

# A goodness-of-fit test for the functional linear model with functional response

Eduardo García-Portugués<sup>1,2</sup>, Javier Álvarez-Liébana<sup>3</sup>,  
Gonzalo Álvarez-Pérez<sup>4,6</sup>, and Wenceslao González-Manteiga<sup>5</sup>

## Abstract

The Functional Linear Model with Functional Response (FLMFR) is one of the most fundamental models to assess the relation between two functional random variables. In this paper, we propose a novel goodness-of-fit test for the FLMFR against a general, unspecified, alternative. The test statistic is formulated in terms of a Cramér–von Mises norm over a doubly-projected empirical process which, using geometrical arguments, yields an easy-to-compute weighted quadratic norm. A resampling procedure calibrates the test through a wild bootstrap on the residuals and the use of convenient computational procedures. As a sideways contribution, and since the statistic requires a reliable estimator of the FLMFR, we discuss and compare several regularized estimators, providing a new one specifically convenient for our test. The finite sample behavior of the test is illustrated via a simulation study. Also, the new proposal is compared with previous significance tests. Two novel real datasets illustrate the application of the new test.

**Keywords:** Bootstrap; Cramér–von Mises statistic; Functional data; Regularization.

## 1 Introduction

The increasing availability of data for continuous processes has boosted the field of Functional Data Analysis (FDA) in the last decades as a powerful tool to take advantage of the complexity and rich structure of this kind of data, difficult to manage for many traditional statistical techniques given their intrinsically infinite dimensionality. Some of the main monographs in FDA are Ramsay and Silverman (2005), Ferraty and Vieu (2006), Horváth and Kokoszka (2012), and Hsing and Eubank (2015).

Regression models with functional covariates and/or responses emerged as natural generalizations of multivariate ones. A specific instance arises when assessing the relation between two functional random variables  $\mathcal{X}$  and  $\mathcal{Y}$  via a general regression model  $\mathcal{Y} = m(\mathcal{X}) + \mathcal{E}$ , where  $\mathcal{E}$  is a functional random error. The main difference with the multivariate case is that here  $m$  is an operator between function spaces, typically of a Hilbertian nature, therefore generalizing the usual Euclidean-Euclidean regression mapping. Nonparametric estimation of  $m$  was addressed by Ferraty et al. (2011) and Lian (2011), who investigated the rates of convergence of kernel and  $k$ -nearest neighbors regression estimates, respectively. Moreover, Ferraty et al. (2012) studied the nonparametric estimation of  $m$  by considering data-driven bases and consistent bootstrap approaches.

However, much of the existing regression literature is concerned with (infinite-dimensional) parametric modeling, where the operator  $m$  is assumed to belong to a given parametric family. As

---

<sup>1</sup>Department of Statistics, Carlos III University of Madrid (Spain).

<sup>2</sup>UC3M-Santander Big Data Institute, Carlos III University of Madrid (Spain).

<sup>3</sup>Department of Statistics and Operations Research and Mathematics Didactics, University of Oviedo (Spain).

<sup>4</sup>Department of Physics, University of Oviedo (Spain).

<sup>5</sup>Department of Statistics, Mathematical Analysis and Optimization, University of Santiago de Compostela (Spain).

<sup>6</sup>Corresponding author. e-mail: gonzaloalvarez@uniovi.es.

an early precedent, the simplest and best-known paradigm is the Functional Linear Model with Scalar Response (FLMSR),  $Y = m_\rho(\mathcal{X}) + \varepsilon$ , where  $\varepsilon$  is a real-valued error and  $m_\rho$  is a linear functional depending on a function  $\rho$ . Within the FLMSR, the so-called Functional Principal Components Regression (FPCR) was introduced by Cardot et al. (1999) as a parsimonious estimation approach. Crambes et al. (2009) proposed a smoothing splines estimator, whereas Aguilera and Aguilera-Morillo (2013) formulated penalized FPCR estimation techniques based on B-splines. Alternatively, functional partial least squares regression was proposed in Preda and Saporta (2005). Some authors have also studied the relation of a functional response and a scalar regressor, see, e.g., Chiou et al. (2003).

In contrast, the Functional Linear Model with Functional Response (FLMFR),  $\mathcal{Y} = m(\mathcal{X}) + \mathcal{E}$ , where  $m$  is a linear operator, has received considerably less attention. When a Hilbertian framework is considered,  $m \equiv m_\beta$  is usually assumed to be a Hilbert–Schmidt operator between  $L^2$  spaces admitting an integral representation in terms of a bivariate kernel  $\beta$ . Ramsay and Silverman (2005) proposed to estimate  $\beta$  based on minimizing the residual sum of squared norms. Motivated by signal transmission problems, Cuevas et al. (2002) provided an estimator considering a fixed and triangular design. An estimator in terms of the Karhunen–Loève expansions of functional response and regressor was discussed in Yao et al. (2005). Crambes and Mas (2013) provided asymptotic results for prediction under the FLMFR through the Karhunen–Loève expansion of the functional regressor, whereas Imaizumi and Kato (2018) derived minimax optimal rates. An estimation based on functional canonical correlation analysis was suggested in He et al. (2010). The FLMFR when both response and covariate are densities was analyzed in Park and Qian (2012).

Several authors have contributed to the Goodness-of-Fit (GoF) framework for regression models, see González-Manteiga and Crujeiras (2013) for a comprehensive review. The first attempts, following the ideas of Bickel and Rosenblatt (1973) in scalar and multivariate contexts, were focused on smoothing-based tests, see Härdle and Mammen (1993). Alternatively, upon the work of Durbin (1973), and aimed at solving the sensitiveness of those approaches to the smoothing parameter, Stute (1997) proposed a GoF test based on the integrated regression function. Extending this work to the high-dimensional context, Escanciano (2006) proposed a GoF test, in terms of a residual marked empirical process based on projections, designed to overcome the poor empirical power inherent to the curse of dimensionality. Promoting these ideas to the FDA context, García-Portugués et al. (2014) and Cuesta-Albertos et al. (2019) derived an easily computable GoF test for the FLMSR in terms of projections. The former proposed a methodology based on the projected empirical estimator of the integrated regression function, whereas the latter considered marked empirical process indexed by a single randomly projected functional covariate, providing a more computationally efficient test.

In addition to the GoF proposals for the FLMSR discussed above, Delsol et al. (2011) formulated a kernel-based test for model assumptions, whereas Bücher et al. (2011) introduced testing procedures well-adapted for the time-variation of directional profiles. Generalized likelihood ratio tests were suggested in McLean et al. (2015) to test the linearity of functional generalized additive models. Staicu et al. (2015) tested the equality of multiple group mean functions for hierarchical functional data. In the context of semi-functional partial linear model, where the scalar response is regressed on multivariate and functional covariates, Aneiros-Pérez and Vieu (2013) tested the simple linear null hypothesis. In the FLMSR setup, a comparative study has been recently provided by Yasemin-Tekbudak et al. (2019), comparing GoF tests in Horváth and Reeder (2013), García-Portugués et al. (2014), McLean et al. (2015), and Kong et al. (2016).

The extension of these GoF proposals to the FLMFR context is currently an open challenge. This model is being applied to a wide range of fields, such as electricity market (Benatia et al., 2017), biology (He et al., 2010) or the study of lifetime patterns (Imaizumi and Kato, 2018), to cite but some,

hence the practical relevance of developing a GoF test for it. Testing the lack of effect, which is actually a particular case of the FLMFR, has received considerable attention: Kokoszka et al. (2008) proposed an FPC-based significance test within the FLMFR; Patilea et al. (2016b) introduced a kernel-based significance test consistent to nonlinear alternatives; Lee et al. (2020) proposed a significance test, within the FLMFR, using an extension to the functional setup of the correlation-based metric adopted in Park et al. (2015). Related testing approaches within the FLMFR include those of Chiou and Müller (2007), which addressed the development of a FPC-based residual diagnostic tool, and Gabrys et al. (2010), that tested if functional residuals are independent and identically distributed (iid). Sharing the aim of the time-domain-based test in Gabrys et al. (2010), Zhang (2016) has recently proposed a Cramér–von Mises test for the functional white noise, with applications to assessing the uncorrelatedness of the residuals in FLMFR and functional autoregressive model fits, but under a frequency-domain framework, in terms of the functional periodogram previously derived in Panaretos and Tavakoli (2013). Empirical likelihood ratio tests were formulated by Wang et al. (2018) for concurrent models. No proposals extending the generalized likelihood ratio test approach seem to exist for the FLMFR. As a consequence, the development of GoF tests for the FLMFR, against unspecified alternatives, is an area still substantially unexplored.

In this paper, we propose a GoF test for the FLMFR, that is, for testing the composite null hypothesis

$$\mathcal{H}_0 : m \in \mathcal{L} = \left\{ m_\beta(\mathcal{X})(t) = \int_a^b \mathcal{X}(s)\beta(s, t) ds : \beta \in L^2([a, b] \times [c, d]) \right\}.$$

Our methodology is based on characterizing  $\mathcal{H}_0$  in terms of the integral regression operator arising from a double projection, of the functional covariate and the response, in terms of finite-dimensional functional directions. The deviation of the resulting empirical process from its expected zero mean is measured by a Cramér–von Mises statistic that integrates on both functional directions and is calibrated via an efficient wild bootstrap on the residuals. We show that our GoF test exhibits an adequate behavior, in terms of size and power, for the composite hypothesis, under two common scenarios: the no effects model and the FLMFR. Besides, since the test can be readily modified for the simple hypothesis  $\beta \equiv 0$ , we compare our GoF test with the procedures from Kokoszka et al. (2008) and Patilea et al. (2016b), obtaining competitive powers. As a by-product contribution, we provide a convenient hybrid approach for the estimation of  $\beta$  based on LASSO (Tibshirani, 1996) regularization and linearly-constrained least-squares. The companion R package `goffda` (García-Portugués and Álvarez-Liébana, 2020) implements all the methods presented in the paper and allows for replication of the real data applications.

The rest of this paper is organized as follows. Section 2 introduces the required background on FDA and the FLMFR, addressing the estimation of the regression operator and providing a brief comparative study between different estimation techniques. Section 3 is devoted to the theoretical, computational, and resampling aspects of the new GoF test. A comprehensive simulation study and a real data application are presented in Sections 4 and 5, respectively. Conclusions are drawn in Section 6. Appendix A contains the proofs of the lemmas and the Supporting Information (SI) provides another data application.

## 2 Functional data and the FLMFR

### 2.1 Functional bases

Given the functional bases  $\{\Psi_j\}_{j=1}^\infty$  and  $\{\Phi_k\}_{k=1}^\infty$  in the separable Hilbert spaces  $\mathbb{H}_1$  and  $\mathbb{H}_2$ , respectively, any elements  $\mathcal{X} \in \mathbb{H}_1$  and  $\mathcal{Y} \in \mathbb{H}_2$  can be represented as  $\mathcal{X} = \sum_{j=1}^\infty x_j \Psi_j$  and  $\mathcal{Y} = \sum_{k=1}^\infty y_k \Phi_k$ , where  $x_j = \langle \mathcal{X}, \Psi_j \rangle_{\mathbb{H}_1}$  and  $y_k = \langle \mathcal{Y}, \Phi_k \rangle_{\mathbb{H}_2}$ , for each  $j, k \geq 1$ . Typical examples are the B-splines basis

(non-orthogonal piece-wise polynomial bases) or the Fourier basis. Both bases are of a deterministic nature and, despite their flexibility, usually require a larger number of elements to adequately represent a functional sample  $\{\mathcal{X}_i\}_{i=1}^n$ . A more parsimonious representation can be achieved by considering data-driven orthogonal bases, being the most popular choice the (empirical) Functional Principal Components (FPC) of  $\{\mathcal{X}_i\}_{i=1}^n$ ,  $\{\hat{\Psi}_j\}_{j=1}^n$ , the eigenfunctions of the sample covariance operator.

