Ankle plantarflexion strength in rearfoot and forefoot runners: a novel clusteranalytic approach

Dominik Liebl[a,*], Steffen Willwacher[b], Joseph Hamill[c], Gert-Peter Brüggemann[b]

[a]*Mathematics Department and ECARES, Université Libre de Bruxelles*
[b]*Institute of Biomechanics and Orthopaedics, German Sports University Cologne*
[c]*Department of Exercise Science, University of Massachusetts, Amherst, MA*

Abstract

The purpose of the present study was to test for differences in ankle palantarflexion strengths of habitually rearfoot and forefoot runners. In order to approach this issue, we revisit the problem of classifying different footfall patterns in human runners. A dataset of 119 subjects running shod and barefoot (speed 3.5 m/s) was analyzed. The footfall patterns were clustered by a novel statistical approach, which is motivated by advances in the statistical literature on functional data analysis. We explain the novel statistical approach in detail and compare it to the classically used strike index of Cavanagh and Lafortune (1980).

The two groups found by the new cluster approach are well interpretable as a forefoot and a rearfoot footfall groups. The subsequent comparison study of the clustered subjects reveals that runners with a forefoot footfall pattern are capable of producing significantly higher joint moments in a maximum voluntary contraction (MVC) of their ankle plantarflexor muscles tendon units; difference in means: 0.28 Nm/kg. This effect remains significant after controlling for an additional gender effect and for differences in training levels.

Our analysis confirms the hypothesis that forefoot runners have a higher mean MVC plantarflexion strength than rearfoot runners. Furthermore, we demonstrate that our proposed stochastic cluster analysis provides a robust and useful framework for clustering foot strikes.

Keywords: Running, biomechanics, foot strike, barefoot, shod, functional data analysis, cluster analysis

---

Highlights

- Novel clusteranalytic approach to find groups in habitual footfall patterns.
- Found clusters are interpretable as forefoot and rearfoot footfall clusters.
- Shod forefoot runners have stronger plantarflexors than shod rearfood runners.

*Corresponding author: Dominik Liebl, Université Libre de Bruxelles, CP 210, local O.9.213, Bd du Triomphe, 1050 Bruxelles-Belgium

Email address: `dominik.liebl@ulb.ac.be` (Dominik Liebl)

1. Introduction

There have been anecdotal claims from coaches and some runners that forefoot running leads to a strengthening of biological structures surrounding the foot and ankle complex and thereby has a protective character with respect to typical running injuries (Lieberman, 2012). Daoud, Geissler, Wang, Saretsky, Daoud, and Lieberman (2012) could show that in a group of competitive shod cross country runners repetitive stress injury rates were significantly lower for forefoot striking runners than for rearfoot striking runners. Another typical observation is that, e.g., sprinters, who are capable of producing high plantarflexion moments, use a forefoot strike even at moderate running speeds. In fact, Williams, McClay, and Manal (2000) showed that habitually forefoot running subjects show significantly higher initial sagital ankle joint plantarflexion moments than habitually rearfoot running subjects. While all this suggests that there might be a systematic relationship between footfall patterns and the strength of the lower extremities, this relationship has not been investigated and tested yet.

This kind of hypothesis demands for data of subjects that were allowed to use their preferred footfall patterns—an issue which demands for an ex-post classification of the subjects' footfall patterns. This ex-post classification step is crucial, but often treated novercally. In fact, reliability of the chosen classification procedure is a necessary precondition for any subsequent inferential comparison study.

Generally, classification of footfall patterns can either be done visually, using sagittal plane high speed video camera data (e.g., Hasegawa, Yamauchi, and Kraemer, 2007), or on basis of quantitative data. The latter is often preferred in academics—presumably, due to its supposed objectivity. Researchers, who want to rely their judgment on quantitative data, usually use a combination of ground reaction force information and foot kinematics, which describe the initial landing of the feet on the ground. For example, Gruber, Umberger, Braun, and Hamill (2013) use the strike index (SI), ankle angle, and the vertical ground reaction force (GRF) at initial foot-ground contact. Classification of the footfall patterns is then done on basis of some ad-hoc decision rules. Unfortunately, just these ad-hoc decision rules run the risk of compromising reliability and, particularly, have to be used with caution in inferential studies. In the following, we use the example of the SI in order to discuss the general problem.

The SI of Cavanagh and Lafortune (1980) is the most established measure to quantify foot strikes, where foot strike classification is performed using a 1/3-decision rule: If the SI indicates an initial foot-ground contact in the rear, middle or front third of the subject's foot, the subject is classified as rearfoot striking (RFS), midfoot striking (MFS) or forefoot striking (FFS). This 1/3-decision rule constitutes a widely accepted, well interpretable ad-hoc classification procedure.

The main problem with respect to the 1/3-decision rule is visualized in Figure 1, where we show two scatter plots of the SI-points of the sample of subjects used in this study. Obviously, the scattered SI-points and the corresponding kernel density estimates indicate that there are actually clusters in the SI data of shod/barefoot running subjects. These clusters are likely to represent distinct footfall strategies. Unfortunately, the 1/3-decision rule does not account for these SI-clusters, but truncates them. This kind of truncation leads to groups of subjects with inhomogeneous SI values and the consequential truncation bias is likely to harm any further statistical inferential study (e.g., Cohen, 1991).

Besides this, there is also a conceptual problem with respect to variables like the SI, ankle angles, or GRF, when processed (only) at the initial foot-ground contact. It is indisputable that all of these variables allow for valuable quantitative descriptions of the footfall pattern. Though, it seems unquestionable that appropriate time-continuous statistics, generally shall do a better

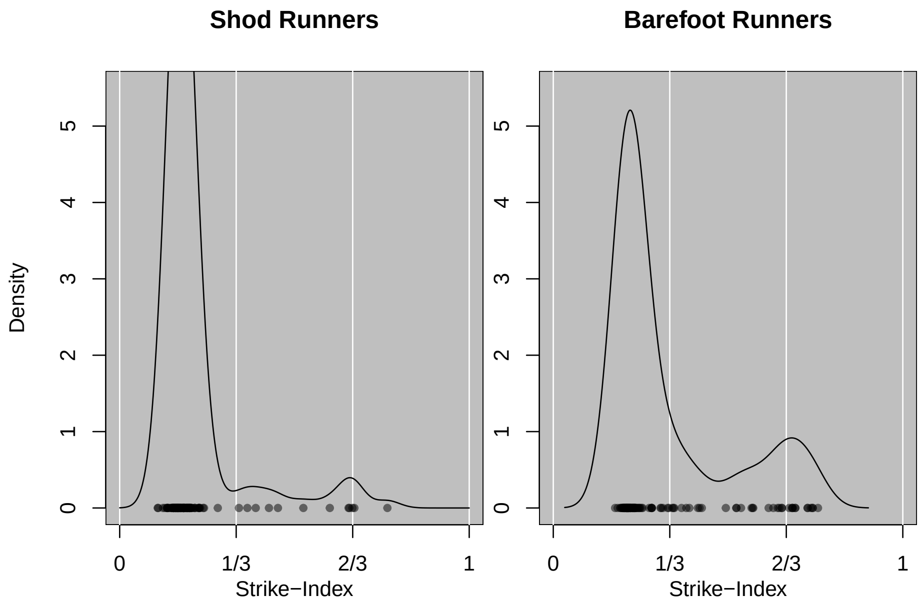

Figure 1: Scatter plots and kernel density estimates of the strike index under both treatment conditions measured for the sample of $n = 119$ subjects used in this study.

job in capturing the information on the dynamic footfall pattern. In fact, the ad-hoc decision rule together with its, let's say, "snap-shot" characteristic the SI can produce rather contradictory results. For example, the SI indicates that 108 shod runners of our dataset perform a RFS, but from these 108 runners 44 runners have initial plantarflexion moments. In the following, we propose solutions to both of the above described problems.

We propose to replace the ad-hoc decision rule by another procedure, which allows for a stochastic grouping of the observed footfall patterns and does not cause any truncation bias. A very simple yet powerful approach is to assume that the unknown distribution within each footfall cluster can be approximated by a Gaussian normal distribution. Besides the favorable simplicity of the Gaussian normal distribution, its uni-modality, symmetry, and rare tail events reflect very well the general demand on a cluster: namely, to represent a homogeneous group of subjects. The combination of Gaussian distributions leads to non-trivial, well interpretable Gaussian mixture distributions, which can be estimated from the data by the so-called EM-algorithm (Dempster, Laird, and Rubin, 1977).

