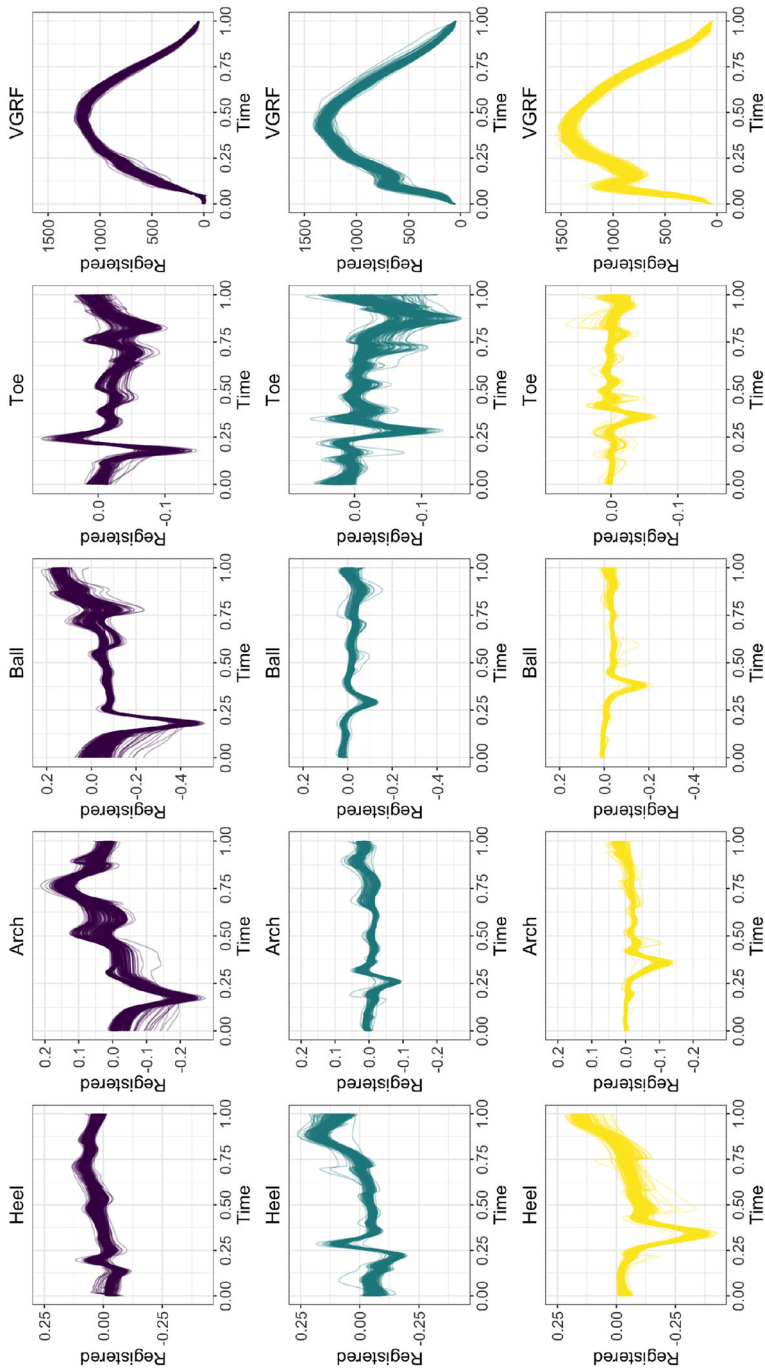


Fig. 4 Curve after registration: Registered curves for three runners are shown in rows; curves from in-shoe sensors at various locations are shown in the first four columns. Unregistered VGRF measurement is shown in the fifth column. Curves are color coded per subject, and each curve represents a stance

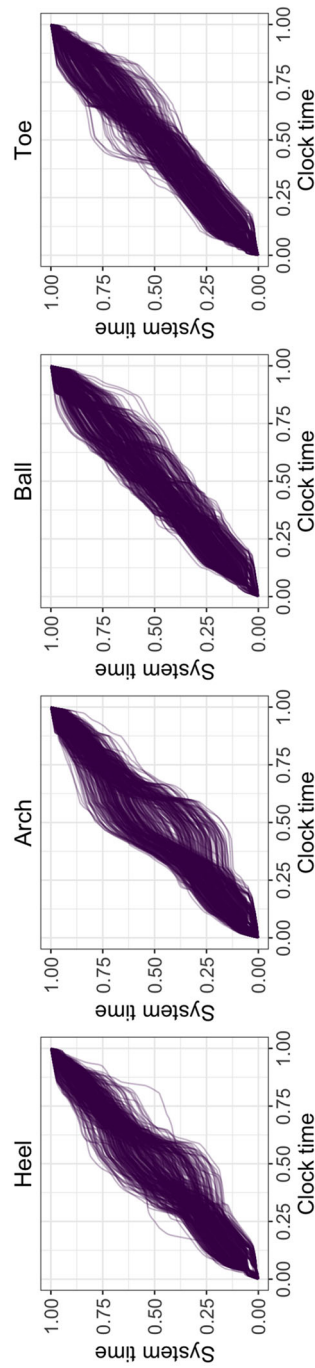


Fig. 5 Warping functions: Each panel represents warping function estimates of Subject 1 of four different in-shoe sensors. The clock time (x-axis) is the curve-specific time before registration, and the system time (y-axis) is the registered time which is common for every curve. Each curve represents a stance

4.2 Sources of Variability

Table 1 presents the amplitude variability, defined in Sect. 3.2, before and after curve registration. As expected, all sensors across subjects have smaller amplitude variability after curve registration. Many sensors have large reductions in amplitude variability, often by as much as 80–90 %; for example, the heel sensor of Subject 2 and the heel and arch sensors of Subject 3. The substantial reduction in amplitude variability indicates that the registration has a substantial effect in understanding the sources of variability in sensor data.