To develop the test, we will consider a  $p$ -truncated basis  $\{\Psi_j\}_{j=1}^p$  in  $\mathbb{H}_1$ , corresponding to the first  $p$  elements of  $\{\Psi_j\}_{j=1}^\infty$ . The projection of  $\mathcal{X}$  on this truncated basis is denoted by  $\mathcal{X}^{(p)} = \sum_{j=1}^p x_j \Psi_j$  and we set  $\mathbf{x}_p := (x_1, \dots, x_p)$ . We will also require to integrate on the functional analogue of the Euclidean  $(p-1)$ -sphere  $\mathbb{S}^{p-1} = \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x}\| = 1\}$ , the  $(p-1)$ -sphere of  $\mathbb{H}_1$  on  $\{\Psi_j\}_{j=1}^\infty$  defined as  $\mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^\infty}^{p-1} := \{f = \sum_{j=1}^p x_j \Psi_j \in \mathbb{H}_1 : \|f\|_{\mathbb{H}_1} = 1\}$ . The relationship between  $\mathbb{S}^{p-1}$  and  $\mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^\infty}^{p-1}$  follows easily (García-Portugués et al., 2014) considering the positive semi-definite matrix  $\mathbf{\Psi} = (\langle \Psi_j, \Psi_\ell \rangle_{\mathbb{H}_1})_{j, \ell=1, \dots, p}$ , whose Cholesky decomposition is  $\mathbf{\Psi} = \mathbf{P}'_p \mathbf{P}_p$ . Then, the  $(p-1)$ -ellipsoid  $\mathbb{S}_{\mathbf{\Psi}}^{p-1} = \{\mathbf{x} \in \mathbb{R}^p : \mathbf{x}' \mathbf{\Psi} \mathbf{x} = 1\}$  is trivially isomorphic with  $\mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^\infty}^{p-1}$  by  $f = \sum_{j=1}^p x_j \Psi_j \in \mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^\infty}^{p-1} \mapsto \mathbf{x}_p \in \mathbb{S}_{\mathbf{\Psi}}^{p-1}$ . Considering also the linear mapping  $\mathbf{x} \in \mathbb{S}^{p-1} \mapsto \mathbf{P}_p^{-1} \mathbf{x} \in \mathbb{S}_{\mathbf{\Psi}}^{p-1}$ , the integration of a functional operator  $\mathcal{T}$  with respect to  $\gamma^{(p)} \in \mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^\infty}^{p-1}$  can be written as

$$\int_{\mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^\infty}^{p-1}} \mathcal{T}(\gamma^{(p)}) d\gamma^{(p)} = \int_{\mathbb{S}_{\mathbf{\Psi}}^{p-1}} \mathcal{T}\left(\sum_{j=1}^p g_j \Psi_j\right) d\mathbf{g}_p = \int_{\mathbb{S}^{p-1}} |\mathbf{P}_p|^{-1} \mathcal{T}\left(\sum_{j=1}^p (\mathbf{P}_p^{-1} \mathbf{g}_p)_j \Psi_j\right) d\mathbf{g}_p, \quad (1)$$

where  $(\mathbf{P}_p^{-1} \mathbf{g}_p)_j$  denotes the  $j$ -th component of the vector  $\mathbf{P}_p^{-1} \mathbf{g}_p$  and  $\mathbf{g}_p$  is the vector of coefficients of  $\gamma^{(p)}$  in the  $p$ -truncated basis. If the basis is orthonormal, then  $\mathbf{\Psi}$  and  $\mathbf{P}_p$  are the identity matrices of order  $p$ , denoted as  $\mathbf{I}_p$ , and  $\mathbf{g}_p \in \mathbb{S}^{p-1}$  without any transformation. Clearly, an analogous development can be established for  $\mathbb{S}_{\mathbb{H}_2, \{\Phi_k\}_{k=1}^\infty}^{q-1}$  by means of  $\mathbf{\Phi} = (\langle \Phi_k, \Phi_\ell \rangle_{\mathbb{H}_2})_{k, \ell=1, \dots, q}$  where  $\{\Phi_k\}_{k=1}^q$  is a  $q$ -truncated basis in  $\mathbb{H}_2$ .

## 2.2 The FLMFR

We consider the context of functional regression with  $\mathbb{H}_2$ -valued functional response  $\mathcal{Y}$  and  $\mathbb{H}_1$ -valued functional covariate  $\mathcal{X}$ :

$$\mathcal{Y} = m(\mathcal{X}) + \mathcal{E}, \quad (2)$$

where the regression operator is defined as  $m(x) = \mathbb{E}[\mathcal{Y} | \mathcal{X} = x]$  and the  $\mathbb{H}_2$ -valued error is such that  $\mathbb{E}[\mathcal{E} | \mathcal{X}] = 0$ . Within this setting, we assume that  $\mathcal{X}$  and  $\mathcal{Y}$  are already centered so there is no need for an intercept term in (2). Particularly, we consider  $L^2$  spaces and assume, in what follows, that  $\mathcal{X} \in \mathbb{H}_1 = L^2([a, b])$  and  $\mathcal{Y} \in \mathbb{H}_2 = L^2([c, d])$ , unless otherwise explicitly mentioned.

In this context, the simplest parametric model is the FLMFR, in which the regression operator  $m : \mathbb{H}_1 \rightarrow \mathbb{H}_2$  is usually assumed to be a Hilbert–Schmidt integral operator, i.e.,  $m$  admits an integral representation  $m_\beta$  given by a bivariate kernel  $\beta \in \mathbb{H}_1 \otimes \mathbb{H}_2 = L^2([a, b] \times [c, d])$  as follows:

$$m_\beta(\mathcal{X})(t) = \int_a^b \beta(s, t) \mathcal{X}(s) ds, \quad t \in [c, d]. \quad (3)$$

In particular, the Hilbert–Schmidt condition directly implies that  $m$  is a compact operator, that is,  $\beta$  can be decomposed in terms of the tensor product of any pair of bases in  $\mathbb{H}_1$  and  $\mathbb{H}_2$ , since such tensor product constitutes a basis on the space of Hilbert–Schmidt operators. As a consequence,

$$\beta = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} b_{jk} (\Psi_j \otimes \Phi_k), \quad b_{jk} = \frac{\langle \beta, \Psi_j \otimes \Phi_k \rangle_{\mathbb{H}_1 \otimes \mathbb{H}_2}}{\|\Psi_j\|_{\mathbb{H}_1}^2 \|\Phi_k\|_{\mathbb{H}_2}^2}, \quad (4)$$

with  $j, k \geq 1$ . For convenience, we denote the linear integral operator in (3) by  $\langle\langle \cdot, \star \rangle\rangle$ , defined as  $\langle\langle \cdot, \star \rangle\rangle: \mathbb{H}_1 \times (\mathbb{H}_1 \otimes \mathbb{H}_2) \rightarrow \mathbb{H}_2$ ,  $\langle\langle \mathcal{X}, \beta \rangle\rangle(t) := \langle \mathcal{X}, \beta(\cdot, t) \rangle_{\mathbb{H}_1}$ . Therefore, the FLMFR from (2)–(3) can be succinctly denoted as

$$\mathcal{Y} = \langle\langle \mathcal{X}, \beta \rangle\rangle + \mathcal{E}. \quad (5)$$

Bearing in mind that  $\mathcal{X} = \sum_{j=1}^{\infty} x_j \Psi_j$  and  $\mathcal{Y} = \sum_{k=1}^{\infty} y_k \Phi_k$ , then

$$\langle\langle \mathcal{X}, \beta \rangle\rangle = \left\langle \left\langle \sum_{j=1}^{\infty} x_j \Psi_j, \sum_{\ell=1}^{\infty} \sum_{k=1}^{\infty} b_{\ell k} (\Psi_{\ell} \otimes \Phi_k) \right\rangle \right\rangle = \sum_{j=1}^{\infty} \sum_{\ell=1}^{\infty} \sum_{k=1}^{\infty} b_{\ell k} x_j \langle \Psi_j, \Psi_{\ell} \rangle_{\mathbb{H}_1} \Phi_k, \quad (6)$$

with  $\langle \Psi_j, \Psi_{\ell} \rangle_{\mathbb{H}_1} = \delta_{j\ell}$ ,  $j, \ell \geq 1$ , for orthonormal bases. From (6) and  $\mathcal{E} = \sum_{k=1}^{\infty} e_k \Phi_k$ ,

$$y_k = \sum_{j=1}^{\infty} \sum_{\ell=1}^{\infty} b_{\ell k} x_j \langle \Psi_j, \Psi_{\ell} \rangle_{\mathbb{H}_1} + e_k, \quad k \geq 1.$$

This (infinite) linear model is usually approached by projecting the variables in the truncated bases  $\{\Psi_j\}_{j=1}^p$  and  $\{\Phi_k\}_{k=1}^q$  (Ramsay and Silverman, 2005, Chapter 16), obtaining the  $(p, q)$ -truncated population version

$$y_k = \sum_{j=1}^p \sum_{\ell=1}^p b_{\ell k} x_j \langle \Psi_j, \Psi_{\ell} \rangle_{\mathbb{H}_1} + e_k, \quad k = 1, \dots, q. \quad (7)$$

Note that an equivalent way of expressing (7) is  $\mathcal{Y}^{(q)} = \langle\langle \mathcal{X}^{(p)}, \beta^{(p,q)} \rangle\rangle + \mathcal{E}^{(q)}$ , where  $\beta^{(p,q)}$  is the projection of (4) into  $\{\Psi_j \otimes \Phi_k\}_{j,k=1}^{p,q}$ .

Now, given an iid centered sample  $\{(\mathcal{X}_i, \mathcal{Y}_i)\}_{i=1}^n$  such that  $\mathcal{Y}_i = \langle\langle \mathcal{X}_i, \beta \rangle\rangle + \mathcal{E}_i$ , the sample version of (7) is expressed in matrix form as

$$\mathbf{Y}_q = \mathbf{X}_p \mathbf{\Psi} \mathbf{B}_{p,q} + \mathbf{E}_q, \quad (8)$$

where  $\mathbf{Y}_q$  and  $\mathbf{E}_q$  are the  $n \times q$  matrices with the coefficients of  $\{\mathcal{Y}_i\}_{i=1}^n$  and  $\{\mathcal{E}_i\}_{i=1}^n$ , respectively, on  $\{\Phi_k\}_{k=1}^q$ ,  $\mathbf{X}_p$  is the  $n \times p$  matrix of coefficients of  $\{\mathcal{X}_i\}_{i=1}^n$  on  $\{\Psi_j\}_{j=1}^p$ , and  $\mathbf{B}_{p,q}$  is the  $p \times q$  matrix of unknown coefficients on  $\{\Psi_j \otimes \Phi_k\}_{j,k=1}^{p,q}$ . Observe that these matrices are centered by columns and hence the model does not have an intercept. Clearly, due to the form of (8), estimators for  $\beta$  in (4) readily follow from the linear model theory. We discuss them next, focusing exclusively on orthonormal bases. This can be done without loss of generality; just replace  $\mathbf{X}_p$  by  $\tilde{\mathbf{X}}_p := \mathbf{X}_p \mathbf{\Psi}$  subsequently for non-orthonormal bases.

### 2.3 Model estimation

FPCR considers in (8) the data-driven bases given by the (empirical) FPC  $\{\hat{\Psi}_j\}_{j=1}^p$  and  $\{\hat{\Phi}_k\}_{k=1}^q$  of  $\{\mathcal{X}_i\}_{i=1}^n$  and  $\{\mathcal{Y}_i\}_{i=1}^n$ , respectively, where  $p, q \leq n$ . The estimator of  $\beta$  is then defined as the least-squares estimator of the  $(p, q)$ -truncated model given in (7) and (8):

$$\hat{\mathbf{B}}_{p,q} = \arg \min_{\mathbf{B}_{p,q}} \|\mathbf{Y}_q - \mathbf{X}_p \mathbf{B}_{p,q}\|^2 = \arg \min_{\beta^{(p,q)}} \sum_{i=1}^n \left\| \mathcal{Y}_i^{(q)} - \langle\langle \mathcal{X}_i^{(p)}, \beta^{(p,q)} \rangle\rangle \right\|^2.$$

Clearly, least-squares estimation gives  $\hat{\mathbf{B}}_{p,q} = (\mathbf{X}_p' \mathbf{X}_p)^{-1} \mathbf{X}_p' \mathbf{Y}_q$ , with  $(\hat{\mathbf{B}}_{p,q})_{jk} = \hat{b}_{jk}$ ,  $j = 1, \dots, p$ ,  $k = 1, \dots, q$ . The estimator of  $\beta^{(p,q)}$  is then  $\hat{\beta}^{(p,q)} = \sum_{j=1}^p \sum_{k=1}^q \hat{b}_{jk} (\hat{\Psi}_j \otimes \hat{\Phi}_k)$ .