The EM-algorithm alternates between an Estimation-step, in which the cluster-wise means and variances are estimated, and a Maximization-step, in which the subjects are newly allocated to the best fitting clusters. The further development of the EM-algorithm suggested by Fraley and Raftery (2002) allows also to determine the number of clusters from the data. This makes the EM-algorithm particularly useful in order to find clusters without prior knowledge on the cluster structure.

Furthermore, we propose to cluster a variable, which captures the whole dynamic of the footfall pattern. Williams et al. (2000) showed that habitually rearfoot and forefoot running subjects significantly differ in the initial courses of their vertical GRF, ankle angles, and sagital ankle joint moments. We use the sagital ankle joint moment courses for our cluster analysis, since these comprise of both other variables, the kinematic vertical GRF and the kinetic ankle angles. Specifically, we use the ankle joint moment functions in the sagittal plane for the initial first 20% of the stance phase; see Figure 2.

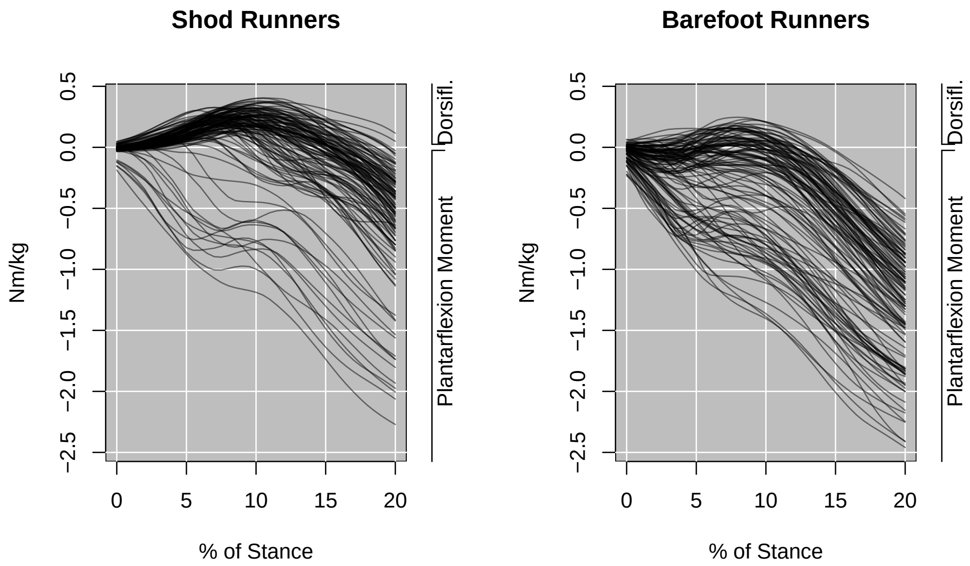

Figure 2: Joint moment curves of $n = 119$ shod and barefoot runners measured for the ankle in the sagittal plane. The x-axis denotes standardized time (in % of stance) and the y-axis denotes weight-standardized moments (in Nm/kg).

We will refer to the ankle joint moment courses as *functional* variables in order to emphasize that we regard them as realizations of so-called functional random variables. The latter differ from multivariate random variables, since they are potentially infinite dimensional objects. Fortunately, in practice it is usually possible to analyze functional data by classical methods from multivariate statistics after applying appropriate dimension reductions. Generally, the most common method for dimension reduction is to select one or more specific characteristic features of the functional data—the SI is a perfect example. In the statistical literature on functional data analysis, the most important method for dimension reduction is functional principal component analysis (FPCA) (Ramsay and Silverman, 2005, Ch. 8). Clusters found in the empirical distribution of the pc-scores correspond to clusters in the original functional data.

This paper has two purposes, where one is accessory to the other. We introduce a novel statistical approach to find clusters in the footfall patterns of human runners. After validation of the revealed rearfoot and forefoot clusters, we use them to investigate our fundamental purpose, i.e., to test for differences in ankle plantarflexor strengths between subjects with habitual rearfoot footfall patterns and those with habitual forefoot patterns. Specifically, we test the null hypothesis that there is no difference in the ankle joint moments of the ankle plantarflexor muscles tendon units between subjects with different footfall patterns against the alternative hypothesis that habitual forefoot footfall subjects have stronger plantarflexor muscles. Hereby, we focus only on the rearfoot and forefoot clusters of the shod treatment, since none of the subjects was a habitual barefoot runner. Furthermore, we control for an additional gender effect and differences in training levels.

In the following, we use the acronyms RFS, MFS, and FFS exclusively in order to refer to the classes determined by the SI and its 1/3-decision rule. The clusters found by the novel statistical approach will be referred to as rearfoot footfall (RFF) and forefoot footfall (FFF) clusters.

2. Methods and Procedures

2.1. Subjects

In the course of this study $n = 119$ male and female subjects (mass: 69.8 ± 11.5 kg; height: 1.75 ± 0.09 m; age: 38.6 ± 13.3 years) were analyzed at a running speed of 3.5 m/s. Runners had an average running experience of 10.3 ± 7.5 years and completed running exercise with an average mileage of 41.4 ± 20.1 km per week. No preselection with respect to the runner's striking behavior was done in order to get a representative sample of the typical striking behavior distribution of human runners.

2.2. Data collection

The force capacities of the right lower leg plantarflexors were determined by a maximum voluntary contraction (MVC) on a dynamometer (Biodex System 3, Biodex Medical Systems, Inc., Shirley, NY, USA). MVC testing was performed at an ankle angle of 90° and with fully extended legs (knee angle 180°). The MVC torques were normalized to body mass.

Subjects were sitting in the device and were advised to apply maximum voluntary force for a period of 5 seconds. Maximum isometric force was determined using a 500 ms sliding window approach. Each subject performed three trials with a self chosen rest period. The best of these trials was used for further analysis—see Arampatzis, Karamanidis, and Albracht (2007) on more details with respect to the testing protocol.

Movement analysis took place at the biomechanics lab located inside the track and field indoor facilities at the German Sport University, Cologne, Germany. Runners were asked to run along a 25 m long runway including a force platform (1250 Hz, 0.6 m × 0.9 m, 8 channel amplifier type 9865, Kistler Instrumente AG, Winterthur, Switzerland) in its middle. Kinematics of the right lower extremity were tracked using a ten camera Vicon Nexus system (250 Hz, Vicon Motion Systems, Oxford, UK) operating at 250 Hz. All runners wore the same kind of running shoes (Brooks Adrenaline, Brooks Sports Inc., Bothell, WA, USA).

The running surface was a 10 mm thick Tartan layer, which is covering the biomechanics lab. In the barefoot condition, a 13 mm thick layer of Ethylene Vinyl Acetate (EVA, Shore A 25) was attached to the Tartan base, in order to allow for barefoot running on a pleasant surface (like natural grass or similar). Runners were advised to run forward and backwards inside the lab in order to get accustomed to each new running condition for at least 200 m. If they needed more time for familiarization, they were given as much time as they needed.

Marker trajectories were filtered using a forth order recursive digital Butterworth low pass filter with a cut-off frequency of 20 Hz; GRF data was smoothed using the same filter but with a cut-off frequency of 100 Hz. We used different cut-off frequencies for trajectory and GRF data in order to stay in line with the already published literature in this field. Still, recent studies suggest to use matching cut-off frequencies in order to avoid artifacts that might result by the fact that accelerations of lower extremity segments do not correspond to measured GRF (e.g., Bisseling and Hof, 2006; Kristianslund, Krosshaug, and van den Bogert, 2012; Bezodis, Salo, and Trewartha, 2013). However, those studies were performed on movements with higher accelerations of the lower extremities than it is the case in our study of subjects running at a moderate speed. Furthermore, the above cited literature focuses on joint moment artifacts at the knee and the hip joint, though, in the calculation of the ankle joint moment exclusively the acceleration of the foot segment is needed. The foot has a relatively small mass, therefore, any potential acceleration measurement error due to strong filtering would only have minor effects on the ankle joint moment calculation. Future studies should investigate the effect of different filtering procedures in less impact prone movements like jogging on lower extremity joints other than the knee in order to highlight if there is a need to apply matching cut-off frequencies in these kind of movements.

