MOX-Report No. 30/2016

Nonparametric inference for functional-on-scalar linear models applied to knee kinematic hop data after injury of the anterior cruciate ligament

Abramowicz, K.; Häger, C.; Pini, A.; Schelin, L.; Sjöstedt de Luna, S.; Vantini, S.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

Nonparametric inference for functional-on-scalar linear models applied to knee kinematic hop data after injury of the anterior cruciate ligament

Konrad Abramowicz[a], Charlotte Häger[b], Alessia Pini[c], Lina Schelin[d,b], Sara Sjöstedt de Luna[a] and Simone Vantini[c]

[a] Department of Mathematical Statistics, Umeå University, Umeå Sweden

konrad.abramowicz@umu.se

sara.sjostedt.de.luna@umu.se

[b] Department of Community Medicine and Rehabilitation, Physiotherapy, Umeå University, Umeå Sweden

charlotte.hager@umu.se

[c] MOX - Department of Mathematics, Politecnico di Milano, Milano Italy

alessia.pini@polimi.it

simone.vantini@polimi.it

[d] Department of Statistics, Umeå School of Business and Economics, Umeå University, Umeå Sweden

lina.schelin@umu.se

Abstract

Motivated by the analysis of the dependence of knee movement patterns during functional tasks on subject-specific covariates, we introduce a distribution-free procedure for testing a functional-on-scalar linear model with fixed effects. The procedure does not only test the global hypothesis on all the domain, but also selects the intervals where statistically significant effects are detected. We prove that the proposed tests are provided with an asymptotic control of the interval-wise error rate, i.e., the probability of falsely rejecting any interval of true null hypotheses. The procedure is applied to one-leg hop data from a study on anterior cruciate ligament injury. We compare knee kinematics of three groups of individuals (two injured groups with different treatments, and one group of healthy controls), taking individual-specific covariates into account.

Keywords: ANCOVA; Functional data; Human movement; Interval-Wise Testing; Permutation test.

1 Introduction

Functional data analysis (FDA) is a dynamically developing research area within the field of statistics. In recent literature, linear models for functional data have been widely studied (see, e.g., Fan and Zhang 2000; Abramovich and Angelini 2006; Cardot et al. 2007; Reiss et al. 2010; Gertheiss et al. 2013; Fogarty and Small 2014; Abramowicz et al. 2014).

Motivated by the analysis of the dependence of knee movement patterns during functional tasks on subject-specific covariates, we consider in the present paper a functional-on-scalar linear model. Specifically, we model a functional response with a set of covariates multiplied by functional parameters. Such models find their applications in a wide range of research fields where modern techniques enable collection of high-resolution time-continuous data. In this context, many of the empirically relevant questions address the effect of covariates on a functional response. They may also desire identification of significant domain subsets, that is, time intervals characterized by significant effects of a specific covariate. The incitament for this work comes from a follow-up study after anterior cruciate ligament injury. Anterior cruciate ligament injuries are common worldwide, especially in sports, and are typically treated either conservatively with physiotherapy or with reconstructive surgery in combination with physiotherapy. We analyze knee-joint kinematics data of sagittal plane, i.e., knee flexion/extension, during a one-leg hop for distance for $n = 95$ individuals (see Figure 1). We compare individuals from the surgery and physiotherapy groups with healthy-knee controls matched on age and sex.

Figure 1 here

Previous studies are indicating differences in the movement patterns between these groups (Tengman et al., 2015; Hébert-Losier et al., 2015). Tengman et al. (2015) has a limitation since it is only considering selected landmarks (events) of the curves. The previous analysis included group effects as well as covariates, but did not take into account the information coming from the functional nature of the data. In Hébert-Losier et al. (2015), the complete functional data was considered, but without taking covariates into account. Both these studies indicate less knee flexion among the individuals treated with physiotherapy compared to the control group. We here overcome both the above mentioned limitations by introducing a statistical tool that both exploits the functional nature of the data, and takes into account possible covariates effects. In this paper, we investigate if these differences are only due to group effects, or if they can be explained by means of additional individual-specific covariates. Further, the introduced methodology enables detection of the intervals where the covariates

have significant effects (domain selection). Such information provides additional insight into the importance of different time segments of the movement. For instance, the active preparation phase with e.g., knee bending, prior to the actual take-off of the jump determines much of the performance, but there is no obvious single event that would necessarily be most representative for comparison of movement control for individuals or groups. Likewise, the control of the knee in the landing phase is essential for maintaining balance during the task but it is not known to a full extent how this is controlled or even how to best assess it. The present method enables comparisons of the whole movement pattern tailored also to individuals and may in addition help in identifying critical intervals within the larger phases of the movement execution.

Parameter estimation of the functional model is handled by least squares estimation, as suggested, for instance, by Ramsay and Silverman (2005). Forming valid tests of various hypotheses about the functional regression parameters, with control of the error rate, is however not straightforward. One solution adopted in the literature is to develop global tests for the parameters of the model (Cuevas et al., 2004; Abramovich and Angelini, 2006; Antoniadis and Sapatinas, 2007; Cardot et al., 2007; Schott, 2007; Cuesta-Albertos and Febrero-Bande, 2010; Staicu et al., 2014; Zhang and Liang, 2014; Kayano et al., 2015). Such tests investigate if a covariate has a significant effect on the response, but does not provide any domain selection. Another approach, proposed in Fan and Zhang (2000); Reiss et al. (2010); Ramsay and Silverman (2005), is to provide point-wise confidence bands for the functional parameters. The results indicate in which parts of the domain that the covariates have an effect, with only a point-wise control of the type I error rate. As clearly discussed in Ramsay and Silverman (2005, pp. 243–244), point-wise limits are not equivalent to confidence regions for the entire estimated curves.

Assuming that data are expressed through a functional basis, inference can be based directly on the expansion coefficients, as proposed, for instance, by Spitzner et al. (2003) and Pini and Vantini (2016). In the latter work, single-component tests are performed, and their p-values adjusted with multiple testing methods. In this way, results are compensated for the many dependent tests performed on the same data set. A drawback with this type of procedures is that they rely on the basis expansion. The test conclusions thus depend on the type of functional basis selected to initialize the methods.

Vsevolozhskaya et al. (2013) and Vsevolozhskaya et al. (2014) proposed an inferential procedure in the functional ANOVA framework based on partitioning the domain into pre-selected sub-intervals. The procedure, relying on permutation tests, results in a selection of sub-intervals where significant differences are observed between groups. However, the test conclusions depend on the initially

chosen partition and do not take into account covariates.

In the framework of testing differences between two-populations, Pini and Vantini (2015) proposed the interval-wise testing (IWT) procedure that tests functional hypotheses and selects intervals of the domain where the null hypothesis is rejected. The method relies on the definition of an adjusted p-value function provided with a control of the interval-wise error rate (IWER), i.e., control of the probability of wrongly rejecting any interval of the domain. In our work, we extend this method to the functional-on-scalar linear model to test various hypotheses on the functional coefficients, and prove that the proposed tests are exact or asymptotically exact. It is important to point out, as tests are based on permutations, that the procedure does not require normality of the residual functions. Moreover, the resulting procedure does not require to specify the covariance structure of the data.

The paper is outlined as follows: in Section 2, we describe the functional-on-scalar linear model, discussing the methodology proposed for functional parameter estimation and inference. Section 3 reports the theoretical properties of the proposed methodology. The proofs of theorems in Section 3 are reported Section A of the Appendix. In Section 4 we provide the implementation details of the proposed procedure. In Section 5, we report the results of the analysis of the kinematics data. Finally, Section B of the Appendix reports some algorithmic details on the permutation scheme used to perform inference, and Section C of the Appendix reports supplementary results on the kinematics data. All computations and plots have been created using R (R Core Team, 2014).