After curve registration, we can better understand the patterns that underly amplitude variability. We conduct FPCA to identify the dominant direction of variation in the registered curves. The first two FPCs for all subjects and sensors are plotted in Fig. 6. Between 24 and 50% of curve variability is explained by the first FPC, and the first two FPCs together explain between 39 and 75%. The FPCs for each subject, even those coming from the same sensor, are very different; this supports the uniqueness of running patterns from subject to subject. It is also noteworthy that for only a few subjects, major patterns of variation coincide with the location of major peaks, while, in most cases, the FPCs are relatively flat in the areas where major features are observed. For example, the peak in the heel sensor does not show up in either the first or the second FPC, suggesting much of the variability in the sensor is not related to the magnitude of the largest force but instead lies elsewhere in the sensor recordings.

4.3 Phase Variability Across Sensors

Using the permutation test framework described in Sect. 3.3 with 1000 permuted datasets, correlations between warping functions within stance across sensors are significant for all subjects and all sensor pairs; the p-value for each pairwise comparison is less than 0.001. These results indicate that the phase variability in-shoe sensors are more similar within a stance than across stances, perhaps suggesting that the process underlying phase variation depends on the stance itself.

Test statistics for each pairwise comparison of all subjects are presented in Table 2. The magnitude of the correlation varies across sensors and, in many cases, the correlation is relatively small. Together with the finding of strong statistical significance, the low observed correlations suggest that although some phase variability in sensors is similar due to processes underlying the stance, a large proportion of phase variability is dissimilar across sensors in the same stance. Stated differently, although the warping

Table 1 Comparison of variability: amplitude variability ($\times 10^5$) before and after registration

	Heel		Arch		Ball		Toe	
	Before	After	Before	After	Before	After	Before	After
Subject 1	38.50	20.70	183.84	49.39	519.77	78.77	48.43	9.01
Subject 2	196.76	28.76	24.44	5.09	45.62	9.22	51.95	19.80
Subject 3	340.98	39.93	30.94	3.48	57.22	4.83	9.09	2.65