The estimator  $\hat{\beta}^{(p,q)}$  critically depends on  $(p, q)$ , hence an automatic data-driven selection of  $(p, q)$  is of most practical interest. A possibility is to extend the predictive cross-validation criterion from Preda and Saporta (2005) to the FLMFR context, at expenses of a likely high computational cost (cross-validation on two indexes). Alternatives based on the generalized cross-validation procedure (Cardot et al., 2003) or a stepwise model selection approach based on the BIC criterion could be studied, but neither the degrees of freedom or the likelihood function are immediate to estimate in the FLMFR setup. A feasible possibility, though not regression-driven, is to select  $p$  and  $q$  as the minimum number of components associated with a certain proportion of Explained Variance ( $EV_p$  and  $EV_q$ ), e.g., such that  $EV_p = EV_q = 0.99$ . This simple rule provides an initial selection which can be subsequently improved.

Regularization techniques provide an estimation alternative that, due to their flexibility and efficient computational implementations (Friedman et al., 2010), have been remarkably popular in the last decades. The so-called elastic-net regularization of  $\mathbf{B}_{p,q}$  gives the estimator

$$\hat{\mathbf{B}}_{p,q}^{(\lambda)} = \arg \min_{\mathbf{B}_{p,q}} \left\{ \frac{1}{2n} \sum_{i=1}^n \|(\mathbf{Y}_q)_i - (\mathbf{X}_p \mathbf{B}_{p,q})_i\|^2 + \lambda \left[ \frac{1-\alpha}{2} \|\mathbf{B}_{p,q}\|_F^2 + \alpha \sum_{j=1}^p \|(\mathbf{B}_{p,q})_j\|_2 \right] \right\},$$

where  $\lambda \geq 0$  is the penalty parameter,  $\alpha \in [0, 1]$ ,  $\|\cdot\|_F$  is the Frobenius norm, and  $(\mathbf{A})_i$  stands for the  $i$ -th row of the matrix  $\mathbf{A}$ . If  $\lambda = 0$ , then we the usual FPCR follows. Cases  $\alpha = 0$  and  $\alpha = 1$  correspond to ridge (henceforth denoted as FPCR-L2) and LASSO (FPCR-L1) regression, respectively. The former does a global penalization in all the entries of  $\mathbf{B}_{p,q}$ , whereas the latter applies a row-wise penalization that effectively zeroes full rows, hence removing predictors. Therefore, the key advantage of the FPCR-L1 is that it enables variable selection:  $p$  and  $q$  are initially fixed but only  $\tilde{p} \leq p$  components are selected. On the other hand, FPCR-L2 exhibits an important advantage when employed within the bootstrap algorithm to be described in Section 3.3: the estimation  $\hat{\mathbf{Y}}_q = \mathbf{X}_p \hat{\mathbf{B}}_{p,q}^{(\lambda)}$  can be re-expressed as  $\hat{\mathbf{Y}}_q = \mathbf{H}^{(\lambda)} \mathbf{Y}_q$ , where  $\mathbf{H}^{(\lambda)} = \mathbf{X}_p (\mathbf{X}_p' \mathbf{X}_p + \lambda \mathbf{I}_p)^{-1} \mathbf{X}_p'$  is the hat matrix for the FPCR-L2 estimator. The lack of an analogous result for the FPCR-L1 estimator notably increases the bootstrapping cost. Finally, note that  $\lambda$  can be selected with reasonable efficiency through leave-one-out cross-validation ( $\hat{\lambda}_{CV}$ ), as implemented in Friedman et al. (2010).

As a way to exploit the advantages of both FPCR-L1 and FPCR-L2, we propose a hybrid approach, termed FPCR-L1-selected (FPCR-L1S) estimator, which firstly implements FPCR-L1 for variable selection, and then performs FPCR estimation with the predictors selected by FPCR-L1 (see Remark 7 on variable selection by FPCR-L1). Therefore, FPCR-L1S has a hat matrix that is very convenient for the latter bootstrap algorithm:

$$\mathbf{H}_C^{(\lambda)} = \tilde{\mathbf{X}}_{\tilde{p}} \left( \tilde{\mathbf{X}}_{\tilde{p}}' \tilde{\mathbf{X}}_{\tilde{p}} \right)^{-1} \tilde{\mathbf{X}}_{\tilde{p}}', \quad (9)$$

where  $\tilde{\mathbf{X}}_{\tilde{p}}$  is the matrix of the coefficients of the  $\tilde{p}$  selected predictors (which can be *non-consecutive* FPC). This variable selection is a crucial advantage, as clearly the number of FPC for representing  $\mathcal{X}$  and  $\mathcal{Y}$  up to a certain EV might not correspond with the best selection of  $(p, q)$  for the estimation of  $\mathbf{B}_{p,q}$ . We denote the scores of the FPCR-L1S estimator as  $\hat{\mathbf{B}}_{\tilde{p},q}^{(\lambda),C}$ .

## 2.4 Comparative study of estimators

A succinct simulation study is conducted for comparing the performance of the four estimators previously described. We used the following common settings: the functional covariates  $\{\mathcal{X}_i\}_{i=1}^n$  are centered and valued in  $[0, 1]$ , the functional errors  $\{\mathcal{E}_i\}_{i=1}^n$  are valued in  $[2, 3]$  (both intervals were discretized in 101 equispaced grid points), the sample size is  $n = 100$ , and 1,000 Monte Carlo replicates were considered. The simulation scenarios are collected in Table 1 and have the following descriptions:

- **CM.** Based on the process used in Crambes and Mas (2013), where  $\mathcal{X}(s) = \sum_{j=1}^{50} \lambda_j \varepsilon_j \Psi_j(s)$ ,  $\varepsilon_j \sim \mathcal{N}(0, 2^2)$ , with  $\lambda_j = (\pi^2(j - \frac{1}{2})^2)^{-1}$  and  $\Psi_j(s) = \sqrt{2} \sin((j - 0.5)\pi s)$ , for each  $j \geq 1$  and  $s \in [0, 1]$ .
- **BM.** Brownian motion with standard deviation equal to 0.15.
- **IK.** Based on the process used in Imaizumi and Kato (2018). Functional covariates are given by  $\mathcal{X}(s) = \sum_{j=1}^{50} j^{-7/4} U_j \Psi_j(s)$ ,  $U_j \sim \mathcal{U}(-\sqrt{5}, \sqrt{5})$ , with  $\Psi_1(s) \equiv 1$  and  $\Psi_j(s) = \sqrt{2} \cos(j\pi s)$ , for each  $j \geq 1$  and  $s \in [0, 1]$ . Functional errors are given by  $\mathcal{E}(t) = \sum_{j=1}^{50} j^{-4/5} \varepsilon_j \Psi_j(t)$ ,  $\varepsilon_j \sim \mathcal{N}(0, 1.5^2)$ , for each  $j \geq 1$  and  $t \in [2, 3]$ .
- **GP.** Gaussian process with covariance function  $\Sigma(s_1, s_2) = 6^2 \exp(-|s_1 - s_2|/0.2)$ .
- **OU.** Ornstein–Uhlenbeck process with unitary drift and stationary standard deviation equal to 0.35.

Scenario	Kernel $\beta(s, t)$	$\mathcal{X}(s)$	$\mathcal{E}(t)$
S1	$(s - a)^2 + (t - c)^2$	CM	BM
S2	$2[\sin(6\pi(s - a)) + \cos(6\pi(t - c))]$	GP	OU
S3	$\beta(s, t) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} b_{j,k} \Phi_j(s) \Psi_k(t)$ with $b_{jk} = 0$ if $j, k \leq 4$ , $b_{jk} = 6(-1)^{j+k} (j - 4)^{-12/5} (k - 4)^{-1/4}$ otherwise	IK	IK

Table 1: Summary of the simulated scenarios.

Table 2 shows the averaged errors  $\|\beta - \hat{\beta}^{(p,q)}\|_{\mathbb{H}_1 \otimes \mathbb{H}_2}$  of all estimators for  $p = 2, 5, 10, 25, 50$  and  $q = 1, 5, 10$ , with  $\lambda$  set as  $\hat{\lambda}_{CV}$ . We summarize next the conclusions:

- There is a weak dependency on  $q$ : parameters  $(p, q)$  do not play a symmetric role (Ramsay and Silverman, 2005). Nonetheless, the influence of  $q$  is more prevalent in S2 and S3, inasmuch as an amount of EV has still to be captured.
- When  $p$  is excessively large, errors skyrocket for FPCR and FPCR-L2, in contrast with FPCR-L1 and FPCR-L1S. This is clearly observed in S1 (low variability and a linear kernel), since the model begins to become promptly overfitted ( $EV_{p=2} > 0.99$  and  $EV_{q=1} = 0.98$ ) and the effective variable selection of FPCR-L1 and FPCR-L1S is clearly manifested ( $\bar{p}/p < 0.05$  as  $p$  increases).
- S2 (high variability and an egg-carton-shape-like kernel) illustrates the situation in which the functional samples are not properly represented with few FPC ( $EV_{p=10} < 0.95$ ). Even though errors are smaller than in S1 (overfitting is mitigated,  $\bar{p}/p \simeq 0.25$  as  $p$  increases), FPCR-L1 (mainly) and FPCR-L1S provide more precise estimations. FPCR slightly outperforms the rest of estimators for small values of  $(p, q)$ .
- A sensible choice of  $(p, q)$  for representing the functional samples might not be so for estimating  $\beta$ . This is illustrated in S3: even though  $\mathcal{X}$  and  $\mathcal{Y}$  are smoother than in S2,  $\bar{p}$  is not much smaller, since the first components are not informative. The number of selected FPC for FPCR-L1 and FPCR-L1S is drastically reduced for large values of  $(p, q)$  ( $\bar{p}/p < 0.05$ , when  $p = 50$  and  $q = 10$ ), since non-consecutive FPC are allowed to be selected, removing the noise from estimating the first null components.

All in all, FPCR-L1 outperforms FPCR-L1S, yet both performances are markedly better than the FPCR and FPCR-L2 ones. Because of this and the key computational advantage the explicit hat matrix (9) delivers, we will adopt FPCR-L1S as our reference estimator.

Scenario	S1			S2			S3			
$q$ ( $EV_q$ )	1 (98%)	5 (> 99%)	10 (> 99%)	1 (92%)	5 (> 99%)	10 (> 99%)	1 (38%)	5 (86%)	10 (93%)	
$EV_p$	> 99%			68.69%			96.66%			
$\bar{p}$	1.0 (0.13)	1.0 (0.13)	1.0 (0.13)	1.8 (0.40)	2.0 (0.14)	2.0 (0.14)	1.0 (0.00)	1.0 (0.00)	1.0 (0.00)	
$p = 2$	FPCR	0.303	0.296	0.296	<b>1.438</b>	<b>1.418</b>	<b>1.418</b>	<b>21.382</b>	21.386	21.387
	L1	0.216	0.216	0.216	1.438	1.425	1.425	21.385	<b>21.385</b>	<b>21.385</b>
	L2	0.300	0.291	0.291	1.438	1.423	1.423	21.385	<b>21.385</b>	<b>21.385</b>
	L1S	<b>0.204</b>	<b>0.203</b>	<b>0.203</b>	1.438	1.418	1.418	21.385	<b>21.385</b>	<b>21.385</b>
$EV_p$	> 99%			87.77%			98.33%			
$\bar{p}$	1.1 (0.27)	1.1 (0.28)	1.1 (0.28)	4.8 (0.47)	4.4 (0.70)	4.4 (0.70)	3.1 (1.40)	1.5 (0.83)	1.5 (0.81)	
$p = 5$	FPCR	2.461	2.660	2.670	<b>1.418</b>	<b>1.303</b>	<b>1.304</b>	<b>10.204</b>	<b>6.696</b>	<b>6.738</b>
	L1	<b>0.239</b>	<b>0.242</b>	<b>0.243</b>	1.418	1.326	1.326	10.299	9.182	9.256
	L2	2.161	2.316	2.324	1.418	1.308	1.308	10.417	10.221	10.335
	L1S	0.308	0.323	0.323	1.418	1.307	1.307	10.230	<b>6.711</b>	<b>6.716</b>
$EV_p$	> 99%			94.46%			> 99%			
$\bar{p}$	1.1 (0.45)	1.1 (0.42)	1.1 (0.42)	8.1 (1.28)	8.9 (1.00)	8.8 (1.00)	5.1 (2.52)	1.9 (1.32)	1.9 (1.28)	
$p = 10$	FPCR	15.297	16.411	16.461	<b>1.416</b>	0.504	0.507	9.643	14.313	15.342
	L1	<b>0.404</b>	<b>0.407</b>	<b>0.408</b>	1.416	0.547	0.548	<b>8.981</b>	8.782	8.868
	L2	13.354	14.194	14.236	1.416	<b>0.503</b>	<b>0.506</b>	9.348	12.468	12.912
	L1S	1.193	1.185	1.186	1.416	0.507	0.509	9.175	<b>6.960</b>	<b>6.978</b>
$EV_p$	> 99%			98.37%			> 99%			
$\bar{p}$	1.2 (0.69)	1.2 (0.67)	1.2 (0.67)	11.5 (3.22)	11.7 (3.02)	11.7 (3.01)	5.9 (3.75)	2.0 (1.80)	1.70 (0.45)	
$p = 25$	FPCR	164.917	176.286	176.757	1.419	1.271	1.291	36.794	111.324	119.420
	L1	<b>2.006</b>	<b>2.004</b>	<b>1.986</b>	<b>1.416</b>	<b>0.622</b>	<b>0.622</b>	<b>10.358</b>	<b>10.383</b>	<b>10.290</b>
	L2	142.442	150.485	150.857	1.419	1.222	1.241	26.367	52.843	53.747
	L1S	9.549	10.505	10.435	1.417	0.936	0.943	16.720	15.679	15.310
$EV_p$	> 99%			99%			> 99%			
$\bar{p}$	1.5 (1.74)	1.4 (1.52)	1.4 (1.48)	13.2 (5.50)	13.6 (5.10)	13.6 (5.11)	6.8 (5.30)	2.2 (2.57)	2.2 (2.60)	
$p = 50$	FPCR	1231.590	1313.864	1317.221	1.445	3.596	3.654	220.034	680.661	729.409
	L1	<b>19.933</b>	<b>17.903</b>	<b>17.703</b>	<b>1.418</b>	<b>0.856</b>	<b>0.852</b>	<b>20.103</b>	<b>20.621</b>	<b>19.626</b>
	L2	1045.237	1098.301	1100.604	1.444	3.456	3.510	135.310	212.072	203.919
	L1S	92.410	92.469	91.647	1.429	2.097	2.105	60.360	73.110	68.900