2 Methodology

2.1 The functional-on-scalar linear model

Suppose we have observed a sample of n continuous squared-integrable random functions $y_i(t)$, $t \in [a, b]$, $i = 1, \ldots, n$, i.e., $y_i(\cdot) \in L^2[a, b] \cap C^0[a, b]$ $\forall i$. We want to study the following functional-on-scalar linear model:

$$y_i(t) = \beta_0(t) + \sum_{l=1}^{L} \beta_l(t) x_{li} + \varepsilon_i(t), \quad \forall i = 1, ..., n,\ t \in [a, b], \tag{2.1}$$

where $x_{1i}, ..., x_{Li} \in \mathbb{R}$ are known scalar covariates and $\beta_l(t)$, $l = 0, ..., L$, are the unknown fixed functional regression parameters. The errors $\varepsilon_i(t)$, $t \in [a, b]$ are independent and identically distributed (with respect to units) zero-mean

random functions (not necessarily Gaussian) with finite total variance, i.e.,

$$\int_a^b \mathbb{E}\left[\varepsilon_i(t)\right]^2 dt < \infty, \quad \forall i = 1, ..., n. \tag{2.2}$$

No further assumption is needed on the covariance structure of the error term.

2.2 Model estimation

The ordinary least squares (OLS) estimators of the functional parameters $\beta_l(t)$, $l = 0, \ldots, L$, can be found by minimizing the sum over units of the squared L^2 distances between the functional data $y_i(t)$ and the quantity $\beta_0(t)+\sum_{l=1}^{L} \beta_l(t)x_{li}$ with respect to $\beta_l(t)$, $l = 0, \ldots, L$ (Ramsay and Silverman, 2005), hence minimizing

$$\sum_{i=1}^{n} \int_a^b \left(y_i(t) - \beta_0(t) - \sum_{l=1}^{L} \beta_l(t) x_{li} \right)^2 \mathrm{d}t. \tag{2.3}$$

From the interchangeability of summation and integration in (2.3) it follows that the minimization can be done separately for each point of the domain, independently of the covariance structure of the errors $\varepsilon_i(t)$. The OLS estimate $\hat{\boldsymbol{\beta}}(t) = (\hat{\beta_0}(t), \ldots, \hat{\beta_L}(t))'$, $t \in [a, b]$, minimizing (2.3) can thus be calculated point-wise for each given t as:

$$\hat{\boldsymbol{\beta}}(t) = \underset{\beta_0(t),\ldots,\beta_L(t)}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i(t) - \beta_0(t) - \sum_{l=1}^{L} \beta_l(t) x_{li} \right)^2. \tag{2.4}$$

For each t, $\hat{\boldsymbol{\beta}}(t)$ is thus the OLS estimator of the corresponding scalar-on-scalar linear model at point t.

Asymptotic properties for the OLS estimates at each point t can be immediately derived from classical results for scalar-on-scalar linear models. In detail, let X_n be the design matrix $X_n \in \mathbb{R}^{(n\times(L+1))}$ ($[X_n]_{i,1} = 1$, $\forall i = 1, ..., n$; $[X_n]_{i,j} = x_{j-1,i}$, $i = 1, ..., n$, $j = 2, ..., L + 1$). Consider the following standard conditions:

C1 The matrix $X_n'X_n$ is non-singular, and the inverse $V = (X_n'X_n)^{-1}$ is s.t. the elements $[V]_{ij} \to 0$ as $n \to \infty$, for all $i, j = 1, ..., L + 1$.

C2 For each $t \in [a, b]$, the regression errors $\varepsilon_i(t)$ satisfy:

$$\sup_{i=1,\ldots,n} \mathbb{E}\left[\varepsilon_i(t)^2\right] < \infty.$$

Under conditions *C1*-*C2*, we have that for each $t \in [a, b]$, the OLS estimate

$$\hat{\boldsymbol{\beta}}(t) = (X_n' X_n)^{-1} X_n' (y_1(t), \ldots, y_n(t))', \tag{2.5}$$

is a strongly consistent estimate of $\boldsymbol{\beta}(t) = (\beta_0(t), \ldots, \beta_L(t))'$ (Lai et al., 1979). Condition *C1* is a sufficient condition for finding an explicit expression of the OLS estimates, and guarantees convergence in probability. Condition *C2* ensures almost sure convergence.

2.3 Model inference

One of the main challenges with inference for functional linear models (2.1) is to perform valid tests of various hypotheses on the functional regression parameters, and to select significant intervals of the domain. For instance, we are interested in the overall model test, that none of the covariates have a significant effect on the response, i.e., the functional version of the classical F-test hypotheses:

$$\begin{cases} H_{0,F} : \beta_l(t) = 0 \quad \forall l \in 1, \ldots, L, \ \forall t \in [a, b] \\ H_{1,F} : \beta_l(t) \neq 0 \quad \text{for some } l \in \{1, \ldots, L\} \text{ and some } t \in [a, b], \end{cases} \tag{2.6}$$

together with tests on single functional parameters, i.e., the functional version of the classical t-test hypotheses:

$$\begin{cases} H_{0,l} : \beta_l(t) = 0 \quad \forall t \in [a, b] \\ H_{1,l} : \beta_l(t) \neq 0 \quad \text{for some } t \in [a, b], \end{cases} \tag{2.7}$$

where $l \in \{0, \ldots, L\}$.

It may also be of interest to test hypotheses on one or more linear combinations of the functional parameters of the regression, specified by a combination matrix C. In detail, let $C \in \mathbb{R}^{(q \times (L+1))}$ be any real-valued full rank matrix, where q denotes the number of hypotheses on the functional regression parameters to be jointly tested, with $1 \leq q \leq L + 1$. Moreover, let $\mathbf{c}_0(t) = (c_{01}(t), ..., c_{0q}(t))'$ be a vector of fixed functions in $L^2[a, b] \cap C^0[a, b]$. The general testing problem can typically then be formulated as follows:

$$\begin{cases} H_{0,C} : C\boldsymbol{\beta}(t) = \mathbf{c}_0(t) \quad \forall t \in [a, b] \\ H_{1,C} : C\boldsymbol{\beta}(t) \neq \mathbf{c}_0(t) \quad \text{for some } t \in [a, b], \end{cases} \tag{2.8}$$

where the j-th element of vector $C\boldsymbol{\beta}(t)$ is a function obtained by means of a linear combination of the functional regression parameters $\beta_l(t)$ with weights $[C]_{jl}$: $[C\boldsymbol{\beta}(\boldsymbol{t})]_j = \sum_{l=0}^{L} [C]_{jl} \beta_l(t)$, $j = 1, ..., q$. There are two important special

cases of the general functional linear hypotheses (2.8):

1. Let $q = L$, $C = C_F = (\mathbf{0}|I_L) \in \mathbb{R}^{L\times(L+1)}$, and $\mathbf{c_0(t)} = \mathbf{0} \in \mathbb{R}^L$, where I_L is the $L \times L$ identity matrix. Then, the hypotheses of (2.8) correspond to the hypotheses in (2.6);

2. For a fixed l, let $q = 1$, $C = C_l \in \mathbb{R}^{1\times(L+1)}$ with $[C_l]_r = 1$ if $r = l$ and 0 otherwise, and $c(t) = 0$. Then, the hypotheses in (2.8) correspond to the hypotheses in (2.7).

For test (2.8), in case of rejection of the null hypothesis $H_{0,C}$, we want to select the intervals on $[a, b]$ where significant differences are detected. In theory, this problem can be addressed by performing an infinite family of tests along the domain $[a, b]$, of the form:

$$\begin{cases} H^t_{0,C} : C\boldsymbol{\beta}(t) = \mathbf{c}_0(t) \\ H^t_{1,C} : C\boldsymbol{\beta}(t) \neq \mathbf{c}_0(t). \end{cases} \tag{2.9}$$

Based on classical results for scalar-on-scalar linear models, it is rather straightforward to test hypotheses (2.9) for each t. The challenge is to control the family-wise error rate arising from the uncountable infinite number of (dependent) hypotheses tests. In this paper, we extend the domain selection IWT procedure by Pini and Vantini (2015) to functional-on-scalar linear models to control the probability of wrongly rejecting each interval characterized by only true hypotheses, i.e., the control of the interval-wise error rate. The domain selection procedure we propose is based on three steps, presented in detail in the following paragraphs.