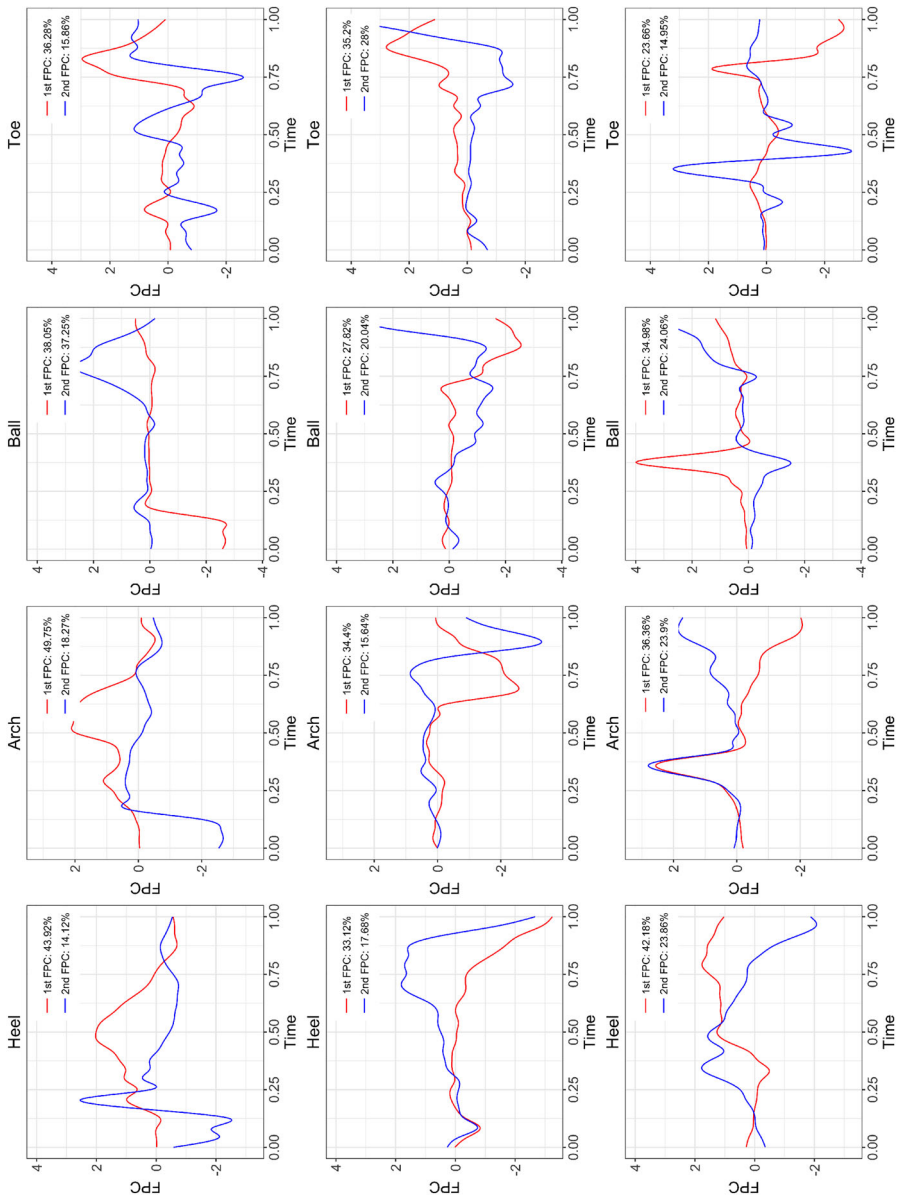


Fig. 6 Functional principal components: First two FPCs of registered curves for three runners are shown in rows; the FPCs in four in-shoe sensors are shown in columns

Table 2 Test statistic of pairwise comparison: for each subject, the upper diagonal elements are the mean correlation within stance across sensors

	Subject 1			Subject 2			Subject 3		
	Arch	Ball	Toe	Arch	Ball	Toe	Arch	Ball	Toe
Heel	0.617	0.240	0.094	0.268	0.264	0.411	0.378	0.280	0.238
Arch		0.315	0.125		0.262	0.258		0.337	0.300
Ball			0.501			0.239			0.276

functions obtained through the registration process are similar for different sensors in the same stance, substantial dissimilarities in warping functions remain.

Recall that Subject 1 is a mid-foot striker and the other subjects are heel strikers. The pairwise correlation values in Table 2 suggest different correlation patterns depending on the running style. In case of mid-foot striker, the heel sensor shows the biggest correlation with arch (0.617) while the smallest correlation appears between heel and toe (0.094). More generally, adjacent sensors have greater strength of correlation within stance. Heel strikers in contrast have roughly uniform correlation across sensors; adjacency does not matter. It is possible that the mid-foot striker has a smoother transition of forces in different parts of the foot within a stance compared to the heel strikers, which may help explain why mid-foot strikers have been hypothesized to have lower risks of certain musculoskeletal injuries. Analyses of additional runners with varying gait patterns will help clarify this hypothesis.

4.4 Association Between In-Shoe Sensors and GRF

In this section, we use the VGRF curves as a response and fit function-on-function regression models. We choose to focus on a single subject for the bulk of our analysis and interpretation to make a detailed case study of a single subject; we focus on Subject 2 to examine a heel striker with a discernible impact transient, which is of broad interest in the analysis of gait and injury. Although we present analyses for Subject 2 in the main text, we include similar results for the remaining two subjects in Appendix.

We start with an exploration of the relationship between VGRF and in-shoe sensors to examine the data-generating mechanism. Figure 7 shows the correlation surface for VGRF and in-shoe sensor curves for Subject 2. These data do not contain a specific pattern, like a clear off diagonal band or an obvious peak along the diagonal (which might suggest a lagged or concurrent model). Keeping that in mind, we proceed with general function-on-function regression models.

We conduct a cross-validation study of models relating the VGRF to sensor predictors. We compare the predictive accuracy of function-on-function regression models using the heel, arch, ball and toe sensors in isolation, and these models using all sensors. We randomly select 20% of the data as a test set, fit five function-on-function regression models, predict VGRFs for the test set, and compute mean integrated squared errors; this process is repeated 100 times, and the results for Subject 2 are shown in Fig. 8.

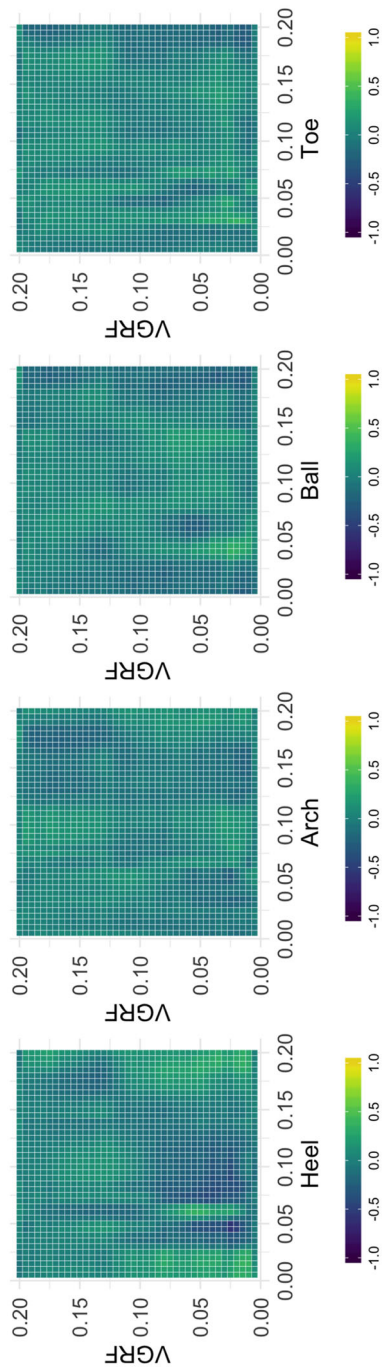


Fig. 7 Correlation between VGRF and in-shoe sensor curves; Correlation between VGRF (y-axis) and in-shoe sensors (x-axis) for Subject 2