Table 2: Averaged  $L^2$  estimation errors. The average number (sd in parentheses) of selected FPC with FPCR-L1 and FPCR-L1S is denoted as  $\bar{p}$ . Boldfaces denote the errors that are not significantly larger than the smallest (on each block), according to a 95%-confidence paired  $t$ -test.

### 3 A GoF test for the FLMFR

#### 3.1 Derivation of the test statistic

Our aim is to verify whether the relation between the functional response and predictor can be explained by the FLMFR in (6), that is, to test the composite null hypothesis

$$\mathcal{H}_0 : m \in \mathcal{L} = \{ \langle \cdot, \beta \rangle : \beta \in \mathbb{H}_1 \otimes \mathbb{H}_2 \}$$

against an unspecified alternative hypothesis  $\mathcal{H}_1 : \mathbb{P}(m \notin \mathcal{L}) > 0$ . Note that  $\mathcal{H}_0$  is equivalent to  $\mathcal{H}_0 : m(\cdot) = \langle \cdot, \beta \rangle$ , where the equality holds for some *unknown*  $\beta \in \mathbb{H}_1 \otimes \mathbb{H}_2$ .

The following lemmas give the characterization of  $\mathcal{H}_0$  in terms of the one-dimensional projections of the response and the predictor.

**Lemma 1** ( $\mathcal{H}_0$  characterization). *Let  $\mathcal{X}$  and  $\mathcal{Y}$  be  $\mathbb{H}_1$ - and  $\mathbb{H}_2$ -valued random variables, respectively, and  $\beta \in \mathbb{H}_1 \otimes \mathbb{H}_2$ . Then, the following statements are equivalent:*

- i.*  $\mathcal{H}_0$  holds, that is,  $m(x) = \langle x, \beta \rangle, \forall x \in \mathbb{H}_1$ .



ii.  $\mathbb{E}[\mathcal{Y} - \langle \langle \mathcal{X}, \beta \rangle \rangle | \mathcal{X} = x] = 0$ , for almost every (a.e.)  $x \in \mathbb{H}_1$ .

iii.  $\mathbb{E}[\mathcal{Y} - \langle \langle \mathcal{X}, \beta \rangle \rangle | \langle \mathcal{X}, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1} = u] = 0$ , for a.e.  $u \in \mathbb{R}$ ,  $\forall \gamma_{\mathcal{X}} \in \mathbb{S}_{\mathbb{H}_1}$ .

iv.  $\mathbb{E}[\langle \mathcal{Y} - \langle \langle \mathcal{X}, \beta \rangle \rangle, \gamma_{\mathcal{Y}} \rangle_{\mathbb{H}_2} | \langle \mathcal{X}, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1} = u] = 0$  almost surely (a.s.), for a.e.  $u \in \mathbb{R}$  and  $\forall \gamma_{\mathcal{X}} \in \mathbb{S}_{\mathbb{H}_1}, \gamma_{\mathcal{Y}} \in \mathbb{S}_{\mathbb{H}_2}$ .

v.  $\mathbb{E} \left[ \langle \mathcal{Y} - \langle \langle \mathcal{X}, \beta \rangle \rangle, \gamma_{\mathcal{Y}} \rangle_{\mathbb{H}_2} \mathbb{1}_{\{\langle \mathcal{X}, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1} \leq u\}} \right] = 0$  a.s., for a.e.  $u \in \mathbb{R}$  and  $\forall \gamma_{\mathcal{X}} \in \mathbb{S}_{\mathbb{H}_1}, \gamma_{\mathcal{Y}} \in \mathbb{S}_{\mathbb{H}_2}$ .

**Lemma 2** ( $\mathcal{H}_0$  characterization on finite-dimensional directions). *Within the setting of Lemma 1, let  $\{\Psi_j\}_{j=1}^{\infty}$  and  $\{\Phi_k\}_{k=1}^{\infty}$  be bases of  $\mathbb{H}_1$  and  $\mathbb{H}_2$ , respectively. Then, the previous statement v is equivalent to*

v'.  $\mathbb{E} \left[ \langle \mathcal{Y} - \langle \langle \mathcal{X}, \beta \rangle \rangle, \gamma_{\mathcal{Y}} \rangle_{\mathbb{H}_2} \mathbb{1}_{\{\langle \mathcal{X}, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1} \leq u\}} \right] = 0$ , for a.e.  $u \in \mathbb{R}$ ,  $\forall \gamma_{\mathcal{X}} \in \mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^{\infty}}^{p-1}, \gamma_{\mathcal{Y}} \in \mathbb{S}_{\mathbb{H}_2, \{\Phi_k\}_{k=1}^{\infty}}^{q-1}$ , and for all  $p, q \geq 1$ .

Hence,  $\mathcal{H}_0$  holds if and only if v' is satisfied. In addition, the former statements iii–iv are equivalent to their iii'–iv' analogues.

We use the characterization given by v in Lemma 1 to detect deviations from  $\mathcal{H}_0$ . We do so by means of the empirical version (from an iid sample from (2)) of the doubly-projected integrated regression function in v, that is, the residual marked empirical process

$$R_n(u, \gamma_{\mathcal{X}}, \gamma_{\mathcal{Y}}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \langle \hat{\mathcal{E}}_i, \gamma_{\mathcal{Y}} \rangle_{\mathbb{H}_2} \mathbb{1}_{\{\langle \mathcal{X}, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1} \leq u\}}, \quad (10)$$

with  $u \in \mathbb{R}$ ,  $\gamma_{\mathcal{X}} \in \mathbb{S}_{\mathbb{H}_1}$ ,  $\gamma_{\mathcal{Y}} \in \mathbb{S}_{\mathbb{H}_2}$  and with residual marks  $\langle \hat{\mathcal{E}}_i, \gamma_{\mathcal{Y}} \rangle_{\mathbb{H}_2} = \langle \mathcal{Y}_i - \langle \langle \mathcal{X}_i, \hat{\beta} \rangle \rangle, \gamma_{\mathcal{Y}} \rangle_{\mathbb{H}_2}$  and jumps  $\langle \mathcal{X}_i, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1}$ ,  $i = 1, \dots, n$ . To measure how close the empirical process (10) is to zero, and following the ideas in Escanciano (2006) and García-Portugués et al. (2014), we consider a Cramér–von Mises (CvM) norm on the space  $\Pi = \mathbb{S}_{\mathbb{H}_2} \times \mathbb{S}_{\mathbb{H}_1} \times \mathbb{R}$ , yielding what we term the Projected Cramér–von Mises (PCvM) statistic:

$$\text{PCvM}_n = \int_{\Pi} [R_n(u, \gamma_{\mathcal{X}}, \gamma_{\mathcal{Y}})]^2 F_{n, \gamma_{\mathcal{X}}}(\text{d}u) \omega_{\mathcal{X}}(\text{d}\gamma_{\mathcal{X}}) \omega_{\mathcal{Y}}(\text{d}\gamma_{\mathcal{Y}}), \quad (11)$$

where  $F_{n, \gamma_{\mathcal{X}}}$  is the empirical cumulative distribution function (ecdf) of  $\{\langle \mathcal{X}_i, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1}\}_{i=1}^n$ , and  $\omega_{\mathcal{X}}$  and  $\omega_{\mathcal{Y}}$  are suitable measures on  $\mathbb{S}_{\mathbb{H}_1}$  and  $\mathbb{S}_{\mathbb{H}_2}$ , respectively. As will be seen in Section 3.2, a key advantage of the PCvM statistic with respect to other possible norms for (10), such as the Kolmogorov–Smirnov norm, is that it admits an explicit representation.

The infinite dimension of  $\mathbb{S}_{\mathbb{H}_1}$  and  $\mathbb{S}_{\mathbb{H}_2}$  makes the functional in (11) unworkable. A way of circumventing this issue, motivated by Lemma 2, is to work with the finite-dimensional directions  $\gamma_{\mathcal{X}}^{(p)}$  and  $\gamma_{\mathcal{Y}}^{(q)}$  expressed on the bases  $\{\Psi_j\}_{j=1}^p$  and  $\{\Phi_k\}_{k=1}^q$ , respectively. For the sake of simplicity, we assume that these bases are orthonormal from now on; see Remark 3 for non-orthogonal bases. Then, the  $(p, q)$ -truncated version of (10) is

$$R_{n,p,q} \left( u, \gamma_{\mathcal{X}}^{(p)}, \gamma_{\mathcal{Y}}^{(q)} \right) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \langle \hat{\mathcal{E}}_i^{(q)}, \gamma_{\mathcal{Y}}^{(q)} \rangle_{\mathbb{H}_2} \mathbb{1}_{\{\langle \mathcal{X}_i^{(p)}, \gamma_{\mathcal{X}}^{(p)} \rangle_{\mathbb{H}_1} \leq u\}} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\mathbf{e}}'_{i,q} \mathbf{h}_q \mathbb{1}_{\{\mathbf{x}'_{i,p} \mathbf{g}_p \leq u\}},$$

where  $u \in \mathbb{R}$ ,  $\mathbf{g}_p \in \mathbb{S}^{p-1}$ ,  $\mathbf{h}_q \in \mathbb{S}^{q-1}$  and  $\hat{\mathbf{e}}'_{i,q}$  represents the  $i$ -th row of the  $n \times q$  matrix of residual coefficients  $\hat{\mathbf{E}}_q$ ,  $\mathbf{g}_p$  and  $\mathbf{h}_q$  are the coefficients of  $\gamma_{\mathcal{X}}^{(p)}$  and  $\gamma_{\mathcal{Y}}^{(q)}$ , respectively, and  $\mathbf{x}_{i,p}$  are the coefficients of  $\mathcal{X}_i^{(p)}$ . Therefore, the  $(p, q)$ -truncated version of (11) is

$$\text{PCvM}_{n,p,q} = \int_{\Pi^{(p,q)}} \left[ R_{n,p,q} \left( u, \gamma_{\mathcal{X}}^{(p)}, \gamma_{\mathcal{Y}}^{(q)} \right) \right]^2 F_{n, \gamma_{\mathcal{X}}^{(p)}}(\text{d}u) \omega_{\mathcal{X}}(\text{d}\gamma_{\mathcal{X}}^{(p)}) \omega_{\mathcal{Y}}(\text{d}\gamma_{\mathcal{Y}}^{(q)}), \quad (12)$$

where  $\Pi^{(p,q)} = \mathbb{S}_{\mathbb{H}_2, \{\Phi_k\}_{k=1}^{\infty}}^{q-1} \times \mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^{\infty}}^{p-1} \times \mathbb{R}$ .