A four segment inverse dynamics model was used to calculate three dimensional joint moments at the ankle joint. The marker protocol included both calibration and tracking markers. The former were removed following neutral standing reference measurement. Calibration markers were positioned on the left and right greater trochanters, right-side medial and lateral femoral condyles, and right-side medial and lateral maleoli, and positioned on the shoe over the first and fifth metatarsal heads. Tracking markers were positioned on the pelvis (right and left anterior and posterior iliac spines), thigh (rigid array of four markers), shank (rigid array of four markers), and calcaneus (three markers placed on the heel of the runner). Holes were cut into the heel-cup (diameter 30 mm) on its medial, lateral and posterior aspect, in order to avoid overestimation of heel movement during the shod condition (Maclean, Davis, and Hamill, 2009). Joint moments were expressed in the shank's anatomical coordinate system and were normalized to body mass.

No advice was made concerning a certain running style to assure that runners were choosing their preferred footfall pattern. A trial was judged valid if a running speed of 3.5 m/s ±5% was detected and no visible change in running technique in order to hit the force platform was observed. To ensure that runners were not accelerating during the analyzed stance phase, the ratio of propulsive to braking impulse of the anterio-posterior ground reaction force component had to be inside a range of 0.9 to 1.1 (Willwacher, Fischer, and Brüggemann, 2011). Running speed was monitored using two light barriers. Stance phases were determined by using a threshold of 20 N of the vertical ground reaction force component. After time normalization to stance durations we took the averages over five trials per subject in order to define the individual joint moment functions. The intra-subject correlations of the single joint moment functions were all greater than 95%.

2.3. Functional mixture model

The used statistical model can be seen as a simplified version of the model based clustering approach for multivariate functional data proposed in Jacques and Preda (2014). The statistical procedure can be divided in two distinct parts: First, (functional) principal component analysis of the joint moment curves. Second, cluster analysis of the principal component scores. The first part of the procedure builds upon recent advances in the literature on functional data analysis (FDA), which focuses on the statistical analysis of functions and curves; a very good introduction to FDA-methods can be found in Ramsay and Silverman (2005). A detailed discussion of the used method can be found in the appendix Appendix A. In the following, we give an easy to understand four-point description of how to implement the procedure and discuss different parametrizations of the subsequently used cluster analysis based on the EM algorithm as proposed by Fraley and Raftery (2002).

Our approach to cluster the joint moment curves can be applied as following:

1. Approximate the ankle joint moment functions $X_1(t), \ldots, X_i, \ldots, X_n(t)$ by their $M \times 1$ dimensional discretization vectors $\hat{X}_i = (\hat{X}_i(t_1), \ldots, \hat{X}_i(t_M))'$.
2. Approximate the model components of Equation (A.4) by (classical) principal component analysis (PCA) applied to the $n \times M$ dimensional data matrix $[\hat{X}_1, \ldots, \hat{X}_n]'$. PCA is a standard method implemented in many statistical software packages. For example, the R-function `princomp()` stores the estimated (and discretized) mean function $\hat{\mu}$ under the name `center`, the pc-score vectors $\hat{\beta}_i = (\hat{\beta}_{i1}, \ldots, \hat{\beta}_{iK})'$ under the name `scores` and the estimated (and discretized) eigenfunctions $\hat{f}_1, \ldots, \hat{f}_K$ under the name `loadings` R Core Team (2013).
3. Decide about a sufficient high number of K eigenfunctions, such that, for example, the first K eigenfunctions $\hat{f}_1, \ldots, \hat{f}_K$ explain more than 95% of the total variance. If the R-function `princomp()` is used, the standard deviations explained by the single eigenfunctions are stored under the name `sdev`.
4. Fit a Gaussian mixture model to the estimated $n \times K$ dimensional data matrix of pc-scores $[\hat{\beta}_1, \ldots, \hat{\beta}_n]'$, where $\hat{\beta}_i = (\hat{\beta}_{i1}, \ldots, \hat{\beta}_{iK})'$. This can be conveniently done using the EM algorithm of Fraley and Raftery (2002) for multivariate Gaussian mixture models, which is implemented in the R-function `Mclust()` of the ad-on R-package `mclust` (Fraley and Raftery, 2007).

The EM algorithm, as described in Fraley and Raftery (2007), fits many different Gaussian mixture models to the set of the estimated pc-scores $[\hat{\beta}_1, \ldots, \hat{\beta}_n]$ and chooses the best model on basis of the Bayesian Information Criterion (BIC). The models differ from each other with respect to the parametrizations of the $K \times K$ dimensional group-wise covariance matrices Σ_g. Once the models are chosen and estimated, a subject is classified to the cluster with its highest probability.

The most simple parametrization is given by $\Sigma_g = \lambda I_K$ (i.e. only one parameter for all G groups) and the most complex is given by $\Sigma_g = \lambda_g D_g A_g D_g'$ (i.e. $K(K+1)/2$ parameters for *each* of the G), where the matrix D_g holds the eigenvectors and the matrix $\lambda_g A_g$ the eigenvalues of the covariance matrix Σ_g. The most simple model essentially allows only for differences in the means of the clusters, but the variances are constrained to be equal. The most complex model additionally allows for rather diverse types of differences in variances: the parameter λ_g controls the volume of the gth cluster, the matrix A_g controls the shape of the gth cluster, and the matrix D_g its orientation. All different parameterizations as well as the usage of the R-package `mclust` are nicely discussed in Fraley and Raftery (2007).

3. Results

Cluster analysis

Before we present our empirical results, we introduce some further notations. Quantities that refer to the shod (barefoot) treatment are marked by a S-superscript (B-superscript), such as, e.g., $X_i^S(t)$ and $X_i^B(t)$. Furthermore, quantities that refer to the FRR (FFF) cluster are marked by an R-index (F-index), such as, e.g., $\hat{\mu}_R^S(t)$ and $\hat{\mu}_F^S(t)$.

For the joint moment functions from the shod treatment $X_i^S(t)$ as well as for those from the barefoot treatment $X_i^B(t)$ only $K = 2$ eigenfunctions suffice to approximate the original functions with 98% accuracy; see the right panels of Figures 3 and 4. The middle panels of Figures 3 and 4 are scatter plots of the 2 dimensional pc-score vectors to which we fit Gaussian mixture models using the EM-algorithm of Fraley and Raftery (2002). For both treatments the Gaussian mixture

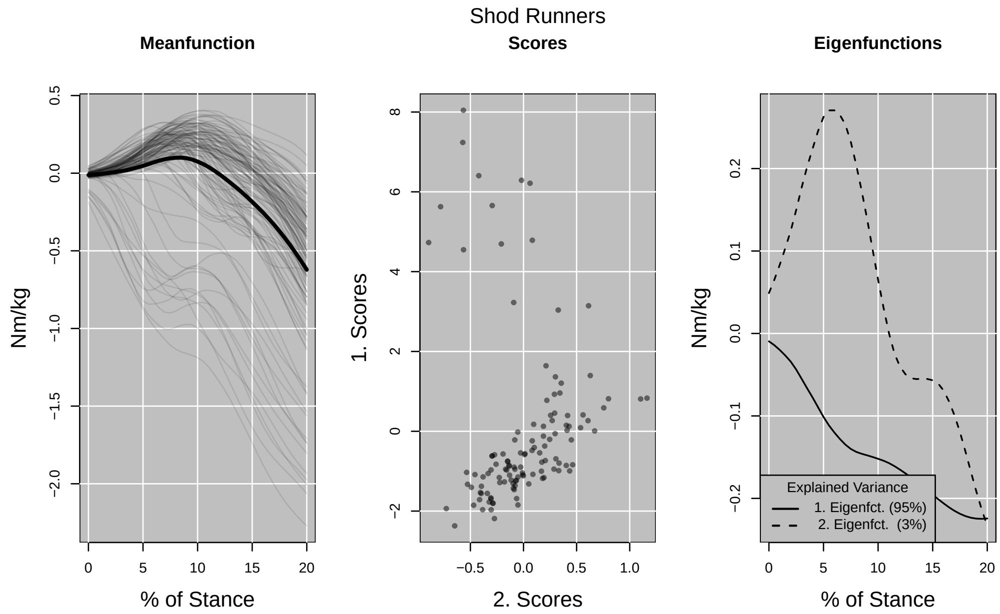

Figure 3: Left Panel: Scatter plot of joint moment functions $X_i^S(t)$ together with the empirical mean function $\hat{\mu}^S(t)$ (thick line). Middle Plot: Scatter plot of the pc-scores $\{\hat{\beta}_1^S, \dots, \hat{\beta}_n^S\}$. Right Plot: First two eigenfunctions $\hat{f}_1^S(t)$ and $\hat{f}_2^S(t)$.

model with the most flexible variance parametrization and two clusters (i.e., $G = 2$) maximize the BIC; see Table 1. This means, for each treatment condition the pc-score vectors are modeled by a Gaussian mixture model with two clusters, where the clusters are allowed to differ with respect to their volumes, shapes, and directions.