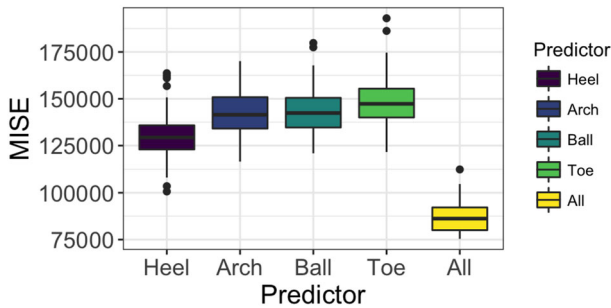


Fig. 8 Cross-validation study: Box plots of mean integrated squared errors of function-on-function regression models. Each box plot is created based on 100 simulations

The cross-validation study suggests that the heel sensor is most important in isolation but that no single sensor performs as well as using all sensors. As a result, we choose to fit a function-on-function regression model specified below:

$$\text{VGRF}_i(t) = \beta_0(t) + \sum_{s=1}^4 \int_0^1 y_i^s(u) \beta^s(u, t) du + \varepsilon_i(t). \quad (3)$$

Figure 9 below shows estimated coefficient surfaces of the function-on-function regression model with all four in-shoe sensors for Subject 2. The coefficient surfaces are interpreted by integrating the product of the predictor and surface at a time t for the $\text{VGRF}(t)$. For example, fixing $t = 0.1$, which is roughly the location of the initial peak, the heel sensor coefficient $\beta^1(u, t)$ suggests that the contrast between sensor values in the middle and edges of the u domain drives the fitted value for the VGRF. In particular, stances with high starting and ending values and low middle values from heel sensor may have higher initial peaks in the VGRF.

Fitted values and residuals from this function-on-function regression model are presented in Fig. 10. The fitted values are similar to the observed VGRF measurements (shown in the second row and fifth column in Fig. 3); the function-on-function regression model captures the major features of observed curves, including the over-

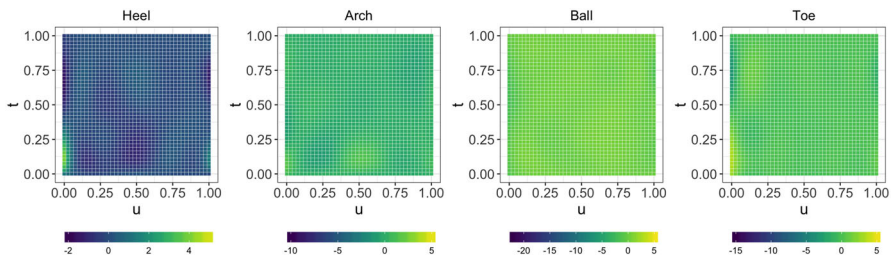


Fig. 9 Function-on-function regression coefficient estimates: Heat map of function-on-function regression coefficient estimates ($\times 10^4$) with all four sensors as predictors. VGRF is measured over t (y-axis) and in-shoe sensors are measured over u (x-axis). Color gradation represents the magnitude of coefficient estimates

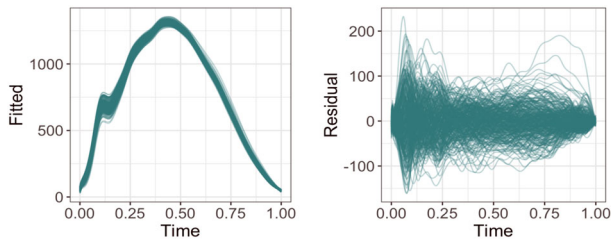


Fig. 10 Model fitting of function-on-function regression: Fitted values (Left) and residuals (Right) of function-on-function regression for Subject 2

all shape and impact transient. This suggests that in-shoe sensor data may be indeed useful for predicting important features or the whole of VGRF curve. However, the residuals indicate that variability in the VGRF is not wholly captured by the function-on-function regression model, especially near the impact transient.

Results for Subjects 1 and 3 also indicate that the heel sensor is the best predictor in isolation but that a model with all sensors has the best performance (Fig. 12 in Appendix) but that the estimated coefficients are distinct across subjects, implying a unique relationship between in-shoe sensors and VGRF curves for each subject (Fig. 13 in Appendix). Heel strikers (Subjects 2 and 3) may have more similar coefficients compared to the mid-foot striker (Subject 1). Exploring this in more detail will require a larger study population composed of both heel and mid-foot strikers, and the use of a function-on-function regression model that allows coefficients to vary across subjects.

5 Discussion

In this project, we explored in-shoe sensor data observed from three healthy subjects obtained during an experiment evaluating gait. We examined both phase and amplitude variabilities in the observed data and illustrated the importance of aligning the observed curves via curve registration. Because the observed phase variability is not expected in stances of healthy individuals, the processing of data using registration methods is an important step in obtaining reliable data from in-shoe sensors; the results of registration also shed light on the properties of the in-shoe sensors and may help to refine the development of this new technology.

In our analyses, we examined the utility of curve registration by comparing the amplitude variability before and after curve registration, which identified a reduction in amplitude variability after curve registration. We further investigated the similarity in estimated warping functions to understand the sources of phase variability across sensors within stance. Our permutation test results indicate that within each stance, time shifts are related across sensors, but that much of the phase variability across stances is dissimilar within the same sensor. This correlation may support the development of a hierarchical approach to understand the shared- and sensor-specific phase variation within a stance.