### 3.2 Computable form of the statistic

The statistic in (12) is now conveniently rewritten for its implementation. First, following Escanciano (2006) and García-Portugués et al. (2014), let us assume that  $\omega_{\mathcal{X}}$  and  $\omega_{\mathcal{Y}}$  in (12) represent uniform measures on  $\mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^{\infty}}^{p-1}$  and  $\mathbb{S}_{\mathbb{H}_2, \{\Phi_k\}_{k=1}^{\infty}}^{q-1}$ , respectively. Second, recall that since both bases are orthonormal, from the transformation defined in (1), we have

$$\text{PCvM}_{n,p,q} = \int_{\mathbb{S}^{q-1} \times \mathbb{S}^{p-1} \times \mathbb{R}} [R_{n,p,q}(u, \mathbf{g}_p, \mathbf{h}_q)]^2 F_{n, \mathbf{g}_p}(du) d\mathbf{g}_p d\mathbf{h}_q, \quad (13)$$

where  $R_{n,p,q}(u, \mathbf{g}_p, \mathbf{h}_q) \equiv R_{n,p,q}(u, \gamma_{\mathcal{X}}^{(p)}, \gamma_{\mathcal{Y}}^{(q)})$ . Using some simple algebra, we obtain

$$\begin{aligned} \text{PCvM}_{n,p,q} &= \int_{\mathbb{S}^{q-1} \times \mathbb{S}^{p-1} \times \mathbb{R}} \frac{1}{n} \left[ \sum_{i=1}^n \hat{\mathbf{e}}'_{i,q} \mathbf{h}_q \mathbb{1}_{\{\mathbf{x}'_{i,p} \mathbf{g}_p \leq u\}} \right]^2 F_{n, \mathbf{g}_p}(du) d\mathbf{g}_p d\mathbf{h}_q \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left[ \int_{\mathbb{S}^{p-1} \times \mathbb{R}} \mathbb{1}_{\{\mathbf{x}'_{i,p} \mathbf{g}_p \leq u\}} \mathbb{1}_{\{\mathbf{x}'_{j,p} \mathbf{g}_p \leq u\}} F_{n, \mathbf{g}_p}(du) d\mathbf{g}_p \right] \int_{\mathbb{S}^{q-1}} (\hat{\mathbf{e}}'_{i,q} \mathbf{h}_q) (\hat{\mathbf{e}}'_{j,q} \mathbf{h}_q) d\mathbf{h}_q \\ &= \frac{1}{n^2} \sum_{i=1}^n \left[ \sum_{j=1}^n \sum_{r=1}^n \int_{\mathbb{S}^{p-1}} \mathbb{1}_{\{(\mathbf{x}_{i,p} - \mathbf{x}_{r,p})' \mathbf{g}_p \leq 0, (\mathbf{x}_{j,p} - \mathbf{x}_{r,p})' \mathbf{g}_p \leq 0\}} d\mathbf{g}_p \right] E_{ij} \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \sum_{r=1}^n \left[ \int_{S_{ijr}} d\mathbf{g}_p \right] E_{ij}, \end{aligned} \quad (14)$$

where we denote  $S_{ijr} := \{\mathbf{z} \in \mathbb{S}^{p-1} : \pi/2 \leq \angle(\mathbf{x}_{i,p} - \mathbf{x}_{r,p}, \mathbf{z}) \leq 3\pi/2, \pi/2 \leq \angle(\mathbf{x}_{j,p} - \mathbf{x}_{r,p}, \mathbf{z}) \leq 3\pi/2\}$  ( $\angle(\mathbf{x}, \mathbf{y})$  stands for the angle between  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$ ) and  $E_{ij} := \int_{\mathbb{S}^{q-1}} (\hat{\mathbf{e}}'_{i,q} \mathbf{h}_q) (\hat{\mathbf{e}}'_{j,q} \mathbf{h}_q) d\mathbf{h}_q$ .

The terms  $\int_{S_{ijr}} d\mathbf{g}_p =: A_{ijr}$  are the same as the ones given in García-Portugués et al. (2014) and they represent surface areas of particular spherical regions, that can either be the whole sphere  $\mathbb{S}^{p-1}$  ( $\mathbf{x}_{i,p} = \mathbf{x}_{j,p} = \mathbf{x}_{r,p}$ ), a hemisphere of  $\mathbb{S}^{p-1}$  (if either  $\mathbf{x}_{i,p} = \mathbf{x}_{j,p}$ ,  $\mathbf{x}_{j,p} = \mathbf{x}_{r,p}$  or  $\mathbf{x}_{i,p} = \mathbf{x}_{r,p}$ ), or a spherical wedge with solid angle

$$\pi - \cos^{-1} \left( \frac{(\mathbf{x}_{i,p} - \mathbf{x}_{r,p})' (\mathbf{x}_{j,p} - \mathbf{x}_{r,p})}{\|\mathbf{x}_{i,p} - \mathbf{x}_{r,p}\| \cdot \|\mathbf{x}_{j,p} - \mathbf{x}_{r,p}\|} \right). \quad (15)$$

Therefore, since the surface area of  $\mathbb{S}^{p-1}$  is equal to  $2\pi^{p/2}/\Gamma(p/2)$ , being  $\Gamma(\cdot)$  the Gamma function, from Escanciano (2006) it follows that

$$A_{ijr} = A_{ijr}^{(\angle)} \frac{\pi^{p/2-1}}{\Gamma(p/2)}, \quad A_{ijr}^{(\angle)} := \begin{cases} 2\pi, & \text{if } \mathbf{x}_{i,p} = \mathbf{x}_{j,p} = \mathbf{x}_{r,p}, \\ \pi, & \text{if } \mathbf{x}_{i,p} \neq \mathbf{x}_{j,p} \text{ and } \mathbf{x}_{i,p} = \mathbf{x}_{r,p} \text{ or } \mathbf{x}_{j,p} = \mathbf{x}_{r,p}, \\ (15), & \text{otherwise.} \end{cases} \quad (16)$$

The term  $E_{ij}$  can be dealt using the next auxiliary lemma regarding integration on the Euclidean sphere, yielding  $E_{ij} = 2\pi^{q/2}/(q\Gamma(q/2)) \hat{\mathbf{e}}'_{i,q} \hat{\mathbf{e}}_{j,q}$ , for each  $i, j = 1, \dots, n$ .

**Lemma 3.** For any vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^q$ ,  $\int_{\mathbb{S}^{q-1}} (\mathbf{x}'\boldsymbol{\omega})(\mathbf{y}'\boldsymbol{\omega}) d\boldsymbol{\omega} = 2\pi^{q/2}/(q\Gamma(q/2)) \mathbf{x}'\mathbf{y}$ .

Substituting these terms into (14), we get an easily computable form of the statistic:

$$\text{PCvM}_{n,p,q} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \sum_{r=1}^n A_{ijr} \frac{2\pi^{q/2}}{q\Gamma(q/2)} \hat{\mathbf{e}}'_{i,q} \hat{\mathbf{e}}_{j,q} = \frac{1}{n^2} \frac{2\pi^{p/2+q/2-1}}{q\Gamma(p/2)\Gamma(q/2)} \text{Tr} \left[ \hat{\mathbf{E}}'_q \mathbf{A}_{\bullet} \hat{\mathbf{E}}_q \right], \quad (17)$$

where  $\text{Tr}(\cdot)$  denotes the trace operator and the elements of the symmetric matrix  $\mathbf{A}_{\bullet}$  are defined as  $(\mathbf{A}_{\bullet})_{ij} := \sum_{r=1}^n A_{ijr}$ , for  $i, j = 1, \dots, n$ .

**Remark 1** (Generalization of the GoF test statistic for the FLMSR). *If  $\mathbb{H}_2 = \mathbb{R}$ , identifiable with the subspace of  $L^2([c, d])$  of constant functions, the FLMSR arises as a particular case of the FLMFR. This is reflected in the statistic (17) which, if  $q = 1$ , yields the PCvM statistic for the FLMSR given in García-Portugués et al. (2014) as a particular case.*

**Remark 2** (Alternative interpretation of  $\text{PCvM}_{n,p,q}$ ). *The statistic (17) can be written as*

$$\text{PCvM}_{n,p,q} = \frac{1}{n^2} \frac{2\pi^{p/2+q/2-1}}{q\Gamma(p/2)\Gamma(q/2)} \sum_{k=1}^q \|(\hat{e}_{1,k}, \dots, \hat{e}_{n,k})\|_{\mathbf{A}_\bullet},$$

where  $\hat{\mathcal{E}}_i^{(q)} = \sum_{k=1}^q \hat{e}_{i,k} \Phi_k$ ,  $i = 1, \dots, n$ , and  $\|\mathbf{v}\|_{\mathbf{A}_\bullet} := (\mathbf{v}' \mathbf{A}_\bullet \mathbf{v})^{1/2}$  is a norm (see Lemma 4) in  $\mathbb{R}^n$  induced by the symmetric matrix  $\mathbf{A}_\bullet$ . Therefore, the statistic is a sum, across the  $q$  dimensions of the truncated response, of the  $\mathbf{A}_\bullet$ -weighted norms of the coefficients of the functional errors on  $\{\Phi_k\}_{k=1}^q$ . If this basis is non-orthonormal, then a similar interpretation can be obtained (see Remark 3).

Observe that  $\|\cdot\|_{\mathbf{A}_\bullet}$  is trivially a semi-norm: since  $\text{PCvM}_{n,p,q}$  is non-negative, then  $\mathbf{A}_\bullet$  must be positive semi-definite. That  $\mathbf{A}_\bullet$  is actually a norm follows from the next lemma.

**Lemma 4.** *Assume that the functional sample  $\{\mathcal{X}_i\}_{i=1}^n$  has pairwise distinct coefficients  $\{\mathbf{x}_{i,p}\}_{i=1}^n$  on an arbitrary  $p$ -truncated basis  $\{\Psi_j\}_{j=1}^p$  of  $\mathbb{H}_1$ . Then, for any sample size  $n \geq 1$ , the  $n \times n$  matrix  $\mathbf{A}_\bullet$  is positive definite.*

**Remark 3** (Statistic for general functional bases). *The statistic in (13) can be expressed in terms of non-orthogonal functional bases as follows:*

$$\begin{aligned} |\mathbf{P}_p| |\mathbf{Q}_q| \text{PCvM}_{n,p,q} &= \int_{\mathbb{S}^{q-1} \times \mathbb{S}^{p-1} \times \mathbb{R}} [R_{n,p,q}(u, \mathbf{P}_p^{-1} \mathbf{g}_p, \mathbf{Q}_q^{-1} \mathbf{h}_q)]^2 F_{n, \mathbf{P}_p^{-1} \mathbf{g}_p}(du) d\mathbf{g}_p d\mathbf{h}_q \\ &= \int_{\mathbb{S}^{q-1} \times \mathbb{S}^{p-1} \times \mathbb{R}} \frac{1}{n} \left[ \sum_{i=1}^n \hat{e}'_{i,q} \mathbf{Q}'_q \mathbf{h}_q \mathbb{1}_{\{\mathbf{x}'_{i,p} \mathbf{P}'_p \mathbf{g}_p \leq u\}} \right]^2 F_{n, \mathbf{P}_p^{-1} \mathbf{g}_p}(du) d\mathbf{g}_p d\mathbf{h}_q, \end{aligned}$$

where  $\Phi = \mathbf{Q}'_q \mathbf{Q}_q$  is the Cholesky decomposition of  $\Phi$  and the second equality stems from  $\langle \mathcal{X}_i^{(p)}, \gamma_{\mathcal{X}}^{(p)} \rangle_{\mathbb{H}_1} = \mathbf{x}'_{i,p} \Psi \mathbf{g}_p$  and  $\langle \hat{\mathcal{E}}_i^{(q)}, \gamma_{\hat{\mathcal{E}}}^{(q)} \rangle_{\mathbb{H}_2} = \hat{e}'_{i,q} \Phi \mathbf{h}_q$ . Then, following the developments preceding (17), it can be shown that

$$\text{PCvM}_{n,p,q} = \frac{1}{n^2} \frac{2\pi^{p/2+q/2-1}}{|\mathbf{P}_p| |\mathbf{Q}_q| q\Gamma(p/2)\Gamma(q/2)} \text{Tr} \left[ (\hat{\mathbf{E}}_q \mathbf{Q}_q)' \mathbf{A}_\bullet (\hat{\mathbf{E}}_q \mathbf{Q}_q) \right], \quad (18)$$

where  $\mathbf{A}_\bullet$  is based on the coefficients of  $\mathcal{X}_1^{(p)}, \dots, \mathcal{X}_n^{(p)}$  on the non-orthonormal basis  $\{\Psi_j\}_{j=1}^p$ .