2.3.1 First step: interval-wise testing

Given any closed interval $\mathcal{I} \subseteq [a, b]$, we want to test:

$$\begin{cases} H^{\mathcal{I}}_{0,C} : C\boldsymbol{\beta}(t) = \mathbf{c}_0(t) \quad \forall t \in \mathcal{I} \\ H^{\mathcal{I}}_{1,C} : C\boldsymbol{\beta}(t) \neq \mathbf{c}_0(t) \quad \text{for some } t \in \mathcal{I}. \end{cases} \tag{2.10}$$

To perform tests of linear hypotheses (2.10), we propose to use the statistic:

$$T^{\mathcal{I}}_C = \int_{\mathcal{I}} T_C(t)\mathrm{d}t, \tag{2.11}$$

where

$$T_C(t) = \left(C\hat{\boldsymbol{\beta}}(t) - \mathbf{c}_0(t)\right)' \left(C\hat{\boldsymbol{\beta}}(t) - \mathbf{c}_0(t)\right), \tag{2.12}$$

and $\hat{\boldsymbol{\beta}}(t)$ is the OLS estimate (2.5). In particular, for the overall model hypotheses on $\mathcal{I}$

$$\begin{cases} H_{0,F}^{\mathcal{I}} : \beta_l(t) = 0 \quad \forall l \in 1, \ldots, L, \forall t \in \mathcal{I} \\ H_{1,F}^{\mathcal{I}} : \beta_l(t) \neq 0 \quad \text{for some } l \in \{1, \ldots, L\} \text{ and } t \in \mathcal{I} \end{cases} \tag{2.13}$$

we use the following statistic:

$$T_F^{\mathcal{I}} = \int_{\mathcal{I}} \sum_{l=1}^{L} (\hat{\beta}_l(t))^2 \mathrm{d}t. \tag{2.14}$$

For the hypotheses of the lth functional regression parameter on $\mathcal{I}$,

$$\begin{cases} H_{0,l}^{\mathcal{I}} : \beta_l(t) = 0 \quad \forall t \in \mathcal{I} \\ H_{1,l}^{\mathcal{I}} : \beta_l(t) \neq 0 \quad \text{for some } t \in \mathcal{I} \end{cases} \tag{2.15}$$

we use the test statistic:

$$T_l^{\mathcal{I}} = \int_{\mathcal{I}} \left(\hat{\beta}_l(t) \right)^2 \mathrm{d}t. \tag{2.16}$$

Note that we have chosen $T_F^{\mathcal{I}}$ and $T_l^{\mathcal{I}}$ to be special cases of $T_C^{\mathcal{I}}$.

2.3.2 Second step: adjusted p-value functions

Let $p_C^{\mathcal{I}}$ denote the p-value of test (2.10) obtained using functional permutation tests based on the Freedman and Lane permutation scheme (Freedman and Lane, 1983), as defined by Pesarin and Salmaso (2010). It is based on the permutation of the estimated residuals under the reduced model, i.e., the linear model constrained to the null hypothesis of the test. The p-value of the permutation test is then obtained by calculating the proportion of permutations leading to a larger value of the test statistic than the test statistic on the observed data. This scheme is the most commonly used scheme for linear models, and presents many advantages compared to other permutation techniques (Davison and Hinkley, 1997; Anderson and Legendre, 1999; Anderson and Robinson, 2001; Zeng et al., 2011; Winkler et al., 2014). In particular, it can be shown empirically that its power is typically higher than the power of tests based on other permutation schemes (Anderson and Legendre, 1999; Winkler et al., 2014). For more details, see Appendix B.

In order to identify significant sub-domains we make use of adjusted p-value functions. The adjusted p-value function $\tilde{p}_C(t)$ at point t for testing general linear hypotheses with contrast C is defined as the supremum p-value of all interval-wise tests on intervals containing t:

$$\tilde{p}_C(t) = \sup_{\mathcal{I} \ni t} p_C^{\mathcal{I}}, \quad t \in [a, b]. \tag{2.17}$$

Analogously, denoting by $p_F^{\mathcal{I}}$, and $p_l^{\mathcal{I}}$ the p-values from testing (2.13) and (2.15), respectively, the adjusted p-value functions for testing hypotheses on the overall model and on the lth functional parameter are defined as

$$\tilde{p}_F(t) = \sup_{\mathcal{I} \ni t} p_F^{\mathcal{I}}; \quad \tilde{p}_l(t) = \sup_{\mathcal{I} \ni t} p_l^{\mathcal{I}}, \quad t \in [a, b], \tag{2.18}$$

respectively.

Due to the nature of permutation tests used here, the adjusted p-value functions are quantized continuous functions with step size decreasing as the sample size n tends to infinity. The continuity of the limiting function is guaranteed by the continuity of the point-wise test statistics and of the observed functions $y_i(t)$'s.

2.3.3 Third step: domain selection

The intervals of the domain presenting a rejection of any of the null hypotheses are obtained by thresholding the corresponding adjusted p-value functions at level α. For example, we select intervals presenting at least one significant effect by thresholding $\tilde{p}_F(t)$ and intervals presenting a significant effect of the lth covariate by thresholding $\tilde{p}_l(t)$.

The introduced domain selection procedure is provided with a (asymptotic) control of the IWER. This type of control implies that the probability of detecting false positive intervals is (asymptotically) controlled at level α. The supporting theoretical results are presented in the following section.

3 Theoretical results

Here we present theoretical properties of the permutation-based inference on functional-on-scalar linear models performed along the line depicted in Section 2. All proofs are reported in Appendix A.

First, we prove that the domain selection IWT procedure for testing general linear hypothesis is provided with an asymptotic control of the IWER. Pini and Vantini (2015) proved that, if all interval-wise tests used to build IWT are exact, the IWT is provided with a control of the IWER. This result can be applied directly in the case of the overall functional model test on the regression model, but has to be extended in the more general case of tests on linear hypotheses (including tests on single functional parameters), since in the latter case the exactness of all tests is only asymptotical.

Theorem 3.1. *Under assumptions (C1-C2), the domain selection procedure for testing general functional linear hypotheses is provided with an asymptotic*

control of the IWER. Formally, the adjusted p-value function $\tilde{p}_C(t)$ *is s.t.,* $\forall \alpha \in (0,1]$*:*

$$\forall \mathcal{I} \subseteq [a,b] : H_{0,C}^{\mathcal{I}} \text{ is true } \Rightarrow \limsup_{n\to\infty} \mathbb{P}\left[\forall t \in \mathcal{I}, \tilde{p}_C(t) \leq \alpha\right] \leq \alpha.$$

Since tests on single functional parameters are specific cases of linear hypotheses, we obtain directly the following corollary.

Corollary 3.2. *Under assumptions (C1-C2), the domain selection procedure for testing hypotheses for single functional regression parameter is provided with an asymptotic control of the IWER.*

Furthermore, the following proposition provides exact results for IWT-based overall functional model tests.

Proposition 3.3. *The domain selection procedure for testing the overall functional model hypotheses is provided with a control of the IWER. Formally, the adjusted p-value function* $\tilde{p}_F(t)$ *is s.t.,* $\forall \alpha \in (0,1]$*:*

$$\forall \mathcal{I} \subseteq [a,b] : H_{0,F}^{\mathcal{I}} \text{ is true } \Rightarrow \mathbb{P}\left[\forall t \in \mathcal{I}, \tilde{p}_F(t) \leq \alpha\right] \leq \alpha.$$

Next, we focus on the property of consistency of the proposed tests. The following theorem states that the probability of detecting every interval $\mathcal{I}$ s.t. $H_{0,C}^t$ is false $\forall t \in \mathcal{I}$, converges to 1 as the sample size increases. This property implies that the probability of truly detecting any point where the null hypothesis is false converges to 1.

Theorem 3.4. *The domain selection procedure for testing general functional linear hypotheses is consistent. Formally,* $\forall \alpha \in (0,1]$*:*

$$\forall \mathcal{I} \subseteq [a,b] \text{ s.t. } H_{0,C}^t \text{ is false } \forall t \in \mathcal{I} \Rightarrow \lim_{n\to\infty} \mathbb{P}\left[\forall t \in \mathcal{I}, \tilde{p}_C(t) \leq \alpha\right] = 1.$$

As a consequence, we also obtain consistency results for the IWT-based tests of the overall functional model and on single functional parameters.

Corollary 3.5. *The domain selection procedure for the overall functional model hypotheses is consistent.*

Corollary 3.6. *The domain selection procedure for single functional parameter hypotheses is consistent.*

4 Details on the implementation

In practice, evaluations of the adjusted p-value functions, even at one fixed point $t \in [a,b]$, is practically unfeasible, since it involves taking supremum over

uncountably many sub-intervals covering point t. We therefore propose a numerical procedure resulting in point-wise constant approximation of the p-value functions and the significance regions chosen accordingly. The approximation error of the procedure is dependent on two parameters. The first parameter, $K \in \mathbb{N}$, defines the size of the initial partition of interval $[a, b]$. The second parameter, $B \in \mathbb{N}$, determines the number of random permutations used in the conditional Monte Carlo algorithm (Pesarin and Salmaso, 2010) to approximate the interval-wise permutation tests p-values.