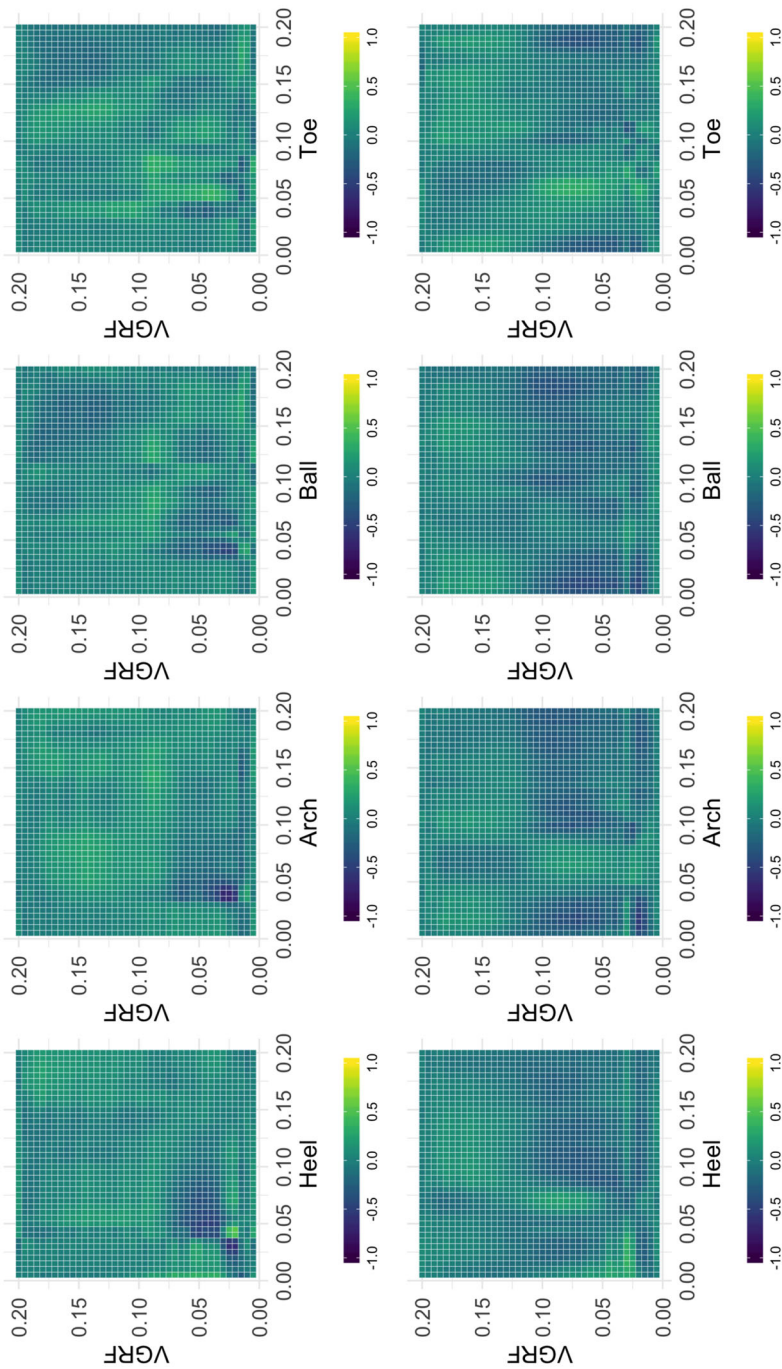


Fig. 11 Correlation between VGRF and in-shoe sensor curves: Correlation between VGRF (y-axis) and in-shoe sensors (x-axis) for Subject 1 (Top) and Subject 3 (Bottom)

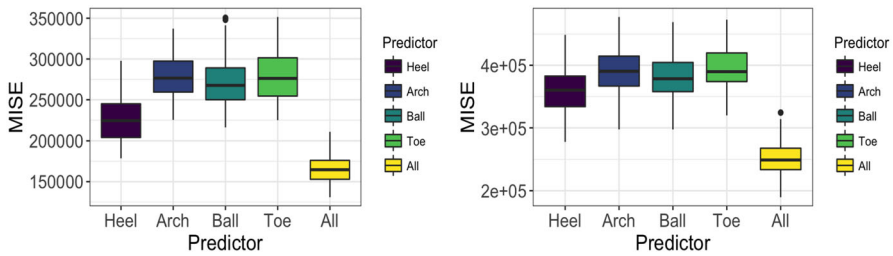


Fig. 12 Cross-validation study: Box plots of mean integrated squared errors of function-on-function regression models for Subject 1 (Left) and Subject 3 (Right). Each box plot is created based on 100 simulations

For examining the association between VGRF and in-shoe sensors, we explored modeling approaches including concurrent models and function-on-function regression models; for poor performance, we omitted the results from concurrent models. The correlation surfaces for VGRF and in-shoe sensor curves did not show a specific pattern, suggesting to fit general function-on-function regression models instead of lagged or concurrent models. The cross-validation study further verified that concurrent models performed uniformly worse than function-on-function regression models.

We used function-on-function regression to evaluate in-shoe sensor data as surrogate of VGRF measurements. Our results provide some initial evidence for a relationship between key features of in-shoe sensors and VGRF curves: according to our results, there is some signal in the in-shoe sensors but subject-to-subject variations in performance and regression coefficients are high, and additional work is needed to provide insights into modeling this variation. It may also be the case that in-shoe sensors become a useful complement to, but not replacement for, in-lab measurements.