Despite the general derivation of the PCvM statistic, we will focus on its application for the data-driven FPC bases  $\{\hat{\Psi}_j\}_{j=1}^n$  and  $\{\hat{\Phi}_k\}_{k=1}^n$ .

### 3.3 Testing in practice and bootstrap resampling

We calibrate the null distribution of the statistic  $\text{PCvM}_{n,p,q}$  in (18) by a wild bootstrap on the residuals. This methodology is consistent in the finite dimensional case, as shown by Stute et al. (1998), and well-adapted for heteroscedastic scenarios.

The bootstrap resampling is detailed within the next algorithm. It describes how to perform our GoF test proposal in practice using FPCR-L1S, as this estimator combines the performance of FPCR-L1 and the computational expediency of FPCR. Adaptations to other estimators described in Section 2.3 are straightforward (but see Remark 4 below).

**Algorithm 1** (Testing in practice). Let  $\{(\mathcal{X}_i, \mathcal{Y}_i)\}_{i=1}^n$  be an iid sample. The GoF test for the FLMFR proceeds as follows:

1. Center the sample and compute the FPC of  $\{\mathcal{X}_i\}_{i=1}^n$  and  $\{\mathcal{Y}_i\}_{i=1}^n$ .
2. Select  $p$  and  $q$  as the minimum number of FPC required for attaining a certain proportion of EV (e.g., such that  $\text{EV}_p = \text{EV}_q = 0.99$ ).
3. Compute the coefficients (scores) of  $\{\mathcal{X}_i\}_{i=1}^n$  and  $\{\mathcal{Y}_i\}_{i=1}^n$  on the  $p$ - and  $q$ -truncated FPC bases, resulting the matrices  $\mathbf{X}_p$  and  $\mathbf{Y}_q$ .
4. Compute the FPCR-L1S estimator  $\hat{\mathbf{B}}_{\tilde{p},q}^{(\lambda),C}$  of  $\beta$  as described in Section 2.3. This automatically selects a subset of  $\tilde{p}$  out of  $p$  FPC coefficients, depending on  $\lambda$ .
5. Obtain the residuals  $\hat{\mathbf{e}}_{i,q} = \mathbf{Y}_{i,q} - \mathbf{X}_{i,\tilde{p}} \hat{\mathbf{B}}_{\tilde{p},q}^{(\lambda),C}$ ,  $i = 1, \dots, n$ , and compute with them the statistic  $\text{PCvM}_{n,\tilde{p},q}$  in (17).
6. Perform the bootstrap resampling. For  $b = 1, \dots, B$ :
  - i. Simulate independent zero-mean and unit-variance random variables  $\{V_i^{*b}\}_{i=1}^n$ . For example, sample  $V^{*b}$  such that  $\mathbb{P}[V^{*b} = (1 \mp \sqrt{5})/2] = (5 \pm \sqrt{5})/10$ .
  - ii. Set the bootstrap errors as  $\mathbf{e}_{i,q}^{*b} := \hat{\mathbf{e}}_{i,q} V_i^{*b}$ ,  $i = 1, \dots, n$ .
  - iii. Set the uncentered bootstrapped responses  $\mathbf{Y}_{i,q}^{*b,u} := \mathbf{X}_{i,\tilde{p}} \hat{\mathbf{B}}_{\tilde{p},q}^{(\lambda),C} + \mathbf{e}_{i,q}^{*b}$ , and center them to imitate the original FPC scores:  $\mathbf{Y}_{i,q}^{*b} := \mathbf{Y}_{i,q}^{*b,u} - \overline{\mathbf{Y}_q^{*b,u}}$ ,  $i = 1, \dots, n$ .
  - iv. From the bootstrap sample  $\{(\mathbf{X}_{i,\tilde{p}}, \mathbf{Y}_{i,q}^{*b})\}_{i=1}^n$ , compute the estimator  $\hat{\mathbf{B}}_{\tilde{p},q}^{*b}$  of  $\hat{\mathbf{B}}_{\tilde{p},q}^{(\lambda),C}$ .
  - v. Obtain the bootstrap residuals  $\hat{\mathbf{e}}_{i,q}^{*b} = \mathbf{Y}_{i,q}^{*b} - \mathbf{X}_{i,\tilde{p}} \hat{\mathbf{B}}_{\tilde{p},q}^{*b}$ ,  $i = 1, \dots, n$ , and compute with them the bootstrapped statistic  $\text{PCvM}_{n,\tilde{p},q}^{*b}$  from (17).
7. Estimate the  $p$ -value by Monte Carlo as usual by  $\#\{\text{PCvM}_{n,\tilde{p},q} \leq \text{PCvM}_{n,\tilde{p},q}^{*b}\}/B$ .

**Remark 4** (Computational tricks). Since  $\mathbf{A}_\bullet$  depends exclusively on the covariate sample, it only needs to be computed once in the testing procedure. In addition, as the wild bootstrap only affects the response, steps iv–v can be efficiently implemented using the hat matrix (9), avoiding costly refittings on each bootstrap iteration. Indeed,  $\hat{\mathbf{E}}_q^{*b} = \mathbf{Y}_q^{*b} - \hat{\mathbf{Y}}_q^{*b} = (\mathbf{I}_q - \mathbf{H}_C^{(\lambda)}) \mathbf{Y}_q^{*b}$ ,  $\hat{\mathbf{Y}}_q^{*b} = \mathbf{X}_{\tilde{p}} \hat{\mathbf{B}}_{\tilde{p},q}^{*b}$ . The same comment holds for FPCR-L2 and FPCR by virtue of  $\mathbf{H}^{(\lambda)}$  (in that case,  $\tilde{p} = p$ ), although not for FPCR-L1 due to its lack of an explicit hat matrix. The GoF test using FPCR-L1 thus requires  $B+1$  LASSO fits.

**Remark 5** (Scores versus functional resampling). The above wild bootstrap performs the resampling on the scores of the residuals in the  $q$ -truncated FPC basis  $\{\hat{\Phi}_k\}_{k=1}^q$ , as from step 4 onwards there is no further mention to the functional nature of the sample. This view could be achieved with extra notation, as the bootstrap errors in step ii can be written as

$$\mathcal{E}_i^{(q)*b} := \sum_{k=1}^q (e_{i,k} V_i^{*b}) \hat{\Phi}_k, \quad \hat{\mathcal{E}}_i^{(q)} = \sum_{k=1}^q e_{i,k} \hat{\Phi}_k.$$

This exposes a subtle point: why not bootstrapping the functional residuals  $\hat{\mathcal{E}}_i = \mathcal{Y}_i - \hat{\mathcal{Y}}_i^{(q)}$  as  $\mathcal{E}_i^{*b} := \hat{\mathcal{E}}_i V_i^{*b}$ ? This would allow to obtain truly functional bootstrap responses  $\{\mathcal{Y}_i^{*b}\}_{i=1}^n$ , yet at expenses of the overhead of recomputing their FPC for each bootstrap replicate. In our experiments, this latter approach did not provide a significant improvement on the calibration of the test over the scores resampling, hence it was discarded in favor of the latter.

**Remark 6** (Selection of the penalty parameter). *A possible data-driven selection for  $\lambda$  in step 4 is  $\hat{\lambda}_{CV}$ . However, we found by simulations that the so-called one standard error rule  $\hat{\lambda}_{1SE}$  (see, e.g., Friedman et al. (2010)) improved the stability of the calibration of  $PCvM_{n,\tilde{p},q}$  under  $\mathcal{H}_0$ . This is coherent with the folklore in smoothing-based GoF tests, where the optimal smoothing parameter for estimating the regression function  $m$  is often not the most appropriate for conducting the test; instead, an oversmoothed estimate of  $m$  (that biases the estimation in exchange for a variance reduction, precisely as  $\hat{\lambda}_{1SE}$  does) is desirable for a better calibration of the statistic.*

**Remark 7** (FPCR-L1 variable selection). *LASSO is a consistent variable selector if the predictors are uncorrelated (Zhao and Yu, 2006). Hence, this result supports the adequateness for combining FPCR (instead of using non-orthogonal bases) with LASSO variable-selection. It also supports ignoring in Algorithm 1 the bootstrapping of the variable selection uncertainty. Indeed, motivated by a comment of one referee, we ran a small simulation study with a modified version of Algorithm 1 that incorporated in Step 6 a bootstrap variable selection using  $\hat{\lambda}_{1SE}^{*b}$ , obtaining very similar powers to the analogs of Tables 8–10 for FPCR-L1S ( $\hat{\lambda}_{1SE}$ ). Clearly, this modification increases the computational requirements by orders of magnitude, which is impractical. In this simulation it was also evidenced that variable selection based on  $\hat{\lambda}_{CV}$  seems to be inconsistent (which may be explained by Shao (1993)’s result), while variable selection based on  $\hat{\lambda}_{1SE}$  seems to behave consistently.*

So far we have only discussed the GoF test for the FLMFR. However, simple adaptations allow to test also the simple hypothesis  $\mathcal{H}_0 : m(\cdot) = \langle \langle \cdot, \beta_0 \rangle \rangle$ , where  $\beta_0 \in \mathbb{H}_1 \otimes \mathbb{H}_2$  now is specified. Algorithm 1 can be straightforwardly adapted. First, replace step 4 by

4'. Compute  $\tilde{p}$  as in step 4. Obtain  $\mathbf{B}_{\tilde{p},q}^0 = (b_{ij}^0)_{ij}$ , the  $\tilde{p} \times q$  matrix of  $\beta_0$  FPC coefficients.

Then, the bootstrap procedure is subsequently adjusted by simply ignoring the estimation steps, that is, by replacing both  $\hat{\mathbf{B}}_{\tilde{p},q}^{(\lambda),C}$  and  $\hat{\mathbf{B}}_{\tilde{p},q}^{*b}$  by  $\mathbf{B}_{\tilde{p},q}^0$ .

Algorithm 1 and its variants (simple hypothesis; FPCR, FPCR-L2, and FPCR-L1 estimators; functional residual resampling) are implemented in the companion R package `goffda` (García-Portugués and Álvarez-Liébana, 2020). The critical parts of the test, such as the computation of the  $\mathbf{A}_\bullet$  matrix and the computation of the PCvM statistic (whose complexity is  $\mathcal{O}(q(n^3 - n^2)/2)$ ), are implemented in C++ for the sake of efficiency.

## 4 Simulation study

The finite sample behaviour of the PCvM test is now illustrated via a comparative study with the available significance tests (Section 4.1) and a simulation study for the composite hypothesis (Section 4.2). We employed the scenarios already described in Table 1 and used the following common settings: discretization of functional samples in 101 equispaced grid points along the domains, sample sizes  $n = 50, 100, 250$ ,  $B = 1,000$ , and 1,000 Monte Carlo replicates. The PCvM test was run using Algorithm 1 with  $EV_p = EV_q = 0.99$ .

The PCvM test was computed using both FPCR and FPCR-L1S, for showing how the overfitting inherent to the former may affect the GoF test. In Section 4.2, FPCR-L1S is employed with both  $\hat{\lambda}_{CV}$  and  $\hat{\lambda}_{1SE}$  for the purpose of illustrating the discussion in Remark 6. When testing for significance, the conclusions reached with both penalty parameters were similar (since an estimator of  $\beta$  is not required), so the results are only reported for  $\hat{\lambda}_{1SE}$ . The search for  $\hat{\lambda}_{CV}$  and  $\hat{\lambda}_{1SE}$  was done among a sequence in  $[10^{-3}, 10^2]$  and, if the minimizer of the objective function was found at its extremes, the interval was expanded.

## 4.1 Simple hypothesis

We compare in this section the significance tests by Kokoszka et al. (2008), Patilea et al. (2016b), and Lee et al. (2020) (henceforth abbreviated as KMSZ, PSS, and LZS, respectively) with our PCvM test for the no effects hypothesis

$$\mathcal{H}_{0,NE} : m(\cdot) = \langle \cdot, \beta_0 \rangle, \quad \beta_0 \equiv 0.$$

Both the KMSZ and PSS tests are based on the FPC of the predictor and response, that are truncated such that  $EV_p = EV_q = 0.99$ . The KMSZ statistic is asymptotically  $\chi_{pq}^2$  distributed under  $\mathcal{H}_{0,NE}$ , this being the distribution employed to calibrate the test. We ran the PSS test as implemented in the `fdapss` (Patilea et al., 2016a) package, with a grid of 50 points for each one-dimensional optimization and the bandwidth chosen as  $h = n^{-2/9}$ , as suggested in Patilea et al. (2016b). A bug in `pss.test` when  $p = 1$  invalidated up to 8.5% of the Monte Carlo replicates, depending on the scenario.