The corresponding probability density functions (pdf) of the 2 dimensional pc-score vectors of treatment $T \in \{S, B\}$ can be formulated as following:

$$\tau_R^T \phi(\mu_R^T, \Sigma_R^T) + \tau_F^T \phi(\mu_F^T, \Sigma_F^T) \quad \text{with} \quad T \in \{S, B\}, \tag{1}$$

where $\phi(\mu, \Sigma)$ denotes the bi-variate Gaussian normal pdf with mean vector μ and covariance matrix Σ. The corresponding parameter estimates are shown in Table 2. Except for the proportion parameters τ_C^T, the estimated parameters are comparable across the treatment conditions $T \in \{S, B\}$. The proportion parameters τ_R^T and τ_F^T with $\tau_F^T = (1 - \tau_R^T)$ quantify which fractions of the $n = 119$ subjects are allocated to the RFF and FFF clusters within treatment $T \in \{S, B\}$. While only $\tau_F^S = 10\%$ of the subjects perform a FFF in the shod condition, $\tau_F^B = 40\%$ perform a FFF in

Number of clusters	1	2	3	4
BIC (Shod)	-650.01	-455.07	-467.27	-492.04
BIC (Bare)	-693.35	-629.94	-647.67	-661.36

Table 1: BIC-values for different numbers of clusters G of the Gaussian mixture model with the most flexible variance parametrizations. Note: The BIC values are computed as in Fraley and Raftery (2002) and have to be maximized.

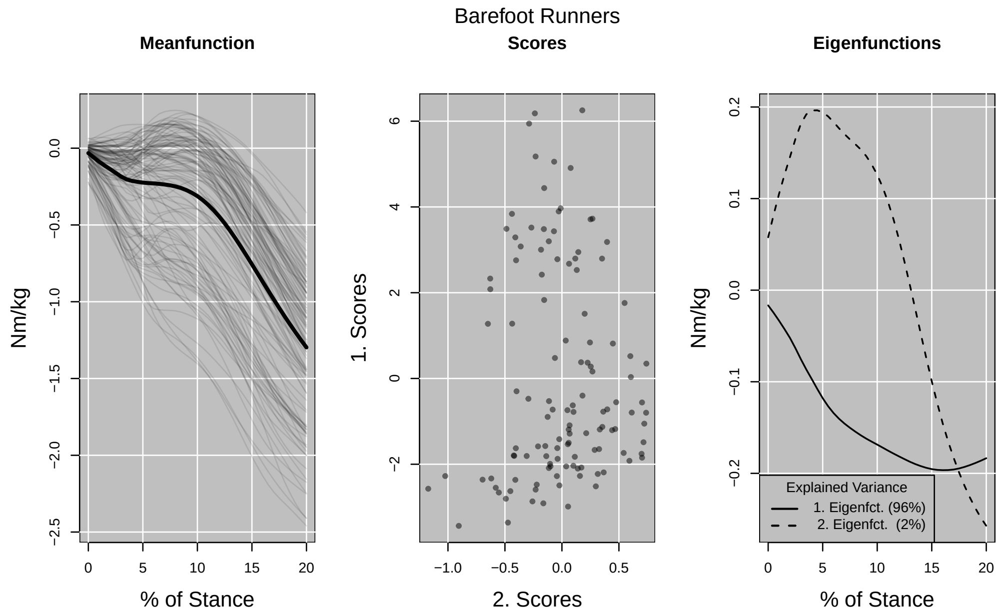

Figure 4: Left Panel: Scatter plot of joint moment functions $X_i^B(t)$ together with the empirical mean function $\hat{\mu}^B(t)$ (thick line). Middle Plot: Scatter plot of the pc-scores $\{\hat{\beta}_1^B, \dots, \hat{\beta}_n^B\}$. Right Plot: First two eigenfunctions $\hat{f}_1^B(t)$ and $\hat{f}_2^B(t)$.

the barefoot condition.

Rearfoot (Shod)	$\mu_R^S = \begin{pmatrix} -0.7 \\ 0.1 \end{pmatrix}$	$\Sigma_R^S = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.1 \end{pmatrix}$	$\tau_R^S = 0.9$
Forefoot (Shod)	$\mu_F^S = \begin{pmatrix} 4.7 \\ -0.2 \end{pmatrix}$	$\Sigma_F^S = \begin{pmatrix} 2.8 & -0.4 \\ -0.4 & 0.2 \end{pmatrix}$	$\tau_F^S = 0.1$
Rearfoot (Bare)	$\mu_R^B = \begin{pmatrix} -1.7 \\ 0.02 \end{pmatrix}$	$\Sigma_R^B = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.2 \end{pmatrix}$	$\tau_R^B = 0.6$
Forefoot (Bare)	$\mu_F^B = \begin{pmatrix} 2.4 \\ -0.03 \end{pmatrix}$	$\Sigma_F^B = \begin{pmatrix} 3.5 & -0.1 \\ -0.1 & 0.1 \end{pmatrix}$	$\tau_F^B = 0.4$

Table 2: Estimated parameters of the bi-variate ($K = 2$) Gaussian mixture densities with two ($G = 2$) clusters.

The mixture densities for the pc-score vectors are given in Figure 5. There, points that are plotted as “u”s indicate subjects with a relative uncertainty of being correctly classified above 10%. The visual inspection confirms the plausibility of the cluster structure and reassures the model choice based on the BIC values. In fact, each of the clusters comprises a homogeneous set of pc-score vectors, which leads to homogeneous clusters of ankle joint moment functions. In order to assess whether it is actually justified to refer to the clusters as RFF and FFF clusters, we conduct a simple ex-post analysis. As already discussed in the introduction, the SI is a well interpretable descriptive measure that quantifies foot strikes. A comparison of the cluster-wise SI boxplots clearly supports the labeling of the pc-score clusters as rearfoot and forefoot footfall

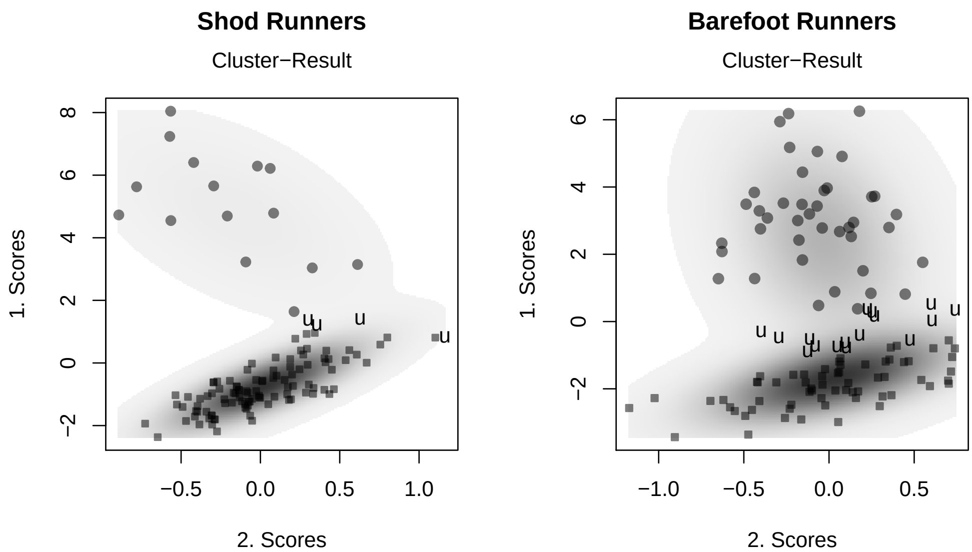

Figure 5: Visual impressions of the mixture densities for the pc-scores from the shod (left panel) and barefoot (right panel) treatment. Darker regions indicate regions with higher densities. Pc-scores of subjects that are classified into the rearfoot/forefoot cluster are plotted as squares/circles. Points that are plotted by u's indicate subjects with a relative uncertainty of being correctly classified above 10%.

clusters; see Figure 6.

The ankle joint moment functions of the RFF and FFF clusters are plotted in Figure 7. There, the mean functions can be interpreted as representative functions of their clusters, since similar functions are in their (close) neighborhood. The corresponding RFS, MFS, and FFS classes based on the SI and its 1/3-decision rule lead to a very different result—particularly, the MFS class shows a high intra-class variability; see Figure 8.