Although we have focused on healthy subjects, it is also of interest to use in-shoe data to diagnose pathologies. Both phase and amplitude variabilities can be an evidence of pathological conditions such as movement disorders (Parkinson's disease and Huntington's disease; Hausdorff et al. [10]), age effects [22], and psychological disorders (major depressive disorder and bipolar disorder; Hausdorff et al. [11]). Including individuals with pathological disorders and investigating their time shifts will emphasize the clinical importance of curve registration and enhance the quality of gait analysis after preprocessing.

Acknowledgements This work was supported in part by NSF CMMI Award #1538447. The last author's research was supported by the Award #R01HL123407 from the National Heart, Lung, and Blood Institute, and by the Award #R01NS097423-01 from the National Institute of Neurological Disorders and Stroke.

Appendix

In Sect. 4.4, we apply functional analysis methods to examine the association between VGRF measurement and in-shoe sensor curves for Subject 2; here, we present the results of functional regression models for Subjects 1 and 3.

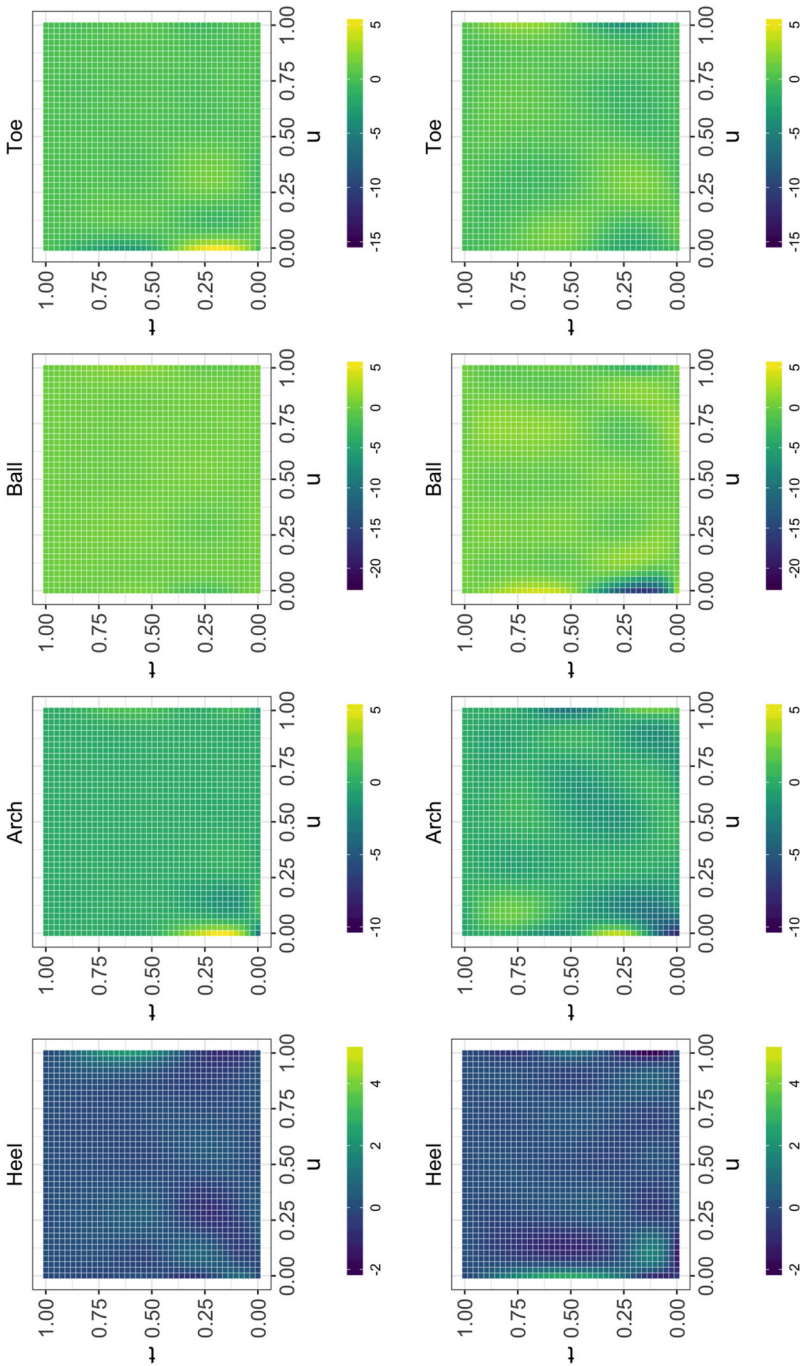


Fig. 13 Function-on-function regression coefficient estimates: Heat map of function-on-function regression coefficient estimates ($\times 10^4$) with all four sensors as predictors for Subject 1 (Top) and Subject 3 (Bottom). VGRF is measured over t (y-axis) and in-shoe sensors are measured over u (x-axis). Color gradation represents the magnitude of coefficient estimates