The following steps describe the procedure for linear hypotheses of the functional parameters with combination matrix C.

Step 1 Partition the domain $[a, b]$ into K equally sized atomic sub-intervals of length $\Delta t = (b - a)/K$, i.e.,

$$\mathcal{P}_i = [a + (i-1)\Delta t, a + i\Delta t] \qquad i = 1, \ldots, K.$$

Let $\mathcal{S}$ be the family (of size $K(K+1)/2$) of all possible *intervals* constructed from the K atomic sub-intervals.

Step 2 For $i = 1, \ldots, K$ approximate the value of the integrated test statistics $T_C^{\mathcal{P}_i}$ by

$$\widehat{T_C^{\mathcal{P}_i}} = T_C\left(a + (i-1)\Delta t\right)\Delta t.$$

Step 3 For all intervals $\mathcal{J} \in \mathcal{S}$ approximate the value of the integrated test statistics by

$$\widehat{T_C^{\mathcal{J}}} = \sum_{\mathcal{P}_i \subset \mathcal{J}} \widehat{T_C^{\mathcal{P}_i}}.$$

The approximations of the integrated test statistics correspond to a rectangle quadrature with step Δt applied to the point-wise test statistic $T_C(\cdot)$.

Step 4 Estimate the p-values $\widehat{p_C^{\mathcal{J}}}$ of the tests on each $\mathcal{J} \in \mathcal{S}$ using the Freedman and Lane permutation scheme (see Appendix B) with B randomly chosen permutations.

Step 5 The adjusted p-value function at point t is estimated by the maximum of the corresponding p-values of all intervals $\mathcal{J}$ in $\mathcal{S}$ containing t, i.e.,

$$\widehat{\tilde{p}_C}(t) = \max\left\{\widehat{p_C^{\mathcal{J}}} : \mathcal{J} \in \mathcal{S} \text{ s.t. } \mathcal{J} \ni t\right\}.$$

Further, by construction, $\widehat{\tilde{p}_C}(t)$ is constant at each atomic sub-interval and therefore, the estimate of $\tilde{p}_C(t)$ for all $t \in [a, b]$ is obtained by evaluating the above maximum K times (one time per each $\mathcal{P}_1, \ldots, \mathcal{P}_K$).

Step 6 Select the significant domain subset by thresholding the obtained approximated adjusted p-value function.

Observe that as K tends to infinity the error arising due to discretization of the domain becomes negligible. This is guaranteed by the integrability of the point-wise test statistic $T_C(t)$ and resulting continuity of the integrated test statistic with respect to the integration limits. Further, as the number of simulated permutations B increases, the Monte Carlo error tends to zero. It is worth to notice that in Step 4, we use the same B permutations for each sub-interval, which decreases the computational complexity of the algorithm.

5 Analysis of knee kinematics

We will now report the results of the analysis of the knee kinematics data briefly presented in the introduction. To investigate potential long term differences in movement patterns following ACL injury, individuals treated either with reconstructive surgery in combination with physiotherapy or physiotherapy alone were tested in a motion laboratory. The resulting functional data consist of the joint angle motion in different movement planes during a one-leg hop, sampled at 240 Hz.

The functional data set that we analyze here corresponds to the knee movement in the sagittal plane, i.e., knee flexion/extension, see Figure 1, during a one-leg hop for distance. Our primary focus is to compare individuals treated either with reconstructive surgery in combination with physiotherapy (ACL_R) or physiotherapy alone (ACL_PT) with healthy-knee controls (CTRL) matched on age and sex, taking individual-specific covariates into account. The three groups are only matched on group level (Hébert-Losier et al., 2015), and hence age and sex are included as possible covariates together with available data about jump length and body mass index (BMI). The knee movement in the sagittal plane for the hop of maximal length (out of three attempts) on the injured leg for the two ACL groups was compared to the non-dominant leg for controls. In the case of group differences, it is of clinical interest to find out in which time intervals of the jump these differences occur. The functional data are represented through B-splines and aligned by means of four landmarks: the event of maximal knee flexion before take-off; the take-off event (the instant at which the foot leaves the ground); the touch-down event (the time instant at which the foot touches the ground); and the event of maximal knee flexion after landing. The aligned data and the landmarks are displayed in the right panel of Figure 1. For more details on data collection and preprocessing we refer to Hébert-Losier et al. (2015). The hop consist of three phases; time preceding the

Table 1: Global p-values of the *initial model* for the overall test (2.6) and for each single functional parameter test (2.7) obtained by performing permutation tests with B=1000 based on test statistics (2.14) and (2.16), respectively.

Test	**Overall**	β_0	β_{Jump}	β_{BMI}	β_{Sex}	β_{Age}	β_{CTRL}	β_R
p-value	0.000	0.000	0.000	0.302	0.277	0.327	0.000	0.037

Table 2: Global p-values of the reduced *final model* for the overall test (2.6) and for each single functional parameter test (2.7) obtained by performing permutation tests with B=1000 based on test statistics (2.14) and (2.16), respectively.

Test	**Overall**	β_0	β_{Jump}	β_{CTRL}	β_R
p-value	0.000	0.000	0.000	0.000	0.029

take-off event (take-off phase); time between the take-off event and the landing event (flight phase); time succeeding the landing event (landing phase).

The initial model we use to describe the knee joint angle motion in the sagittal plane, $y_i(t)$, is the following:

$$y_i(t) = \beta_0(t) + \beta_{Jump}(t)x_{Jump,i} + \beta_{BMI}(t)x_{BMI,i} + \beta_{Age}(t)x_{Age,i} + \beta_{Sex}(t)x_{Sex,i} + \beta_{CTRL}(t)x_{CTRL,i} + \beta_R(t)x_{R,i} + \varepsilon_i(t), i = 1, ..., 95,$$

where the covariates $x_{Jump,i}$, $x_{BMI,i}$, and $x_{Age,i}$ indicate the jump length, BMI, and age, respectively of individual i. Furthermore, $x_{Sex,i}$, $x_{CTRL,i}$ and $x_{R,i}$, are indicator functions attaining the value 1 if individual i is a man, in the control group and in the reconstructive surgery group, respectively, otherwise 0.

As in classical multiple regression, we start with the initial model using all available covariates, and then use backward elimination to reduce it to a final model which includes only the significant covariates. We start by performing a global overall significance test, that simultaneously tests if at least one of the coefficients is significant somewhere on the domain, c.f. (2.6). Separate global tests are then performed for each coefficient as in eq. (2.7). Table 1 presents the p-values obtained by performing permutation tests with $B = 1000$ and test statistics (2.14) and (2.16), where the interval $\mathcal{I}$ corresponds to the whole domain. Starting from these results, we apply the backward elimination procedure and stepwise remove the covariates with the largest p-value (one at a time, and reestimate the reduced model) until only the significant coefficients remain in the model. We use 5% significance level at all steps of the procedure. The final model includes the continuous covariate Jump and the group indicators. The corresponding p-values are presented in Table 2.

After assessing the overall significance of the covariates on the whole do-

main, we apply the introduced IWT-based procedure, to study in which parts of the domain the coefficients are significant. To illuminate all between-group differences, we use suitable combination matrices and introduce an additional hypothesis comparing the performance of the control and reconstructive surgery groups, see below. The adjusted p-value functions are computed according to Section 4 with $K = 300$, and $B = 1000$.

Functional parameter estimates together with estimated adjusted p-value functions for the overall functional model and single functional parameter tests are given for the final model in Figure 2. For illustrative purposes, we also present the results of the ITW-based procedure for the full model in Appendix C.

Figure 2 here

The first row in Figure 2 shows the individual knee joint angle kinematics curves (left) and the estimated adjusted p-value functions for the overall functional tests (right), indicating the presence of at least one significant effect in the majority of the jump. The grey shaded parts of the domain (left) correspond to significant effects at the 5% level (i.e., the points t with associated adjusted p-values ≤ 0.05).