The LZS test estimates the functional martingale difference divergence that characterizes the conditional mean dependence of  $\mathcal{X}$  and  $\mathcal{Y}$ . Hence, unlike the previous competitors and our approach, it does not require from an FPC-based dimension reduction.

Notation	Model	$(\delta_1, \delta_2, \delta_3)$
$\mathcal{H}_{0,NE}$ (no effects)	$\mathcal{Y}(t) = \mathcal{E}(t)$	None
$\mathcal{H}_{1,FR}^h$ (FLMFR)	$\mathcal{Y}(t) = \delta_h \langle \mathcal{X}, \beta \rangle + \mathcal{E}(t)$	S1: (0.035, 0.08, 0.15) S2: (0.01, 0.02, 0.03) S3: (1, 1.3, 1.6)
$\mathcal{H}_{1,C}^h$ (FLCFR)	$\mathcal{Y}(t) = \delta_h \tilde{\beta}_j(t) \mathcal{X}(t) + \mathcal{E}(t), j = 1, 2, 3$	S1: (0.025, 0.05, 0.15) S2: (0.2, 0.6, 1) S3: (0.01, 0.025, 0.05)
$\mathcal{H}_{1,NLQ}^h$ (non linear, quadratic)	$\mathcal{Y}(t) = \delta_h \Delta(\mathcal{X})(t) + \mathcal{E}(t)$ $\Delta(\mathcal{X})(t) = \mathcal{X}^2(a + (t - c)\frac{b-a}{d-c}) - 1$	S1: (0.025, 0.075, 0.15) S2: (0.02, 0.04, 0.1) S3: (0.2, 0.35, 0.55)
$\mathcal{H}_{1,NLT}^h$ (non linear, trigonometric)	$\mathcal{Y}(t) = \delta_h \Delta(\mathcal{X})(t) + \mathcal{E}(t)$ $\Delta(\mathcal{X})(t) = (\sin(2\pi t) - \cos(2\pi t)) \ \mathcal{X}\ _{\mathbb{H}_1}^2$	S1: (0.025, 0.075, 0.15) S2: (0.02, 0.04, 0.1) S3: (0.2, 0.35, 0.55)

Table 3: Summary of null and alternative hypotheses. Concurrent models are given by functions  $\tilde{\beta}_1(t) = \sqrt{|\sin(\pi t) - \cos(\pi t)|}$  (S1),  $\tilde{\beta}_2(t) = \log(t - a + 0.5)$  (S2), and  $\tilde{\beta}_3(t) = (t - 0.5)^3$  (S3).

We assume here that  $\mathbb{H}_1 = \mathbb{H}_2 = L^2([0, 1])$ . As reflected in Table 3, four kind of deviations from  $\mathcal{H}_{0,NE}$  were generated: FLMFR, concurrent model (degenerated FLMFR, denoted as FLCFR), and two nonlinear alternatives. The empirical rejection rates are given in Tables 4–6. They contain only the results of the FPCR-based PCvM test since the FPCR-L1S version gave almost identical rejection rates. Their analysis reveals the following insights:

- Regarding the calibration, the PCvM and LZS tests are the only without repeated miscalibrations in any scenario: an over-rejection happens in S2 ( $\mathcal{H}_{0,NE}$  in Table 5) for the PSS test, while the KMSZ test has difficulties in S2 and S3 ( $\mathcal{H}_{0,NE}$  in Tables 5–6).
- Concerning linear alternatives (FLMFR and concurrent), the LZS and KMSZ tests seem to be the most powerful in S1, but the KMSZ test is notably the most powerful approach in S2 under the FLMFR alternative, an outcome somehow expected given the test nature. However, the KMSZ test may fail under linear alternatives for sparse scenarios ( $\mathcal{H}_{1,FR}$  and  $\mathcal{H}_{1,C}$  in Table 6), providing empirical powers smaller than the nominal level. The behaviour is worse for the LZS test, even under larger deviations from  $\mathcal{H}_{0,NE}$ . A possible explanation is that the noise introduced with the null FPCs is not removed due to the lack of dimension reduction in the

test. With respect to the comparison of the PCvM and PSS tests, the former is more powerful than the latter under concurrent models ( $\mathcal{H}_{1,C}$  in Tables 4–6) in all scenarios and for all sample sizes. In the case of FLMFR alternatives, this is also the case (unless for minor exceptions) for S1 and S2 ( $\mathcal{H}_{1,FR}$  in Tables 4–5). In S3, the PSS test attains perfect empirical power, even for  $n = 50$  and the smallest deviation from the null hypothesis, manifesting a sharp difference with respect to its behavior for S2 (almost blind for linear alternatives).

- Concerning nonlinear alternatives, as expected, KMSZ exhibits a poor performance detecting them, except for S2 under  $\mathcal{H}_{1,NLQ}$ . The PSS, LZS, and PCvM tests correctly detect all the nonlinear alternatives, the former being on overall more powerful in S3, the second one in S1 and S2.

$n$	KMSZ			PSS			LZS			PCvM		
	50	100	250	50	100	250	50	100	250	50	100	250
$\mathcal{H}_{0,NE}$	<b>0.053</b>	<b>0.055</b>	<b>0.057</b>	<b>0.049</b>	<b>0.047</b>	<b>0.048</b>	<b>0.048</b>	<b>0.044</b>	<b>0.049</b>	<b>0.042</b>	0.034	<b>0.050</b>
$\mathcal{H}_{1,FR}^1$	0.083	<b>0.178</b>	<b>0.495</b>	0.067	0.074	0.160	0.339	0.567	0.943	<b>0.087</b>	0.128	0.282
$\mathcal{H}_{1,FR}^2$	<b>0.384</b>	<b>0.836</b>	<b>1.000</b>	0.177	0.316	0.718	0.916	0.998	1.000	0.292	0.516	0.923
$\mathcal{H}_{1,FR}^3$	<b>0.955</b>	<b>1.000</b>	<b>1.000</b>	0.551	0.885	0.997	1.000	1.000	1.000	0.718	0.973	<b>1.000</b>
$\mathcal{H}_{1,C}^1$	<b>0.146</b>	<b>0.378</b>	<b>0.890</b>	0.066	0.069	0.152	0.339	0.576	0.946	0.082	0.121	0.272
$\mathcal{H}_{1,C}^2$	<b>0.527</b>	<b>0.936</b>	<b>1.000</b>	0.113	0.195	0.472	0.822	0.989	1.000	0.171	0.340	0.778
$\mathcal{H}_{1,C}^3$	<b>0.976</b>	<b>1.000</b>	<b>1.000</b>	0.708	0.969	<b>1.000</b>	1.000	1.000	1.000	0.511	0.864	<b>1.000</b>
$\mathcal{H}_{1,NLQ}^1$	0.050	<b>0.065</b>	0.070	0.052	0.061	<b>0.116</b>	0.113	0.218	0.677	<b>0.053</b>	0.045	0.074
$\mathcal{H}_{1,NLQ}^2$	0.125	0.171	0.168	<b>0.143</b>	<b>0.362</b>	<b>0.876</b>	0.581	0.970	1.000	0.086	0.171	0.686
$\mathcal{H}_{1,NLQ}^3$	0.246	0.274	0.255	<b>0.553</b>	<b>0.959</b>	<b>1.000</b>	0.876	1.000	1.000	0.233	0.721	<b>1.000</b>
$\mathcal{H}_{1,NLT}^1$	<b>0.100</b>	<b>0.135</b>	<b>0.129</b>	0.050	0.050	0.059	0.093	0.133	0.502	0.047	0.039	0.064
$\mathcal{H}_{1,NLT}^2$	<b>0.194</b>	<b>0.217</b>	0.196	0.068	0.132	<b>0.791</b>	0.632	0.987	1.000	0.080	0.107	0.483
$\mathcal{H}_{1,NLT}^3$	0.217	0.237	0.216	<b>0.446</b>	<b>0.949</b>	<b>1.000</b>	0.932	1.000	1.000	0.245	0.743	<b>1.000</b>

Table 4: Scenario S1. Empirical rejection rates for the KMSZ, PSS, LZS, and PCvM tests for  $n = 50, 100, 250$  and the deviations in Table 3. Under  $\mathcal{H}_{0,NE}$ , the rejection rates are boldfaced if they lie in the 95%-confidence interval of the nominal level, 0.05. Under  $\mathcal{H}_1$ , boldfaces denote the empirical powers that are not significantly smaller than the largest, for each deviation and sample size, according to a 95%-confidence paired  $t$ -test.

$n$	KMSZ			PSS			LZS			PCvM		
	50	100	250	50	100	250	50	100	250	50	100	250
$\mathcal{H}_{0,NE}$	0.006	0.033	<b>0.043</b>	0.093	0.070	0.068	<b>0.054</b>	<b>0.048</b>	<b>0.047</b>	0.030	<b>0.036</b>	<b>0.045</b>
$\mathcal{H}_{1,FR}^1$	0.025	<b>0.201</b>	<b>0.932</b>	<b>0.091</b>	0.078	0.064	0.064	0.064	0.131	0.036	0.056	0.107
$\mathcal{H}_{1,FR}^2$	0.058	<b>0.521</b>	<b>1.000</b>	<b>0.094</b>	0.078	0.057	0.125	0.259	0.952	0.065	0.168	0.900
$\mathcal{H}_{1,FR}^3$	0.083	0.657	<b>1.000</b>	0.095	0.075	0.058	0.342	0.846	1.000	<b>0.180</b>	<b>0.729</b>	<b>1.000</b>
$\mathcal{H}_{1,C}^1$	0.023	<b>0.112</b>	<b>0.567</b>	<b>0.087</b>	0.085	0.059	0.868	0.999	1.000	0.032	0.056	0.104
$\mathcal{H}_{1,C}^2$	<b>0.120</b>	<b>0.874</b>	<b>1.000</b>	0.082	0.092	0.073	1.000	1.000	1.000	0.059	0.176	0.655
$\mathcal{H}_{1,C}^3$	<b>0.955</b>	<b>1.000</b>	<b>1.000</b>	0.093	0.070	0.070	1.000	1.000	1.000	0.381	0.898	<b>1.000</b>
$\mathcal{H}_{1,NLQ}^1$	0.050	<b>0.174</b>	<b>0.305</b>	<b>0.080</b>	0.072	0.156	0.102	0.167	0.626	0.043	0.082	0.282
$\mathcal{H}_{1,NLQ}^2$	<b>0.083</b>	<b>0.337</b>	0.552	0.077	0.196	0.894	0.225	0.648	1.000	0.060	0.227	<b>0.975</b>
$\mathcal{H}_{1,NLQ}^3$	0.084	0.420	0.689	<b>0.250</b>	<b>0.983</b>	<b>1.000</b>	0.502	0.989	1.000	0.086	0.532	<b>1.000</b>
$\mathcal{H}_{1,NLT}^1$	0.007	0.039	0.041	<b>0.074</b>	<b>0.083</b>	0.047	0.098	0.145	0.486	0.039	0.067	<b>0.190</b>
$\mathcal{H}_{1,NLT}^2$	0.010	0.044	0.046	0.067	0.131	0.767	0.253	0.655	1.000	<b>0.069</b>	<b>0.244</b>	<b>0.961</b>
$\mathcal{H}_{1,NLT}^3$	0.010	0.042	0.067	<b>0.385</b>	<b>0.998</b>	<b>1.000</b>	0.625	0.997	1.000	0.180	0.758	<b>1.000</b>

Table 5: Scenario S2. The description of Table 4 applies.