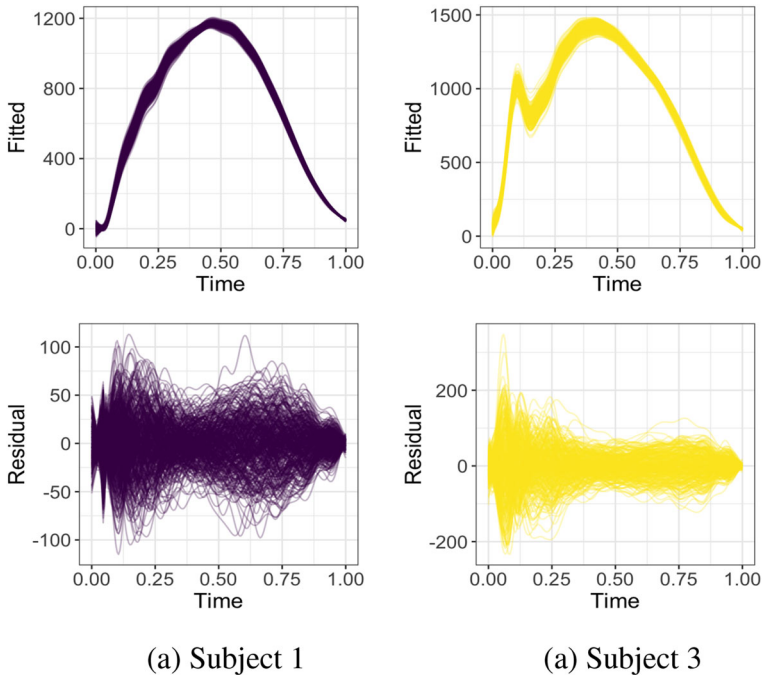


Fig. 14 Model fitting of function-on-function regression: Fitted values (Top) and residuals (Bottom) of function-on-function regression for Subjects 1 and 3

Analogous to Fig. 7, Fig. 11 shows the correlations between VGRF and in-shoe sensor curves for Subjects 1 and 3. Similar to Subject 2, these two subjects show neither a clear off-diagonal band nor an obvious peak along the diagonal.

The results in Fig. 12 of cross-validation study for Subjects 2 and 3 are similar to those in Fig. 8 from the main manuscript. For both Subjects 1 and 3, the heel seems most useful, but the model with all four predictors performs the best.

Figures 13 and 14 are analogous to Figs. 9, and 10 in the main manuscript, respectively. Rows in Fig. 13 show the estimated coefficients of function-on-function regression models with all four in-shoe sensors as predictors for Subjects 2 and 3. The estimated coefficients are distinctively different across subjects, implying the unique relationship between VGRF and in-shoe sensors for each subject.

References

1. Benedetti M, Merlo A, Leardini A (2013) Inter-laboratory consistency of Gait analysis measurements. *Gait Posture* 38(4):934–939
2. Bilodeau RA, Fullwood DT, Colton JS, Yeager JD, Bowden AE, Park T (2015) Evolution of nano-junctions in piezoresistive nanostrand composites. *Compos Part B Eng* 72:45–52
3. Crane EA, Cassidy RB, Rothman ED, Gerstner GE (2010) Effect of registration on cyclical kinematic data. *J Biomech* 43(12):2444–2447