Given the two clusters found under each of the two treatments, we can identify three different types of subjects: (i) 60% of all subjects remain RFF under both treatments. (ii) 13% of all subjects remain FFF under both treatments. (iii) 27% of all subjects perform a RFF when running shod, but a FFF when running barefoot. The corresponding three groups, say: "Remain RFF", "Switch RFF to FFF", and "Remain FFF", are visualized in Figure 9.

The mean functions of the "Remain RFF" group and the "Switch RFF to FFF" group are relatively similar to each other; see the left and middle panel of Figure 9. In order to test the null hypothesis of equal mean functions against the alternative hypothesis that the mean function of the Remain RFF group lays above the mean function of the Switch RFF to FFF group, we test for equality in means of the corresponding pc-scores. Both t-tests are highly significant ($p < .01$) such that we can accept the alternative hypothesis of a lower mean function in the Switch group.

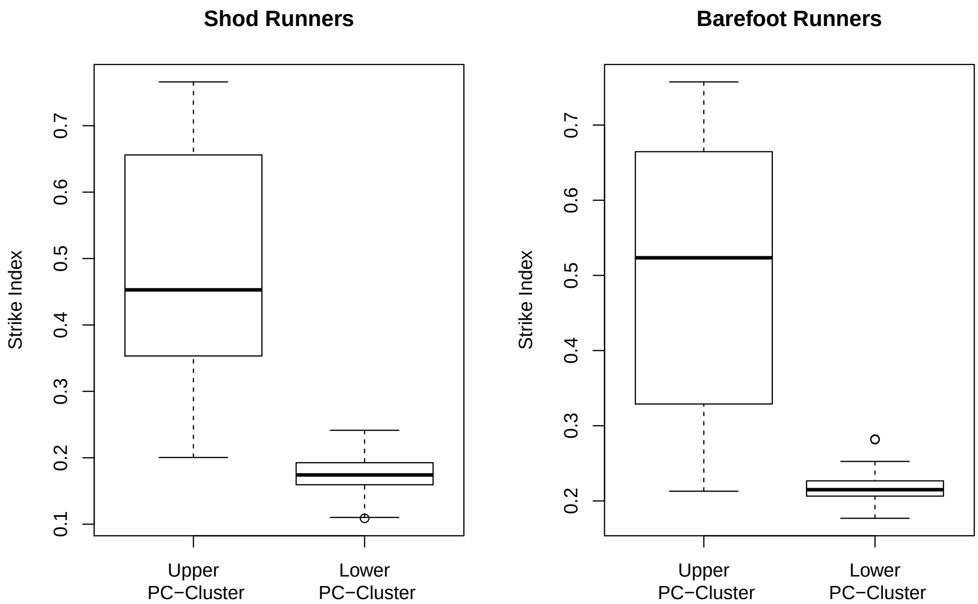

Figure 6: Boxplots of the cluster-wise SI values.

Comparison study

In this section, we test our fundamental null hypothesis on whether RFF and FFF subjects have equally strong plantarflexor muscles. We test this null hypothesis against the hypothesis that FFF subjects are capable of producing significantly higher ankle joint moments of their ankle plantarflexor muscles tendon units. In order to approach this, we focus only on the RFF and FFF clusters found under the shod treatment, since none of the subjects was a habitual barefoot runner.

In total $n_R^S = 104$ and $n_F^S = 15$ subjects were classified as RFF and FFF runners. Our alternative hypothesis implies a positive location shift of the pdf that belongs to the FFF subjects. This positive location shift is likely to be caused by differences in the compositions of the RFF and FFF clusters. For example, 50% of the subjects in the RFF cluster are male runners, while the FFF cluster consist of 60% male runners. The positive location shift might just as well be caused by this higher share of male runners. Furthermore, it is important to control for differences in the training levels of the subjects, which was measured by the co-variable "weekly kilometers run".

In order to test for a positive location shift in the pdf of FFF subjects under consideration of the discussed control variables, we use a two-way ANCOVA model with the factors "gender" and "strike-type" and the co-variable "weekly km run". For this analysis the total number of individuals is reduced from $n = 119$ to $n = 118$, since for one subject the statement on the co-variable is missing. In the following, we formalize the two-way ANCOVA model using dummy variables, as suggested in Gujarati (1970):

$$Y_{ijk} = \beta_0 + \beta_S D_S + \beta_G D_G + \beta_X X_{ijk} + \varepsilon_{ijk}, \quad (2)$$

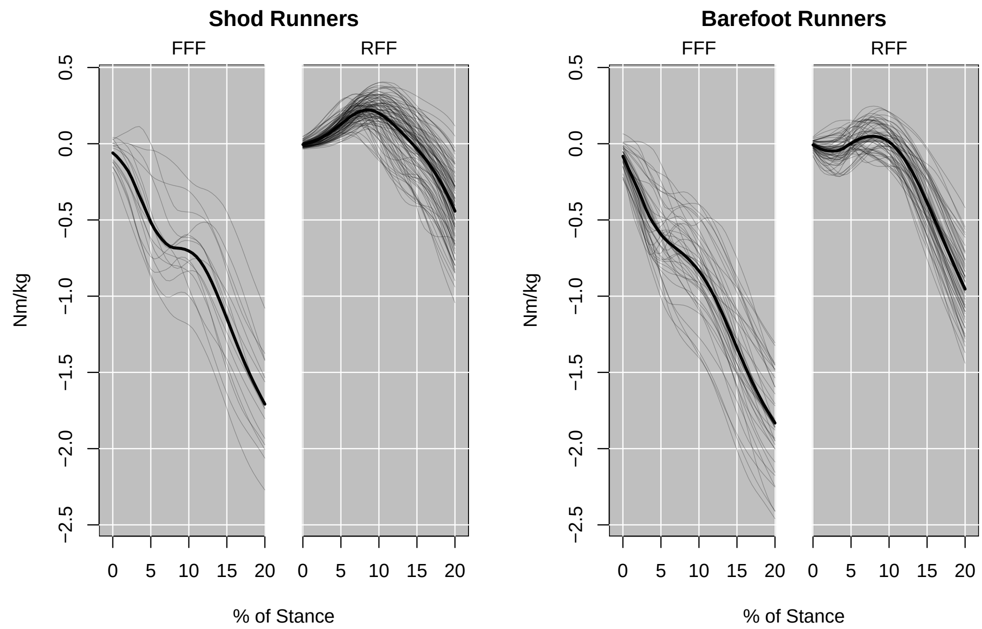

Figure 7: Joint moment curves of $n = 119$ shod and barefoot runners, separately plotted for rearfoot and forefoot striking subjects together with the estimated mean functions (thick lines) of each cluster.

where the index $i = 1, \ldots, n$ refers to the individuals, the index $j = 1, 2$ refers to the different footfall types, and the index $k = 1, 2$ refers to the different genders. The dummy variable D_S (D_G) is equal to one, if the corresponding subject is a FFF subject (male subject) and zero otherwise. The dependent variable Y denotes the MVC torques of plantarflexors and the independent random variable X denotes the co-variable "weekly km run". The β-parameters can be interpreted as following:

β_0: intercept for female, RFF subjects
β_S: differential intercept for FFF subjects
β_G: differential intercept for male subjects
β_X: slope coefficient of Y with respect to X

We use the rather simple Model (2) with a common slope-parameter for all groups and without interaction effects, since it is not significantly worse in explaining the variations of the dependent variable Y, than the corresponding completely unrestricted ANCOVA model with groupwise slope-parameters and interaction effects; (F-test, $p > .10$). The null hypothesis on whether Model (2) explains the variations of the dependent variable Y equally well as the further restricted model with $\beta_S = \beta_G = 0$, can be rejected in favor of Model (2) (F-test, $p \leq .01$).