As expected, the jump length has a significant effect throughout all three jump phases, in the majority of the domain (Figure 2, row 3). The functional parameter estimate $\hat{\beta}_{Jump}(t)$ is positive in intervals containing the maximal flexion during both take-off and landing, and negative directly after take-off and just before touch-down, confirming that the movement is more pronounced for longer jumps. The parts of the domain where the effect is non significant is expected, since $\beta_{Jump}(\cdot)$ is anticipated to change sign.

The last three rows of Figure 2 present results of group comparisons, testing respectively,

$$H_{0,CTRL-PT} : \beta_{CTRL}(t) = 0,\ H_{1,CTRL-PT} : \beta_{CTRL}(t) \neq 0,$$
$$H_{0,R-PT} : \beta_{R}(t) = 0,\ H_{1,R-PT} : \beta_{R}(t) \neq 0,$$
$$H_{0,CTRL-R} : \beta_{CTRL}(t) - \beta_{R}(t) = 0,\ H_{1,CTRL-R} : \beta_{CTRL}(t) - \beta_{R}(t) \neq 0,$$

with parameter estimates to the left and corresponding adjusted p-value functions to the right. The ACL_{PT} group is significantly different with respect to the control group during both take-off and landing (Figure 2, row 4). The functional parameter estimate $\hat{\beta}_{CTRL}(\cdot)$ associated to the differences indicates less flexion in the ACL_{PT} group during these two phases compared to individuals in the control group, because there $\hat{\beta}_{CTRL}(t)$ is positive. These results are in line with previously reported results (Tengman et al., 2015; Hébert-Losier et al., 2015),

indicating significant differences between physiotherapy treated individuals and controls. We do not detect any significant differences between the reconstructive surgery group and the controls (Figure 2, row 6). The overall test of $\beta_R(\cdot)$ indicates that the reconstructive surgery group is significantly different from the ACL_{PT} group (Table 2). However, there is not sufficient evidence to identify in which parts of the domain significant differences between the two groups occur (the corresponding adjusted p-value function never goes below the 5% level), cf. Figure 2, row 5.

We observe significant group-differences between 0% and 56% of the take-off phase and between 36% and 100% of the landing phase. We thus validate the clinical hypotheses that the preparation of the jump in the take-off phase and the stabilization in the landing phase are of particular interest in relation to movement control after injury. Our analysis confirms that the events of maximal flexion, analyzed e.g., in Tengman et al. (2015), provides some insight into how the groups may differ. However, to only base analyses on comparisons of movement data taken at one particular point in time, provides limited information, and with the present method we are able to detect significant differences in large parts of the take-off and landing phases.

6 Discussion and conclusions

In this work, we introduce a non-parametric methodology to test the functional parameters of a functional-on-scalar linear model with fixed effects. We provide interval-wise testing procedures based on permutations, to test hypotheses on the functional regression parameters, including domain selection. We show that our proposed IWT-based procedure is asymptotically exact and consistent.

Due to the non-parametric nature of the testing procedures that we propose, the test statistics (2.11), (2.14) and (2.16) can be replaced by the integrated versions of other feasible point-wise test statistics, given that they are continuous functions on $[a, b]$. The continuity of the point-wise test statistics with respect to t is required to guarantee that the numerical procedure described in Section 4 provides a proper estimate of the adjusted p-value functions. If integrable point-wise test statistics are to be used, more sophisticated numerical algorithms that can deal with improper integration need to be used to estimate the adjusted p-value functions.

The IWT procedure used in this paper provides control of the IWER. As shown by Pini and Vantini (2015) the control can also be extended to the complementary sets of all intervals.

The analysis of the knee kinematics data set showed that the effect of jump length on knee kinematics is significantly different from zero, while the effects

of BMI, sex, and age are not. In line with previous findings, even after having discounted for the jump length, the physiotherapy group remains significantly different with respect to the control group during take-off and landing. Our detected significant domain segments confirm the importance of the landmarks analysed earlier by Tengman et al. (2015), in the problem of identifying group differences, simultaneously indicating statistical differences on wider segments of time domain.

Estimation and interval-wise testing of functional-on-scalar linear models are of interest in many recent applications. For instance, it could be applied for comparing pulmonary volume over time of different individuals, like the data analyzed by Fogarty and Small (2014); for comparing hemolysis curves - the percent hemolysis as a function of time - at various treatment levels (Vsevolozhskaya et al., 2014); or for modeling the functional connectivity between brain regions as a function of distance between the regions, as in Reiss et al. (2010). In all mentioned cases, the methodology proposed in this work would additionally allow the selection of the intervals of the domain (i.e., time or space intervals) presenting significant effects of the covariates.

It would be of interest to extend the domain-selective inference described here to functional-on-functional linear models. In such a framework, the functional regression coefficients $\beta_l(t)$ can be replaced by functional linear operators, and the concept of interval-wise inference has to be extended accordingly. Such a model would allow to introduce the effects of time-varying covariates on the functional responses, and could be of interest in applications where also the covariates change in time. Another possible extension of the methodology proposed in this paper would be to incorporate random effects in the model.

References

Abramovich, F. and Angelini, C. (2006), 'Testing in mixed-effects FANOVA models', *J. Statist. Plann. Inference* **136**(12), 4326–4348.

Abramowicz, K., Häger, C., Hébert-Losier, K., Pini, A., Schelin, L., Strandberg, J. and Vantini, S. (2014), An inferential framework for domain selection in functional ANOVA, *in* E. G. Bongiorno, E. Salinelli, A. Goia and P. Vieu, eds, 'Contributions in infinite-dimensional statistics and related topics', Società Editrice Esculapio, pp. 13–18.

Anderson, M. J. and Legendre, P. (1999), 'An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model', *J. Stat. Comput. Sim.* **62**(3), 271–303.

Anderson, M. J. and Robinson, J. (2001), 'Permutation tests for linear models', *Aust. NZ J. Stat.* **43**(1), 75–88.

Antoniadis, A. and Sapatinas, T. (2007), 'Estimation and inference in functional mixed-effects models', *Comput. Statist. Data Anal.* **51**(10), 4793–4813.

Cardot, H., Prchal, L. and Sarda, P. (2007), 'No effect and lack-of-fit permutation tests for functional regression', *Comput. Statist.* **22**(3), 371–390.

Cuesta-Albertos, J. A. and Febrero-Bande, M. (2010), 'A simple multiway ANOVA for functional data', *TEST* **19**(3), 537–557.

Cuevas, A., Febrero, M. and Fraiman, R. (2004), 'An ANOVA test for functional data', *Comput. Statist. Data Anal.* **47**(1), 111–122.

Davison, A. C. and Hinkley, D. V. (1997), *Bootstrap methods and their application*, Cambridge university press.

Fan, J. and Zhang, J. T. (2000), 'Two-step estimation of functional linear models with applications to longitudinal data', *J. R. Stat. Soc. Ser. B Stat. Methodol.* **62**(2), 303–322.

Fogarty, C. B. and Small, D. S. (2014), 'Equivalence testing for functional data with an application to comparing pulmonary function devices', *Ann. Appl. Stat* **8**(4), 2002–2026.

Freedman, D. and Lane, D. (1983), 'A nonstochastic interpretation of reported significance levels', *J. Bus. Econ. Stat.* **1**(4), 292–298.

Gertheiss, J., Goldsmith, J., Crainiceanu, C. and Greven, S. (2013), 'Longitudinal scalar-on-functions regression with application to tractography data', *Biostatistics* **14**(3), 447–461.

Hébert-Losier, K., Pini, A., Vantini, S., Strandberg, J., Abramowicz, K., Schelin, L. and Häger, C. K. (2015), 'One-leg hop kinematics 20 years following anterior cruciate ligament ruptures: data revisited using functional data analysis', *Clinical Biomechanics* .

Kayano, M., Matsui, H., Yamaguchi, R., Imoto, S. and Miyano, S. (2015), 'Gene set differential analysis of time course expression profiles via sparse estimation in functional logistic model with application to time-dependent biomarker detection', *Biostatistics* .

Lai, T. L., Robbins, H. and Wei, C. Z. (1979), 'Strong consistency of least squares estimates in multiple regression', *J. Multivariate Anal.* **9**(3), 343–361.

Pesarin, F. and Salmaso, L. (2010), *Permutation tests for complex data: theory, applications and software*, John Wiley & Sons Inc.