4. Daoud AI, Geissler GJ, Wang F, Saretzky J, Daoud YA, Lieberman DE (2012) Foot strike and injury rates in endurance runners: a retrospective study. *Med Sci Sports Exerc* 44(7):1325–1334
5. Earls C, Hooker G (2017) Combining functional data registration and factor analysis. *J Comput Graph Stat* 26(2):296–305
6. Earls C, Hooker G (2017) Variational bayes for functional data registration, smoothing, and prediction. *Bayesian Anal* 12(2):557–582
7. Gasser T, Kneip A (1995) Searching for structure in curve samples. *J Am Stat Assoc* 90(432):1179–1188
8. Goldsmith J, Greven S, Crainiceanu C (2013) Corrected confidence bands for functional data using principal components. *Biometrics* 69(1):41–51
9. Goldsmith J, Scheipl F, Huang L, Wrobel J, Gellar J, Harezlak J, McLean MW, Swihart B, Xiao L, Crainiceanu C, Reiss PT (2016) Refund: regression with functional data. R package version 0.1-16. <https://CRAN.R-project.org/package=refund>
10. Hausdorff JM, Cudkowicz ME, Firtion R, Wei JY, Goldberger AL (1998) Gait variability and basal Ganglia disorders: stride-to-stride variations of Gait cycle timing in Parkinson's disease and Huntington's disease. *Mov Disord* 13(3):428–437
11. Hausdorff JM, Peng CK, Goldberger AL, Stoll AL (2004) Gait unsteadiness and fall risk in two affective disorders: a preliminary study. *BMC Psychiatry* 4(1):39
12. Helwig NE, Hong S, Hsiao-Weckler ET, Polk JD (2011) Methods to temporally align Gait cycle data. *J Biomech* 44(3):561–566
13. Hyldahl RD, Evans A, Kwon S, Ridge ST, Robinson E, Hopkins JT, Seeley MK (2016) Running decreases knee intra-articular cytokine and cartilage oligomeric matrix concentrations: a pilot study. *Eur J Appl Physiol* 116(11–12):2305–2314
14. Johnson OK, Kaschner GC, Mason TA, Fullwood DT, Hansen G (2011) Optimization of nickel nanocomposite for large strain sensing applications. *Sens Actuators A Phys* 166(1):40–47
15. Karamanidis K, Arampatzis A, Brüggemann GP (2004) Reproducibility of electromyography and ground reaction force during various running techniques. *Gait Posture* 19(2):115–123
16. Kneip A, Gasser T (1992) Statistical tools to analyze data representing a sample of curves. *Ann Stat* 20:1266–1305
17. Kneip A, Ramsay JO (2008) Combining registration and fitting for functional models. *J Am Stat Assoc* 103(483):1155–1165
18. Kneip A, Li X, MacGibbon K, Ramsay J (2000) Curve registration by local regression. *Can J Stat* 28(1):19–29
19. Marron J, Ramsay JO, Sangalli LM, Srivastava A (2014) Statistics of time warpings and phase variations. *Electr J Stat* 8(2):1697–1702
20. Marron JS, Ramsay JO, Sangalli LM, Srivastava A (2015) Functional data analysis of amplitude and phase variation. *Stat Sci* 30(4):468–484
21. Merrell AJ, Fullwood DT, Bowden AE, Remington TD, Stolworthy DK, Bilodeau A (2013) Applications of nano-composite piezoelectric foam sensors. In: ASME 2013 conf on smart materials, adaptive structures and intelligent systems
22. Owings TM, Grabiner MD (2004) Variability of step kinematics in young and older adults. *Gait Posture* 20(1):26–29
23. Peter A, Rangarajan A (2006) Shape analysis using the Fisher-Rao Riemannian metric: unifying shape representation and deformation. In: 3rd IEEE international symposium on biomedical imaging: nano to macro, 2006. IEEE, pp 1164–1167
24. Pietrosimone B, Loeser RF, Blackburn JT, Padua DA, Harkey MS, Stanley LE, Luc-Harkey BA, Ulici V, Marshall SW, Jordan JM (2017) Biochemical markers of cartilage metabolism are associated with walking biomechanics 6-months following anterior cruciate ligament reconstruction. *J Orthop Res* 35:2288–2297
25. Ramsay J, Li X (1998) Curve registration. *J R Stat Soc Ser B* 60(2):351–363
26. Ramsay JO (1998) Estimating smooth monotone functions. *J R Stat Soc Ser B* 60(2):365–375
27. Rice H, Jamison S, Davis I (2016) Footwear matters: influence of footwear and foot strike on load rates during running. *Med Sci Sports Exerc* 48(12):2462–2468
28. Rosquist PG, Collins G, Merrell AJ, Tuttle NJ, Tracy JB, Bird ET, Seeley MK, Fullwood DT, Christensen WF, Bowden AE (2017) Estimation of 3d ground reaction force using nanocomposite piezo-responsive foam sensors during walking. *Ann Biomed Eng* 45:2122–2134

29. Sadeghi H, Allard P, Shafie K, Mathieu PA, Sadeghi S, Prince F, Ramsay J (2000) Reduction of gait data variability using curve registration. *Gait Posture* 12(3):257–264
30. Sadeghi H, Mathieu PA, Sadeghi S, Labelle H (2003) Continuous curve registration as an intertrial gait variability reduction technique. *IEEE Trans Neural Syst Rehabil Eng* 11(1):24–30
31. Scheipl F, Staicu AM, Greven S (2015) Functional additive mixed models. *J Comput Graph Stat* 24(2):477–501
32. Seeley MK, Son SJ, Kim H, Hopkins JT (2017) Walking mechanics for patellofemoral pain subjects with similar self-reported pain levels can differ based upon neuromuscular activation. *Gait Posture* 53:48–54
33. Seliktar R, Yekutieli M, Bar A (1979) Gait consistency test based on the impulse-momentum theorem. *Prosthet Orthot Int* 3(2):91–98
34. Silverman BW (1995) Incorporating parametric effects into functional principal components analysis. *J R Stat Soc Ser B* 57:673–689
35. Srivastava A, Klassen E, Joshi SH, Jermyn IH (2011a) Shape analysis of elastic curves in Euclidean spaces. *IEEE Trans Pattern Anal Mach Intell* 33(7):1415–1428
36. Srivastava A, Wu W, Kurtek S, Klassen E, Marron J (2011b) Registration of functional data using Fisher-Rao metric. [arXiv:1103.3817](https://arxiv.org/abs/1103.3817)
37. Teng HL, Wu D, Su F, Pedroia V, Souza RB, Ma CB, Li X (2017) Gait characteristics associated with a greater increase in medial knee cartilage $t1\rho$ and $t2$ relaxation times in patients undergoing anterior cruciate ligament reconstruction. *Am J Sports Med* 45:3262–3271
38. Thies SB, Tresadern PA, Kenney LP, Smith J, Howard D, Goulermas JY, Smith C, Rigby J (2009) Movement variability in stroke patients and controls performing two upper limb functional tasks: a new assessment methodology. *J Neuroeng Rehab* 6(1):2
39. Tucker JD, Wu W, Srivastava A (2013) Generative models for functional data using phase and amplitude separation. *Comput Stat Data Anal* 61:50–66
40. Tucker JD, Wu W, Srivastava A (2014) Analysis of proteomics data: phase amplitude separation using an extended Fisher-Rao metric. *Electron J Stat* 8(2):1724–1733
41. Wu W, Srivastava A (2014) Analysis of spike train data: alignment and comparisons using the extended Fisher-Rao metric. *Electron J Stat* 8(2):1776–1785
42. Yao F, Müller HG, Wang JL (2005) Functional data analysis for sparse longitudinal data. *J Am Stat Assoc* 100(470):577–590