The dummy variable notation in Model (2) makes it easy to formalize our null hypothesis, since a test of no footfall type effect and no gender effect against the alternatives of a positive FFF effect and a positive male effect is equivalent to test whether $\beta_S = 0$ and $\beta_G = 0$ against $\beta_S > 0$ and $\beta_G > 0$. These hypothesis can be tested by t-tests, where we adjust the p-values of the

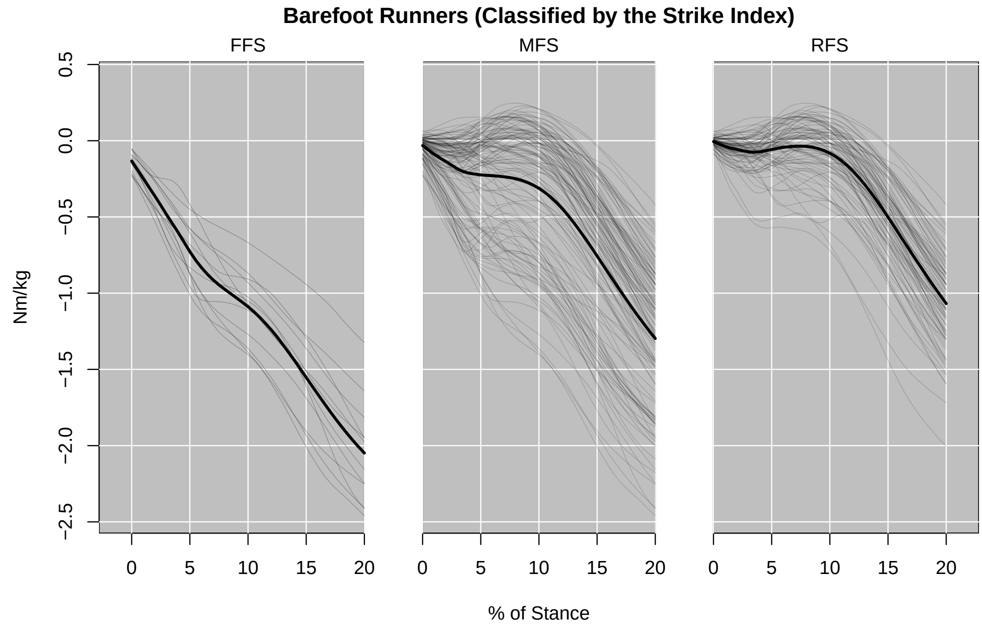

Figure 8: Joint moment curves of n = 119 barefoot runners, separately plotted for rearfoot, midfoot and forefoot striking subjects according to the classical strike index of Cavanagh and Lafortune (1980). The estimated cluster-wise mean functions are plotted as thick lines.

t-tests using the method suggested in Benjamini and Hochberg (1995), in order to control for the false discovery rate of this multiple testing problem. The estimation results are shown in Table 3.

Parameter	Estimate	Std. Error	t-value	Adj. p-value
β_0	2.29	0.12	18.56	0.00 (two-tailed)
β_G	0.32	0.10	3.24	0.00 (one-tailed)
β_S	0.28	0.15	1.86	0.04 (one-tailed)
β_X	0.00	0.00	−1.07	0.29 (two-tailed)

Table 3: Estimation results of the parameters $\beta_0, \beta_S, \beta_G$ for Model (2) (second column) with the corresponding standard deviations (third column) and t-test statistics (fourth column). The p-values (fifth column) are adjusted p-values in order to control the false discovery rate as suggested in Benjamini and Hochberg (1995).

Both null hypothesis can be rejected in favor of the alternative hypothesis (t-tests, adj. $p < .05$). This means, in tendency, FFF/male subjects are capable of producing higher ankle joint moments of their ankle plantarflexor muscles tendon units than RFF/female subjects with differences in the respective group-wise means by 0.28 Nm/kg and 0.32 Nm/kg; see Table 3.

The normality assumption cannot be rejected in either of the groups (Shapiro-Wilk tests, $p > .5$). Nevertheless, in order to take into account a possible miss judgment with respect to the

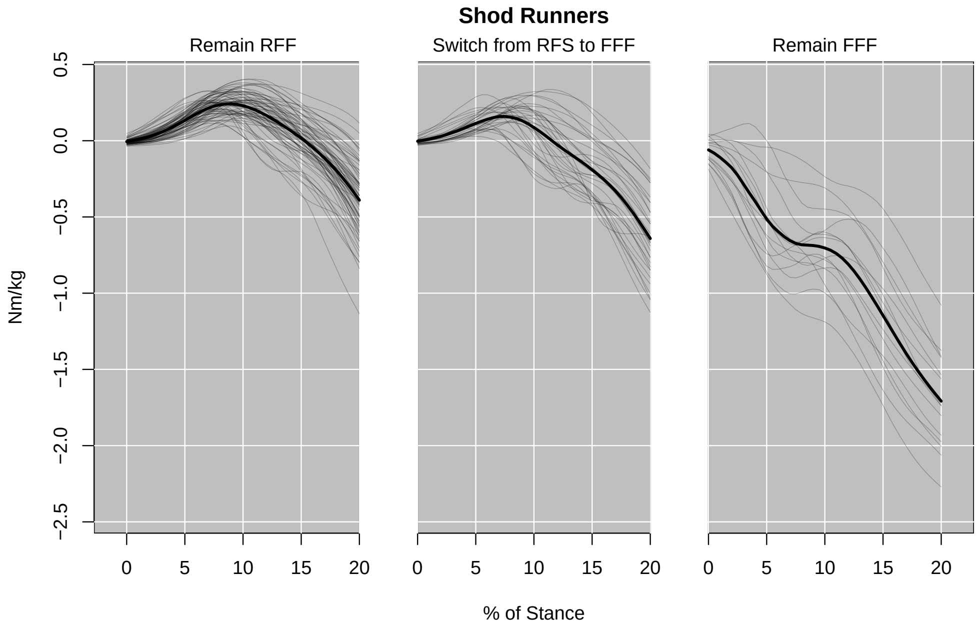

Figure 9: Left Panel: Joint moment curves of shod running subjects that use a rearfoot strike under the shod and barefoot treatment. Middle Panel: Joint moment curves of shod running subjects that use a rearfoot strike under the shod treatment, but a forefoot strike under the barefoot treatment. Right Panel: Joint moment curves of shod running subjects that use a forefoot strike under the shod and barefoot treatment.

normality assumption, we repeated the above analysis using a two-way ANCOVA model based on the ranks of the dependent variable Y and the co-variable X, and additionally conducted the test procedure suggested by Quade (1967). Both methodologies are robust against violations of the normality assumption; see Conover and Iman (1982). As this additional statistical analysis confirms our above results with respect to a significance level of 5%, we do not report any further details.

4. Discussion

The most interesting characteristic of our dataset is that the subjects were not instructed to perform a specific footfall pattern, but were allowed to run with their habitually preferred patterns. On the one hand, this situation brings the drawback of having to conduct an ex-post cluster analysis. On the other hand, it provides the unique possibility to analyze physical differences between RFF and FFF subjects. Both issues are discussed in this section.

Cluster analysis

The first of two purposes of the present study was to introduce a novel statistical procedure to find clusters in footfall patterns of human runners. As motivated in the introduction, the main problem with the 1/3-decision rule is its rigid geometrical specification, which generally does not take into account excising clusters in the data. This leads to truncated samples, which harm any statistical inference due to the so-called truncation bias. By contrast, the stochastic specification of the proposed cluster procedure allows for overlapping clusters and assigns each subject to the cluster with its highest probability.

The introduced procedure heavily relies on the functional version of principal component analysis, which takes into account the typical functional (or dynamic) nature of our biomechanical dataset. Although functional data typically arise in biomechanical studies, FPCA is still rarely used in the literature on biomechanics. There, most often used approach to dimension reduction is the selection of so-called characteristic features extracted from the functional data; see, for example, Vardaxis, Allard, Lachance, and Duhaime (1998), Wu, Wang, and Liu (2007), Lau, Tong, and Zhu (2009), and Saripalle, Paiva, Cliett III, Derakhshani, King, and Lovelace (2014).

We use the sagital joint moments at the ankle joint from a four segment inverse dynamics model in order to find clusters in human foot strike patterns. Nevertheless, the introduced cluster procedure can also be used with other kinds of functional biomechanical data—as long as the data contain useful information for the differentiation of footfall patterns. Promising examples are functional measurements from 3-D kinematic studies of the angle of the plantar surface or functional GRF data; see Williams et al. (2000). In fact, the ankle angle might indeed be less prone to inaccuracies of center of pressure calculation with low ground reaction forces. If more than one set of functional data is available, it is advisable to use the multivariate version of FPCA as proposed in Jacques and Preda (2014).

We consider only the first 20% of the stance phase, which roughly spans the time span of a typical foot landing. The smoothness of functional data implies that changes in the considered ranges of the domains of functional data will not cause abrupt changes in the statistical results. Correspondingly, our results essentially remain unchanged when considering the first 15%, 25%, or 30% of the stance phase instead of the first 20%.