Pini, A. and Vantini, S. (2015), Interval-wise testing for functional data, Technical Report 30/2015, MOX, Politecnico di Milano.

Pini, A. and Vantini, S. (2016), 'The interval testing procedure: a general framework for inference in functional data analysis.', *Biometrics* .

R Core Team (2014), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *http://www.R-project.org/*

Ramsay, J. O. and Silverman, B. W. (2005), *Functional data analysis*, Springer, New York.

Reiss, P. T., Huang, L. and Mennes, M. (2010), 'Fast function-on-scalar regression with penalized basis expansions', *Int. J. Biostat.* **6**(1).

Schott, J. R. (2007), 'Some high-dimensional tests for a one-way MANOVA', *J. Multivariate Anal.* **98**(9), 1825–1839.

Spitzner, D. J., Marron, J. S. and Essick, G. K. (2003), 'Mixed-model functional ANOVA for studying human tactile perception', *J. Amer. Statist. Assoc.* **98**(462), 263–272.

Staicu, A.-M., Li, Y., Crainiceanu, C. M. and Ruppert, D. (2014), 'Likelihood ratio tests for dependent data with applications to longitudinal and functional data analysis', *Scand. J. Stat.* **41**(4), 932–949.

Tengman, E., Grip, H., Stensdotter, A. K. and Häger, C. K. (2015), 'Anterior cruciate ligament injury about 20 years post-treatment: A kinematic analysis of one-leg hop', *Scandinavian journal of medicine & science in sports* .

Vsevolozhskaya, O. A., Greenwood, M. C., Bellante, G. J., Powell, S. L., Lawrence, R. L. and Repasky, K. S. (2013), 'Combining functions and the closure principle for performing follow-up tests in functional analysis of variance', *Comput. Statist. Data Anal.* **67**, 175–184.

Vsevolozhskaya, O., Greenwood, M. and Holodov, D. (2014), 'Pairwise comparison of treatment levels in functional analysis of variance with application to erythrocyte hemolysis', *Ann. Appl. Stat.* **8**(2), 905–925.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. and Nichols, T. E. (2014), 'Permutation inference for the general linear model', *NeuroImage* **92**, 381–397.

Zeng, C., Pan, Z., MaWhinney, S., Barón, A. E. and Zerbe, G. O. (2011), 'Permutation and F distribution of tests in the multivariate general linear model', *Am. Stat.* **65**(1), 31–36.

Zhang, J. and Liang, X. (2014), 'One-way ANOVA for functional data via globalizing the pointwise F-test', *Scand. J. Stat.* **41**(1), 51–71.

Corresponding author:
Alessia Pini
MOX - Department of Mathematics
Politecnico di Milano
Via Bonardi 9, 20133 Milano, Italy
alessia.pini@polimi.it

A Proofs

We here provide the proofs of the theorems stated in Section 3. We first report the theoretical properties of the functional interval-wise tests (2.10), (2.13) and (2.15) based on the Freedman and Lane scheme and on integrated test statistics (2.12), (2.14) and (2.16). Then, we prove that the IWT-based tests of linear hypotheses on the functional-on-scalar linear model is provided with an asymptotic control of the IWER and that they are consistent. Additionally we show that the IWT-based F-test on the regression model is provided with an exact control of the IWER.

A.1 Interval-wise tests

We start by showing asymptotic exactness of functional tests on linear hypotheses for a given interval $\mathcal{I}$.

Lemma A.1. *Under assumptions (C1-C2), and for each interval $\mathcal{I} \subseteq [a, b]$, the functional test of linear hypotheses on the regression parameters* (2.10) *based on statistic $T_C^{\mathcal{I}}$* (2.11) *is asymptotically exact.*

Proof. Let $H_{0,C}^{\mathcal{I}}$ hold, i.e., $C\boldsymbol{\beta}(t) = \mathbf{c}_0(t)$, $\forall t \in \mathcal{I}$. Under the null hypothesis, and for any $t \in \mathcal{I}$, the model can be reduced by solving the linear system $C\boldsymbol{\beta}(t) = \mathbf{c}_0(t)$. In particular, since C has full rank, $q \leq L + 1$ regression parameters can be removed from the model. Let $\mathcal{Q}$ denote the set of indexes removed. The reduced model is then $y_i(t) = \sum_{r \notin \mathcal{Q}} \beta_r(t) a_r(t) x_{ri} + \varepsilon_i(t)$, where $x_{0i} = 1$, $a_r(t)$ is a fixed known function (depending only on the solution of linear systems $C\boldsymbol{\beta}(t) = \mathbf{c}_0(t)$), and $\varepsilon_i(t)$ are *i.i.d.* and zero-mean random functions.

The generalization to the functional case of the Freedman and Lane permutation scheme is based on the joint permutations (the same for each $t \in \mathcal{I}$) of the residuals $\hat{\varepsilon}_{i,C}(t) = y_i(t) - \sum_{r \notin \mathcal{Q}} \hat{\beta}_{r,C}(t) a_r(t) x_{ri}$, where $\hat{\beta}_{r,C}(t)$, $r \notin \mathcal{Q}$ are the OLS estimates of the parameters $\beta_r(t)$, $r \notin \mathcal{Q}$ under the reduced model. Under conditions (*C1-C2*), we have strong consistency of the OLS parameters estimates, i.e., in our case: $\hat{\beta}_{r,C}(t) \xrightarrow{a.s.} \beta_r(t)$, $\forall r \notin \mathcal{Q}$, and $\forall t \in \mathcal{I}$. Hence, we also have the strong convergence of the residuals, i.e., $\hat{\varepsilon}_{i,C}(t) \xrightarrow{a.s.} \varepsilon_i(t)$, $\forall i = 1, ..., n$ and $\forall t \in \mathcal{I}$.

The errors $\varepsilon_i(t)$ of the linear model are jointly exchangeable. Hence, the likelihood of every joint permutation is invariant, and equal to $1/n!$. So, the test based on the joint permutations of the errors $\varepsilon_i(t)$ is exact. As $\hat{\varepsilon}_{i,C}(t) \xrightarrow{a.s.} \varepsilon_i(t)$, $\forall t \in \mathcal{I}$, the residuals are jointly asymptotically exchangeable, i.e., the likelihood of every joint permutation is asymptotically invariant, and converges to $1/n!$. Hence, the test based on joint permutations of the residuals is asymptotically exact.

□

Asymptotical exactness for the test of hypothesis (2.15) for the lth functional parameter, based on statistic $T_l^{\mathcal{I}}$ (2.16), is a direct consequence of the above

lemma. In addition to asymptotic results for functional tests on any linear hypothesis, we prove exactness for the overall functional model test of hypothesis (2.13):

Lemma A.2. *For each interval $\mathcal{I} \subseteq [a, b]$, the functional overall model test of hypothesis* (2.13) *based on statistic $T_F^{\mathcal{I}}$* (2.14) *is exact.*

Proof. Under $H_{0,F}^{\mathcal{I}}$, we have $y_i(t) = \beta_0(t) + \varepsilon_i(t)$, $\forall t \in \mathcal{I}$. The estimated residuals of this model are $\hat{\varepsilon}_{i,0}(t) = \beta_0(t) + \varepsilon_i(t) - \hat{\beta}_0(t)$, where $\hat{\beta}_0(t) = \bar{y}(t)$ is the sample mean of the responses $y_i(t)$. Note that the quantity $\beta_0(t) + \varepsilon_i(t) - \hat{\beta}_0(t)$ is permutationally invariant. Hence, the independence between the random functions $\varepsilon_i(t)$ implies the exchangeability with respect to units of the residual functions $\hat{\varepsilon}_{i,0}(t)$ under $H_{0,F}^{\mathcal{I}}$. Thus, the test is exact, as it is based on the permutation of exchangeable quantities (Pesarin and Salmaso, 2010). □

In the next step, we verify the consistency of functional tests on linear hypotheses.

Lemma A.3. *For each interval $\mathcal{I} \subseteq [a, b]$, the functional test of linear hypotheses on the regression parameters* (2.10) *based on the test statistic $T_C^{\mathcal{I}}$* (2.11) *is consistent.*

Proof. The statement follows directly from the fact that the test statistic $T_C^{\mathcal{I}}$ is stochastically greater under $H_{1,C}^{\mathcal{I}}$ than under $H_{0,C}^{\mathcal{I}}$ (Pesarin and Salmaso, 2010). □

As a direct implication of Lemma A.3, we get consistency for the overall functional model test of hypothesis (2.13) and the l:th functional parameter test of hypothesis (2.15) based on test statistics $T_F^{\mathcal{I}}$ and $T_l^{\mathcal{I}}$, respectively.

A.2 Properties of domain selection IWT procedures

We start by proving Theorem 3.1, establishing asymptotic interval-wise control of the domain selection IWT procedure for tests of linear hypotheses.

Proof of Theorem 3.1. Let $\mathcal{I} \subseteq [a, b]$ be an interval associated to a true null hypothesis, i.e., let $H_{0,C}^{\mathcal{I}}$ hold. Consider a point $t \in \mathcal{I}$, and let $\mathcal{S}_t$ denote the set of all intervals containing the point t. The IWT-adjusted p-value associated to point t is $\tilde{p}_C(t) = \max_{\mathcal{J} \in \mathcal{I}_t} p_C^{\mathcal{J}}$, where $p_C^{\mathcal{J}}$ is the p-value of the permutation test on interval $\mathcal{J}$. In particular, since $\mathcal{I} \in \mathcal{S}_t$, we have that $\forall t \in \mathcal{I}$, $\tilde{p}_C(t) \geq p_C^{\mathcal{I}}$. Therefore, $\mathbb{P}_{H_{0,C}^{\mathcal{I}}}[\forall t \in \mathcal{I} : \tilde{p}_C(t) \leq \alpha] \leq \mathbb{P}_{H_{0,C}^{\mathcal{I}}}[p_C^{\mathcal{I}} \leq \alpha]$. Since all tests are asymptotically exact (Lemma A.1), we have:

$$\lim_{n \to \infty} \mathbb{P}_{H_{0,C}^{\mathcal{I}}} \left[p_C^{\mathcal{I}} \leq \alpha \right] = \alpha,$$

and therefore,

$$\limsup_{n\to\infty} \mathbb{P}_{H^{\mathcal{I}}_{0,C}} \left[\forall t \in \mathcal{I} : \tilde{p}_C(t) \leq \alpha\right] \leq \alpha.$$

□

Assertion of Proposition 3.3 follows directly from the results of Pini and Vantini (2016) and the fact that interval-wise tests used to build the procedure are exact (Lemma A.2).

Further, since all functional interval-wise tests are consistent (Lemma A.3), the IWT procedure is also consistent, due to the result of Pini and Vantini (2015) (Theorem 3). As special cases of Theorem 3.4, Corollary 3.5 and Corollary 3.6 follow immediately.

B The Freedman and Lane permutation scheme

In this section, we give some details of the implementation of the Freedman and Lane permutation scheme for testing linear hypotheses on the regression model for each interval $\mathcal{I} \subseteq [a, b]$. In detail, we start from the restriction of functional linear model (2.1) to interval $\mathcal{I}$

$$y_i(t) = \sum_{l=0}^{L} \beta_l(t) x_{li} + \varepsilon_i(t), \quad \forall i = 1, \ldots, n;\ t \in \mathcal{I} \tag{B.1}$$

with $x_{0i} = 1,\ \forall i$, and describe the extention of the Freedman and Lane permutation scheme to the functional case, for implementing the functional test (2.10).

The Freedman and Lane permutations are based on the following steps:

(i) the residuals of the reduced model (that is the linear model under the null hypothesis) are estimated;

(ii) the residuals of the reduced model are permuted;

(iii) the permuted responses are computed, through the reduced model and permuted residuals.

For more details about this method, we refer to Freedman and Lane (1983); Anderson and Legendre (1999). In the following subsections we present the details of the scheme for the hypotheses introduced in our paper.

B.1 Tests on general linear hypotheses

Under the null hypothesis (2.10), the model (B.1) can be reduced by solving the linear system $C\boldsymbol{\beta}(t) = \mathbf{c}_0(t)$. In particular, since C has full rank, q regression

parameters can be removed from the model, by expressing them in terms of the others. Let $\mathcal{Q}$ denote the set of indexes of the removed regression parameters. The reduced model is then:

$$y_i(t) = \sum_{r \notin \mathcal{Q}} \beta_r(t) \tilde{x}_{ri} + \varepsilon_i(t), \tag{B.2}$$

i.e., responses can be written in terms of a linear combination of modified covariates $\tilde{x}_{ri} = a_r(t) x_{ri}$, where $a_r(t)$ are fixed known coefficients, depending only on the solution of the linear system $C\boldsymbol{\beta}(t) = \mathbf{c}_0(t)$, and $\varepsilon_i(t)$ are *i.i.d.* and zero-mean errors.

The residuals of the reduced model can then be estimated as $\hat{\varepsilon}_{i,C}(t) = y_i(t) - \sum_{r \notin \mathcal{Q}} \hat{\beta}_{r,C}(t) \tilde{x}_{ri}$, where $\hat{\beta}_{r,C}(t)$ are the OLS estimates of parameters $\beta_r(t)$, $r \notin \mathcal{Q}$, of model (B.2). Then, the residuals $\hat{\varepsilon}_{i,C}(t)$ are permuted, and the permuted responses are computed using the permuted residuals $\hat{\varepsilon}^*_{i,C}(t)$ in the reduced model (B.2):

$$y^*_i(t) = \sum_{r \notin \mathcal{Q}} \hat{\beta}_{r,C}(t) \tilde{x}_{ri} + \hat{\varepsilon}^*_{i,C}(t). \tag{B.3}$$

B.2 Overall model tests

In the case of tests on the overall functional model hypothesis (2.13), under the null hypothesis $H^{\mathcal{I}}_{0,F}$ all regression parameters except the intercept are null. So, the reduced model is:

$$y_i(t) = \beta_0(t) + \varepsilon_i(t).$$

The estimated residuals of the reduced model are $\hat{\varepsilon}_{i,F}(t) = y_i(t) - \bar{y}(t)$, where $\bar{y}(t)$ is the sample mean of the responses $y_i(t)$. Therefore, using the permuted residuals $\hat{\varepsilon}^*_{i,F}(t)$, we get:

$$y_i{}^*(t) = \bar{y}(t) + \hat{\varepsilon}^*_{i,F}(t).$$

Note that in this case permuting the estimated residuals $\hat{\varepsilon}_{i,F}(t)$ is equivalent to permuting the responses $y_i(t)$, $i = 1, \ldots, n$.

B.3 Single parameter tests

In the case of tests of hypothesis on the l:th functional parameter (2.15), the model under the null hypothesis $H^{\mathcal{I}}_{0,l}$ reduces to:

$$y_i(t) = \beta_0(t) + \sum_{r \neq l} \beta_r(t) x_{ri} + \varepsilon_i(t).$$

The estimated residuals of such model are $\hat{\varepsilon}_{i,l}(t) = y_i(t) - \hat{\beta}_{0,l}(t) - \sum_{r \neq l} \hat{\beta}_{r,l}(t) x_{li}$, where $\hat{\beta}_{r,l}(t)$ are the OLS estimates of the parameters of the reduced model. Then, the permuted responses are:

$$y_i^*(t) = \hat{\beta}_{0,l}(t) + \sum_{r \neq l} \hat{\beta}_{r,l}(t) + \hat{\varepsilon}_{i,l}^*(t), \tag{B.4}$$

where $\hat{\varepsilon}_{i,l}^*(t)$ are the permuted residuals.

C Full model results

Figure 3 presents the results for the full model initially estimated for the knee data.

Figure 3 here

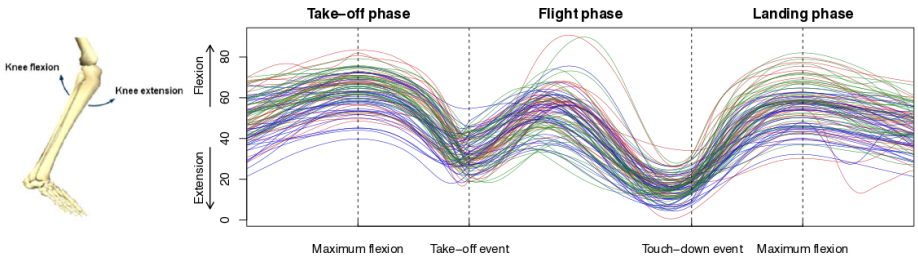

Figure 1: Knee flexion and extension in the sagittal plane (left) and flexion/extension angle curves of the physiotherapy (blue), reconstructive surgery (red) and control (green) groups (right).