There is a typical difference between the proposed FPCA based procedure and the more classical procedure based on characteristic features extracted from the functional data. For example, in the very left plot of Figure 7 there is one ankle joint moment function with an initial positive internal ankle moment. If, for example, the sign of the initial internal ankle plantar flexion moment would be used as a characteristic feature in order to differentiate between rearfoot and forefoot strikes, this particular joint moment function would have been classified to the rearfoot strike group. By contrast, the cluster analysis on basis of pc-score vectors classifies this joint moment function to the FFF cluster, since here the whole course is considered, such that single characteristics do not carry much weight. However, the relative uncertainty of this observation being correctly classified is 70 times higher than the corresponding uncertainty of all other observations in this cluster together.

The application of the novel cluster approach to our dataset of $n = 119$ subjects leads to two, well interpretable clusters with low intra-cluster variations; see Figure 7. The ex-post analysis of the two clusters validates clearly that we can refer to them as RFF and FFF clusters; see boxplots in Figure 6. However, it has to be emphasized that the RFF/FFF clusters and the RFS/MFS/RFS classes arise from different concepts. Yet, they are closely related with respect to their aims. In fact, the 90% frequency of RFF shod running subjects (see τ_R^S in Table 2) equals the 90% frequency of RFS shod running subjects published in Larson, Higgins, Kaminski, Decker, Preble, Lyons, McIntyre, and Normile (2011), who analyze a comparable set of recreational sub-elite runners.

Though, classification of the foot strikes based on the SI and its 1/3-decision rule generally produces a rather different, ambiguous results; see Figure 8. Particularly, the MFS class is very heterogeneous, such that the corresponding mean function cannot be regarded as a representative joint moment functions for MFS subjects. The latter issue is critical: it reflects the truncation bias, which demands for an appropriate statistical treatment; see, e.g., Cohen (1991). In the introduction we point to the contradictory classification result (SI-method) that 44 of 108 shod SI-classified RFS runners have initial plantarflexion moments. By contrast, after excluding subjects that have a relative uncertainty of being correctly classified above 10% (see also Figure 5), only 4 of 105 shod RFF runners have initial plantarflexion moments.

A possible drawback of the proposed cluster analysis based on functional data is that it is more involved than the SI and its 1/3-decision rule—although our step-by-step instructions in Section 2.3 shall provide guidance. Furthermore, we cannot find a third cluster of subjects with a midfoot footfall pattern. In fact, the choice of Gaussian mixture models with only two clusters on basis of the BIC is rather clear-cut; see Table 2. However, this result does not exclude the general existence of a separate, say, MFF cluster. It rather suggests that MFF patterns are not distinguishable from FFF patterns on basis of the proposed procedure.

Finally, a very interesting feature of our study is that the two clusters found under both treatments allow to determine how many subjects switch their footfall patterns from one treatment condition to the other. While the majority (87%) of all subjects use their preferred footfall pattern across the treatments, there is a non-negligible share of 13% of the subjects, which switch from RFF when running shod to FFF when running barefoot. Although the ankle joint moment functions of the "Remain RFF" subjects are rather similar to those of the "Switch from RFF to FFF" subjects, a test on the equality of the corresponding mean functions can be rejected at a significance level < 0.01. A further, deeper analysis of this group of subjects that are capable of switching their footfall patterns is not within the scope of this study, but will be of interest for further research.

Comparison study

As hypothesized, FFF subjects were capable to produce significantly higher plantarflexion torques than RFF subjects (difference in means: 0.28 Nm/kg). This means, the maximum voluntary contraction of the ankle plantarflexor muscles tendon units between representative RFF and FFF subjects (assuming masses of 75 kg) differs on average by 21 Nm. This effect remains significant after controlling for an additional gender effect and differences in training levels measured by the co-variable "weekly kilometers run".

Of course, from this result we cannot conclude whether the higher force capacities of FFF subjects are an adaptation to the higher mechanical demand (i.e., higher plantarflexion moments) that is placed upon their ankle joints, or if higher force capacities are a prerequisite for using a FFF pattern. Essentially, the latter research question coincides with the research question on whether the higher plantarflexion moments during FFF running are sub- or suprathreshold stimuli on the plantarflexion muscle tendon units. The higher plantarflexion moments during a single step might be considered as too small to induce an adaptation in force capacities of ankle plantarflexors.

Nonetheless, there are converse research findings for the case of minimal footwear, which tend to trigger a forefoot striking behavior. For example, Brüggemann, Potthast, Braunstein, and Niehoff (2005) showed that the use of minimal footwear in warm up running induces improvements in strength and anatomical cross sectional areas of triceps surae and extrinsic foot muscles. The latter considerations would suggest an interaction effect between the dummy variable "strike-type" and the co-variable "weekly km run". The hypothesis of a suprathreshold

FFF stimulus on the plantarflexion muscle tendon units would be supported, if the co-variable "weekly km run" has a positive effect on the MVC torques of plantarflexors in the case of FFF runners, but a non-significant effect in the case of a RFS runners.

The reported result that the rather simple Model (2) without interaction effects is not significantly worse in explaining the variations of the dependent variable Y, than the completely unrestricted ANCOVA model with interaction effects does not generally rule out the existence of interaction effects. Longitudinal studies are needed to resolve this issue in greater detail.

5. Conclusions

We introduce a novel statistical cluster procedure to find groups in habitually performed footfall patterns. The proposed procedure relies on the functional version of principal component analysis and tries to find systematic differences in the initial ankle joint moment functions. The found clusters are well interpretable as rearfoot footfall and forefoot footfall clusters and show low intra-cluster variations.

These RFF and FFF clusters are used to test our null hypothesis that subjects with habitual rearfoot footfall patterns and subjects with habitual forefoot footfall patterns equal in their ankle plantarflexor strength. This null hypothesis can be rejected at a 5% significance level in favor of the alternative that the mean ankle plantarflexor strength of FFF subjects is significantly higher. This effect remains significant after controlling for an additional gender effect and for differences in training levels.

Conflict of interest statement

We acknowledge that all authors do not have any conflict of interest and were fully involved in the study and preparation of the manuscript.

Acknowledgment

Parts of the study were financially supported by Brooks Sports Inc., Bothell, WA, USA. This sponsor had no influence on data analysis and interpretation. Further, the sponsor was not involved in the writing and submitting process of the manuscript. The authors are grateful to the reviewers for constructive comments and support.

Appendix A. Functional mixture model

We assume that the joint moment function $X_i(t)$ of a subject i is a smooth random function generated by a functional Gaussian mixture model with unknown group labels g, unknown mixture probabilities τ_g, and group-specific mean joint moment functions $\mu_g(t)$. The model can be written as

$$X_i(t) = \sum_{g=1}^{G} \tau_g \left(\mu_g(t) + Y_i(t, g)\right) \text{ with } t \in [0, 1],\ i \in \{1, \ldots, n\}, \tag{A.1}$$

where G is the number of groups. Here, the time interval $t \in [0, 1]$ is set to the unit-interval, which goes without loss of generality. The first term of the right hand side of Equation (A.1) models the group-specific mean functions μ_g and the second term the group-specific stochastic deviations $Y_i(t, g)$ from the mean functions with $E(Y_i(t, g)) = 0$ for all $t \in [0, 1]$ and $g \in \{1, \ldots, G\}$. Our assumption on the smoothness of the random functions $X_i(t)$ is formalized by requiring that the functions $X_i(t)$ are square integrable, i.e., $X_i(t) \in L^2[0,1]$.

Since we do not observe the group labels g, but only the mixture of all G groups we can just as well write Equation (1) in an aggregated version as following:

$$X_i(t) \quad = \quad \mu(t) + Y_i(t) \quad \text{with} \quad t \in [0,1], \tag{A.2}$$

where $\mu(t) = \sum_{g=1}^{G} \tau_g \, \mu_g(t)$ and $Y_i(t) = \sum_{g=1}^{G} \tau_g \, Y_i(t,g)$.

In the following it will be convenient to write the functional random variable $Y_i(t)$ using its Karhunen Loève decomposition

$$Y_i(t) \quad = \quad \sum_{k=1}^{\infty} \beta_{ik} f_k(t), \tag{A.3}$$

where the series of eigenfunctions functions $(f_k)_{k \in \{1,2,\dots\}}$ forms a complete orthonormal basis system, i.e., $\int_0^1 f_k(t) f_l(t)\, dt = 1$ for all $k = l$ and zero else, and the scores $\beta_{ik} = \int_0^1 Y_i(t) f_k(t) dt$ are univariate Gaussian random variables with $E(\beta_{ik}) = 0$ for all $k \in \{1,2,\dots\}$ and $E(\beta_{ik}^2) \to 0$ sufficiently fast as k goes to infinity.