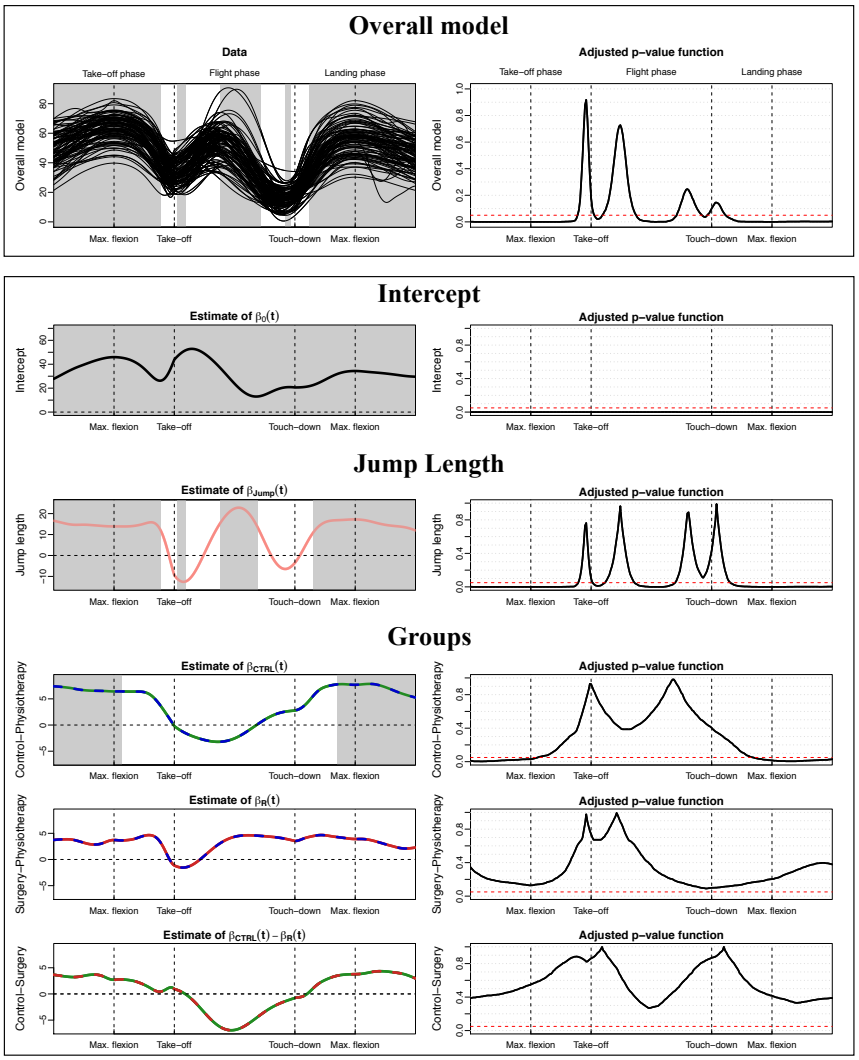

Figure 2: Functional parameter estimates and estimated adjusted p-value functions for the final (reduced) functional-on-scalar model. Top panel: Knee joint angle kinematics data (functional response variable) for the 95 individuals (left) and estimated adjusted p-value function for the overall functional model tests (right). Bottom panel: Functional parameter estimates (left) and corresponding estimated adjusted p-value functions (right) for the intercept, jump length, and group effects, respectively. Grey-shaded parts (left) correspond to significant effects at the 5% level, i.e. the points t with associated adjusted p-values less or equal to 0.05.

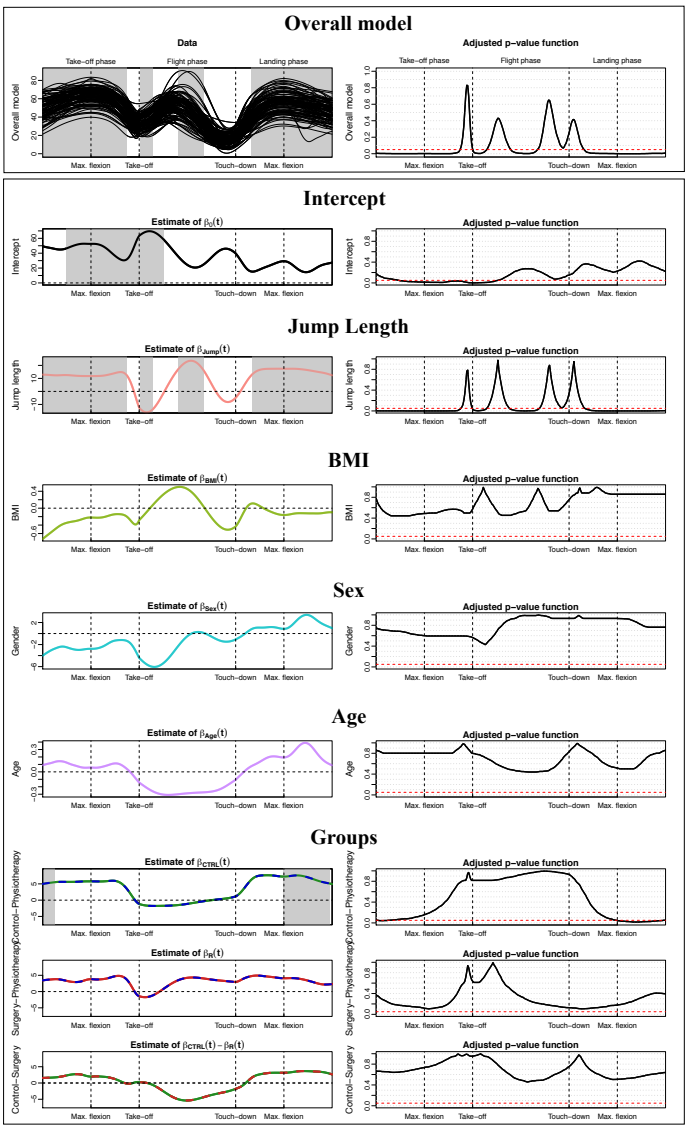

Figure 3: Functional parameter estimates and estimated adjusted p-value functions for the initial (full) functional-on-scalar model. Top panel: Knee joint angle kinematics data (functional response variable) for the 95 individuals (left) and estimated adjusted p-value function for the overall functional model tests (right). Bottom panel: Functional parameter estimates (left) and corresponding estimated adjusted p-value functions (right) for the intercept, jump length, BMI, sex, age and group effects, respectively. Grey-shaded parts (left) correspond to significant effects at the 5% level, i.e. the points t with associated adjusted p-values less or equal to 0.05.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

29/2016 Miglio, E.; Parolini, N.; Penati, M.; Porcù, R.
GPU parallelization of brownout simulations with a non-interacting particles dynamic model

28/2016 Antonietti, P.F.; Dal Santo, N.; Mazzieri, I.; Quarteroni, A.
A high-order discontinuous Galerkin approximation to ordinary differential equations with applications to elastodynamics

26/2016 Brunetto, D.; Calderoni, F.; Piccardi, C.
Communities in criminal networks: A case study

27/2016 Repossi, E.; Rosso, R.; Verani, M.
A phase-field model for liquid-gas mixtures: mathematical modelling and Discontinuous Galerkin discretization

25/2016 Baroli, D.; Cova, C.M.; Perotto, S.; Sala, L.; Veneziani, A.
Hi-POD solution of parametrized fluid dynamics problems: preliminary results

24/2016 Pagani, S.; Manzoni, A.; Quarteroni, A.
A Reduced Basis Ensemble Kalman Filter for State/parameter Identification in Large-scale Nonlinear Dynamical Systems

23/2016 Fedele, M.; Faggiano, E.; Dedè, L.; Quarteroni, A.
A Patient-Specific Aortic Valve Model based on Moving Resistive Immersed Implicit Surfaces

22/2016 Antonietti, P.F.; Facciola', C.; Russo, A.;Verani, M.
Discontinuous Galerkin approximation of flows in fractured porous media

21/2016 Ambrosi, D.; Zanzottera, A.
Mechanics and polarity in cell motility

20/2016 Wilhelm, M.; Sangalli, L.M.
Generalized Spatial Regression with Differential Regularization