The eigenfunctions $f_k(t)$ can be determined via the eigendecomposition of the covariance operator $\Gamma(t) = \int_0^1 \gamma(s,t) x(s) ds$ with the covariance function $\gamma(s,t) = E\left[(X_i(s) - \mu(s))\,(X_i(t) - \mu(t))\right]$. More detailed discussions can be found, e.g., in Chapter 8 of the textbook of Ramsay and Silverman (2005).

It is well known that the first K eigenfunctions $f_1(t), \dots, f_K(t)$ determine an orthonormal basis system, which allows for the best K-dimensional approximation of Equation (A.3), such that we can write without much loss of accuracy

$$Y_i(t) \quad = \quad \sum_{k=1}^{K} \beta_{ik} f_k(t) = (\beta_{i1}, \dots, \beta_{iK}) \begin{pmatrix} f_1(t) \\ \vdots \\ f_K(t) \end{pmatrix} \tag{A.4}$$

with $K < \infty$.

In practical applications, usually $K = 2$ or $K = 3$ is sufficient to achieve over 95% accuracy. We can use this K-dimensional system of eigenfunctions $f_1(t), \dots, f_K(t)$ in order to reduce the dimension of the (potentially infinite dimensional) joint moment functions $X_i(t)$. In other words, instead of working with the functions $X_i(t)$, we can analyze the corresponding pc-scores $\beta_{i1}, \dots, \beta_{iK}$ using standard multivariate methods with negligible loss of information.

Given our model assumptions in Equation (A.1) it can be seen easily that the $K \times 1$ dimensional pc-score vectors $\beta_i = (\beta_{i1}, \dots, \beta_{iK})'$ of Equation (A.4) can be written as a mixture of G group-specific pc-score vectors, i.e., $\beta_i = \sum_{g=1}^{G} \tau_g \beta_i^g$, where the group-specific pc-score vectors β_i^g are K-variate Gaussian random variables with $K \times 1$ dimensional mean vectors $E(\beta_i^g)$ and $K \times K$ dimensional covariance matrices $V(\beta_i^g)$.

This is exactly the situation, which is assumed for (multivariate) Gaussian mixture models, which can be easily estimated by the EM-algorithm (see, e.g., Fraley and Raftery (2002)). This means that once we have estimated the overall pc-score vectors β_i, we can cluster them using standard multivariate Gaussian mixture models.

Note that, although the vector of the individual pc-scores β_i is an ordinary K-variate random variable, it determines the specific *shape* of the joint moment function $X_i(t)$. Therefore, finding clusters in a sample of the K dimensional scores $\{\beta_1, \dots, \beta_n\}$ corresponds to finding shape-specific clusters in the sample of joint moment functions $X_i(t)$.

Of course, in practice the functions and integrals in Equations (A.1) to (A.4) have to be approximated by vectors and sums. The arising discretization error can be kept reasonably small by using a sufficiently dense grid of discretization points $0 \leq t_1 < \cdots < t_M \leq 1$. Hereby, "sufficiently dense" means to balance the following trade off: On the one hand side the number of discretization points M has to be large (often $M \approx n$) such that all important features of the functions $X_i(t)$ are cached by the discretization vectors. On the other hand, M has to be strictly smaller than the number of functional observations n in order to guarantee a stable principal components analysis. Otherwise, the dual estimation approach proposed by Benko, Härdle, and Kneip (2009) has to be used. Many further particularities of FPCA can be found in Chapter 8 of the textbook Ramsay and Silverman (2005).

References

Arampatzis, A., Karamanidis, K., & Albracht, K. (2007). Adaptational responses of the human achilles tendon by modulation of the applied cyclic strain magnitude. *Journal of Experimental Biology*, 210, 2743–2753.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple t esting. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.

Benko, M., Härdle, W., & Kneip, A. (2009). Common functional principal components. *The Annals of Statistics*, 37, 1–34.

Bezodis, N. E., Salo, A. I., & Trewartha, G. (2013). Excessive fluctuations in knee joint moments during early stance in sprinting are caused by digital filtering procedures. *Gait & Posture*, 38, 653–657.

Bisseling, R. W., & Hof, A. L. (2006). Handling of impact forces in inverse dynamics. *Journal of Biomechanics*, 39, 2438–2444.

Brüggemann, G. P., Potthast, W., Braunstein, B., & Niehoff, A. (2005). Effect of increased mechanical stimuli on foot muscles functional capacity. *In Proceedings of the ISB XXth Congress-ASB 29th Annual Meeting: 31 July-5 August 2005* (pp. 579–582).

Cavanagh, P., & Lafortune, M. (1980). Ground reaction forces in distance running. *Journal of Biomechanics*, 13, 397–406.

Cohen, A. C. (1991). *Truncated and Censored Samples: Theory and Applications*. CRC press.

Conover, W. J., & Iman, R. L. (1982). Analysis of covariance using the rank transformation. *Biometrics*, 38, 715–724.

Daoud, A., Geissler, G., Wang, F., Saretsky, J., Daoud, Y., & Lieberman, D. (2012). Foot strike and injury rates in endurance runners: a retrospective study. *Medicine and Science in Sports and Exercise*, 44, 1325–34.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*. Series B (Methodological), 39, 1–38.

Fraley, C., & Raftery, A. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97, 611–631.

Fraley, C., & Raftery, A. (2007). Model-based methods of classification: using the mclust software in chemometrics. *Journal of Statistical Software*, 18, 1–13.

Gruber, A. H., Umberger, B. R., Braun, B., & Hamill, J. (2013). Economy and rate of carbohydrate oxidation during running with rearfoot and forefoot strike patterns. *Journal of Applied Physiology*, 115, 194–201.

Gujarati, D. (1970). Use of dummy variables in testing for equality between sets of coefficients in linear regressions: a generalization. *The American Statistician*, 24, 18–22.

Hasegawa, H., Yamauchi, T., & Kraemer, W. J. (2007). Foot strike patterns of runners at the 15-km point during an elite-level half marathon. *Journal of Strength and Conditioning Research*, 21, 888–93.

Jacques, J., & Preda, C. (2014). Model-based clustering for multivariate functional data. *Computational Statistics and Data Analysis*, 71, 92–106.

Kristianslund, E., Krosshaug, T., & van den Bogert, A. J. (2012). Effect of low pass filtering on joint moments from inverse dynamics: Implications for injury prevention. *Journal of Biomechanics*, 45, 666–671.

Larson, P., Higgins, E., Kaminski, J., Decker, T., Preble, J., Lyons, D., McIntyre, K., & Normile, A. (2011). Foot strike patterns of recreational and sub-elite runners in a long-distance road race. *Journal of Sports Sciences*, 29, 1665–1673.

Lau, H.-y., Tong, K.-y., & Zhu, H. (2009). Support vector machine for classification of walking conditions of persons after stroke with dropped foot. *Human Movement Science*, 28, 504–514.

Lieberman, D. (2012). What we can learn about running from barefoot running: an evolutionary medical perspective. *Exercise and Sport Sciences Reviews*, 40, 63–72.

Maclean, C. L., Davis, I. S., & Hamill, J. (2009). Influence of running shoe midsole composition and custom foot orthotic intervention on lower extremity dynamics during running. *Journal of Applied Biomechanics*, 25, 54–63.

Quade, D. (1967). Rank analysis of covariance. *Journal of the American Statistical Association*, 62, 1187–1200.

R Core Team (2013). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org/`

Ramsay, J. O., & Silverman, B. W. (2005). *Functional Data Analysis*. (2nd ed.). Springer Series in Statistics. New York: Springer.

Saripalle, S. K., Paiva, G. C., Cliett III, T. C., Derakhshani, R. R., King, G. W., & Lovelace, C. T. (2014). Classification of body movements based on posturographic data. *Human Movement Science*, 33, 238–250.

Vardaxis, V. G., Allard, P., Lachance, R., & Duhaime, M. (1998). Classification of able-bodied gait using 3-d muscle powers. *Human Movement Science*, 17, 121–136.

Williams, D., McClay, I. S., & Manal, K. T. (2000). Lower extremity mechanics in runners with a converted forefoot strike pattern. *Journal of Applied Biomechanics*, 16, 210–218.

Willwacher, S., Fischer, K., & Brüggemann, G. P. (2011). Is average speed control sensitive enough to ensure nonaccelerated running in the analysis of distance running mechanics? *In ISBS-Conference Proceedings* (pp. 579–582).

Wu, J., Wang, J., & Liu, L. (2007). Feature extraction via KPCA for classification of gait patterns. *Human Movement Science*, 26, 393–411.