$n$	KMSZ			PSS			LZS			PCvM		
	50	100	250	50	100	250	50	100	250	50	100	250
$\mathcal{H}_{0,NE}$	0.006	<b>0.036</b>	0.026	<b>0.046</b>	0.071	<b>0.052</b>	<b>0.054</b>	<b>0.049</b>	<b>0.040</b>	<b>0.047</b>	<b>0.041</b>	<b>0.037</b>
$\mathcal{H}_{1,FR}^1$	0.010	0.041	0.052	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	0.058	0.073	0.107	0.055	0.108	0.398
$\mathcal{H}_{1,FR}^2$	0.014	0.040	0.056	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	0.063	0.079	0.134	0.062	0.119	0.582
$\mathcal{H}_{1,FR}^3$	0.027	0.044	0.058	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	0.057	0.076	0.135	0.067	0.136	0.675
$\mathcal{H}_{1,C}^1$	0.007	0.037	0.047	0.057	0.099	0.110	0.054	0.049	0.040	<b>0.069</b>	<b>0.120</b>	<b>0.217</b>
$\mathcal{H}_{1,C}^2$	0.008	0.057	0.185	0.141	0.293	0.589	0.054	0.049	0.040	<b>0.252</b>	<b>0.503</b>	<b>0.891</b>
$\mathcal{H}_{1,C}^3$	0.020	0.234	0.870	0.459	0.781	0.998	0.055	0.049	0.040	<b>0.756</b>	<b>0.979</b>	<b>1.000</b>
$\mathcal{H}_{1,NLQ}^1$	0.004	0.027	0.031	<b>0.061</b>	<b>0.132</b>	<b>0.374</b>	0.083	0.100	0.261	0.059	0.080	0.197
$\mathcal{H}_{1,NLQ}^2$	0.006	0.030	0.033	<b>0.120</b>	<b>0.408</b>	<b>0.956</b>	0.139	0.296	0.897	0.096	0.200	0.824
$\mathcal{H}_{1,NLQ}^3$	0.007	0.035	0.036	<b>0.349</b>	<b>0.903</b>	<b>1.000</b>	0.309	0.785	1.000	0.201	0.627	<b>1.000</b>
$\mathcal{H}_{1,NLT}^1$	0.005	0.028	0.034	<b>0.054</b>	<b>0.082</b>	<b>0.178</b>	0.073	0.082	0.199	0.052	0.070	0.156
$\mathcal{H}_{1,NLT}^2$	0.005	0.028	0.033	0.077	<b>0.252</b>	<b>0.986</b>	0.131	0.253	0.940	<b>0.082</b>	0.177	0.816
$\mathcal{H}_{1,NLT}^3$	0.008	0.032	0.030	<b>0.345</b>	<b>0.973</b>	<b>1.000</b>	0.344	0.873	1.000	0.207	0.700	<b>1.000</b>

Table 6: Scenario S3. The description of Table 4 applies.

We report some illustrative average running times of the four tests when  $n = 100$  and  $B = 1,000$ . We do so only for S3, whose running times for all the tests are approximately between S1 and S2, and under  $\mathcal{H}_{0,NE}$  and  $\mathcal{H}_{1,FR}^3$  (similar results were obtained under other alternatives). For the KMSZ test (does not requires bootstrap calibration), the average running times (in seconds) were 0.0086s ( $\mathcal{H}_{0,NE}$ ) and 0.0085s ( $\mathcal{H}_{1,FR}^3$ ). For the PSS and LZS tests, 24.6s and 17.5s, and 0.5s and 0.4s, respectively. For the PCvM test (employs the same estimator as PSS), 0.5s and 0.2s. The comparison was done in a core with 1.8 GHz.

As a conclusion, in the considered scenarios, the PCvM test properly calibrates  $\mathcal{H}_{0,NE}$ , is competitive against the competing tests for all the alternatives (eventually being the most powerful in certain of them), and matches or improves the omnibus LZS and PSS tests in computational expediency.

## 4.2 Composite hypothesis

We consider now  $\mathbb{H}_1 = L^2([0, 1])$  and  $\mathbb{H}_2 = L^2([2, 3])$  and two different null (linear) hypotheses: no effects model and FLMFR. The same two nonlinear deviations from the linearity, weighted by different intensity parameters, are again considered as alternatives. Table 7 summarizes all the hypothesis tested. The conclusions from the results collected in Tables 8–10 are the following:

- As argued in Remark 6, for the PCvM test based in FPCR-L1S,  $\hat{\lambda}_{1SE}$  provides better calibration of the null hypothesis than  $\hat{\lambda}_{CV}$ . The latter statistic encounters serious difficulties to be calibrated, specially in S2–S3 (Tables 9–10) and under  $\mathcal{H}_{0,NE}$ .
- The PCvM test based on FPCR over-rejects under irregular/sparse scenarios like S2 and S3 ( $\mathcal{H}_{0,NE}$  and  $\mathcal{H}_{0,FR}$  in Table 9;  $\mathcal{H}_{0,FR}$  in Table 10). In the case of S2, this phenomena likely arises from the overfitting (already discussed in Section 2.4) associated with the FPCR estimator. For S3, the first scores for estimating  $\beta$  are null coefficients, and therefore, the information coming from the FPC (incorrectly) suggests that  $\mathcal{H}_{0,FR}$  is related to a null surface (i.e., FPC suggest that  $\mathcal{H}_{0,NE}$  holds) and so rejection of  $\mathcal{H}_{0,FR}$  happens. This issue was the main motivation for developing FPCR-L1S and use it as a flexible estimator of  $\beta$  within the PCvM test.
- With respect to the power, the referred over-rejection of the FPCR-based PCvM test unfairly provides greater empirical powers to this test with respect to FPCR-L1S based tests. Concerning the use of  $\hat{\lambda}_{CV}$ , only marginal advantages are provided by  $\hat{\lambda}_{CV}$  in specific situations. Finally, as expected, empirical powers tends to one as  $n$  and the deviation index  $h$  increase.



Notation	Model	$(\delta_1, \delta_2, \delta_3)$
$\mathcal{H}_{0,NE}$ (no effects)	$\mathcal{Y}(t) = \mathcal{E}(t)$	None
$\mathcal{H}_{0,FR}$ (FLMFR)	$\mathcal{Y}(t) = \frac{1}{2}\langle\langle\mathcal{X}, \beta\rangle\rangle + \mathcal{E}(t)$	
$\mathcal{H}_{1,NLQ}^h$ (non linear, quadratic)	$\mathcal{Y}(t) = \langle\langle\mathcal{X}, \beta\rangle\rangle + \delta_h \Delta(\mathcal{X})(t) + \mathcal{E}(t)$ $\Delta(\mathcal{X})(t) = (\mathcal{X}^2(a + (t - c)\frac{b-a}{d-c}) - 1)$	S1: (0.02, 0.04, 0.1) S2: (0.01, 0.02, 0.03) S3: (0.02, 0.15, 0.5)
$\mathcal{H}_{1,NLT}^h$ (non linear, trigonometric)	$\mathcal{Y}(t) = \langle\langle\mathcal{X}, \beta\rangle\rangle + \delta_h \Delta(\mathcal{X})(t) + \mathcal{E}(t)$ $\Delta(\mathcal{X})(t) = (\sin(2\pi t) - \cos(2\pi t)) \ \mathcal{X}\ _{\mathbb{H}_1}^2$	S1: (0.03, 0.05, 0.1) S2: (0.035, 0.045, 0.055) S3: (0.025, 0.2, 0.45)

Table 7: Summary of null and alternative hypotheses, for S1–S3.

$n$	FPCR			FPCR-L1S ( $\hat{\lambda}_{1SE}$ )			FPCR-L1S ( $\hat{\lambda}_{CV}$ )		
	50	100	250	50	100	250	50	100	250
$\mathcal{H}_{0,NE}$	<b>0.041</b>	<b>0.042</b>	<b>0.049</b>	<b>0.043</b>	0.031	<b>0.050</b>	0.010	0.014	0.013
$\mathcal{H}_{0,FR}$	<b>0.042</b>	<b>0.043</b>	<b>0.050</b>	0.028	<b>0.046</b>	<b>0.045</b>	0.030	<b>0.037</b>	<b>0.042</b>
$\mathcal{H}_{1,NLQ}^1$	0.063	0.091	0.177	0.045	0.092	0.179	0.046	0.080	0.166
$\mathcal{H}_{1,NLQ}^2$	0.142	0.271	0.626	0.122	0.254	0.620	0.115	0.244	0.605
$\mathcal{H}_{1,NLQ}^3$	0.596	0.929	1.000	0.566	0.917	1.000	0.568	0.919	1.000
$\mathcal{H}_{1,NLT}^1$	0.048	0.057	0.115	0.035	0.059	0.120	0.037	0.050	0.106
$\mathcal{H}_{1,NLT}^2$	0.087	0.166	0.642	0.068	0.155	0.623	0.068	0.141	0.608
$\mathcal{H}_{1,NLT}^3$	0.555	0.953	1.000	0.496	0.943	1.000	0.505	0.941	1.000

Table 8: Scenario S1. Empirical rejection rates for the PCvM test, based on FPCR and FPCR-L1S, for  $n = 50, 100, 250$ . Under  $\mathcal{H}_{0,NE}$  and  $\mathcal{H}_{0,FR}$ , the rejection rates are boldfaced if they lie in the 95%-confidence interval of the nominal level, 0.05.

$n$	FPCR			FPCR-L1S ( $\hat{\lambda}_{1SE}$ )			FPCR-L1S ( $\hat{\lambda}_{CV}$ )		
	50	100	250	50	100	250	50	100	250
$\mathcal{H}_{0,NE}$	1.000	0.949	0.364	0.026	<b>0.036</b>	<b>0.043</b>	0.014	0.020	0.014
$\mathcal{H}_{0,FR}$	0.997	0.876	0.308	0.091	<b>0.047</b>	<b>0.037</b>	0.298	0.109	<b>0.047</b>
$\mathcal{H}_{1,NLQ}^1$	0.978	0.825	0.648	0.101	0.125	0.380	0.351	0.250	0.390
$\mathcal{H}_{1,NLQ}^2$	0.993	0.962	0.990	0.235	0.463	0.929	0.504	0.593	0.934
$\mathcal{H}_{1,NLQ}^3$	0.999	1.000	1.000	0.585	0.910	1.000	0.844	0.969	1.000
$\mathcal{H}_{1,NLT}^1$	0.979	0.772	0.390	0.120	0.214	0.313	0.438	0.277	0.238
$\mathcal{H}_{1,NLT}^2$	0.996	0.970	0.991	0.650	0.891	0.985	0.772	0.899	0.985
$\mathcal{H}_{1,NLT}^3$	1.000	1.000	1.000	0.910	0.975	1.000	0.967	0.995	1.000

Table 9: Scenario S2. The description of Table 8 applies.

As before, we report some illustrative average timings for S3 under the same conditions. For the PCvM-FPCR test, the timings were 0.8s ( $\mathcal{H}_{0,FR}$ ) and 0.8s ( $\mathcal{H}_{1,NLQ}$ ). The PCvM-FPCR-L1S ( $\hat{\lambda}_{1SE}$ ) test took 13.1s and 11.9s, and the  $\hat{\lambda}_{CV}$  variant, 11.5s and 8.8s.

As a conclusion, the obtained empirical results evidence that the PCvM test based on FPCR-L1S with  $\lambda$  selected by  $\hat{\lambda}_{1SE}$  is a well-calibrated, flexible, and computationally efficient test that is consistent against a broad class of alternatives to the FLMFR.

$n$	FPCR			FPCR-L1S ( $\hat{\lambda}_{1SE}$ )			FPCR-L1S ( $\hat{\lambda}_{CV}$ )		
	50	100	250	50	100	250	50	100	250
$\mathcal{H}_{0,NE}$	<b>0.042</b>	<b>0.047</b>	<b>0.046</b>	<b>0.045</b>	<b>0.043</b>	<b>0.037</b>	0.003	0.008	0.006
$\mathcal{H}_{0,FR}$	0.082	0.235	0.957	<b>0.044</b>	<b>0.055</b>	<b>0.057</b>	0.009	0.029	0.125
$\mathcal{H}_{1,NLQ}^1$	0.346	0.963	1.000	0.047	0.106	0.398	0.024	0.160	0.411
$\mathcal{H}_{1,NLQ}^2$	0.471	0.988	1.000	0.054	0.129	0.597	0.039	0.203	0.572
$\mathcal{H}_{1,NLQ}^3$	0.967	1.000	1.000	0.170	0.571	1.000	0.250	0.575	1.000
$\mathcal{H}_{1,NLT}^1$	0.359	0.963	1.000	0.047	0.107	0.399	0.025	0.160	0.406
$\mathcal{H}_{1,NLT}^2$	0.576	0.997	1.000	0.062	0.145	0.710	0.070	0.230	0.683
$\mathcal{H}_{1,NLT}^3$	0.978	1.000	1.000	0.118	0.443	1.000	0.200	0.445	1.000

Table 10: Scenario S3. The description of Table 8 applies.

## 5 Real data application

We apply our GoF test to a real dataset with functional predictor and response (see Figure 1), openly accessible as the object `aemet_temp` from the `goffda` package. Another application is given in the SI for the dataset considered in Benatia et al. (2017). Along both applications, we used  $B = 10,000$  bootstrap replicates to calibrate all the bootstrap-based tests and the PCvM test was run using Algorithm 1 with FPCR-L1S,  $EV_p = EV_q = 0.99$ , and  $\hat{\lambda}_{1SE}$ . For both applications, the same qualitative results were obtained with FPCR or FPCR-L2.

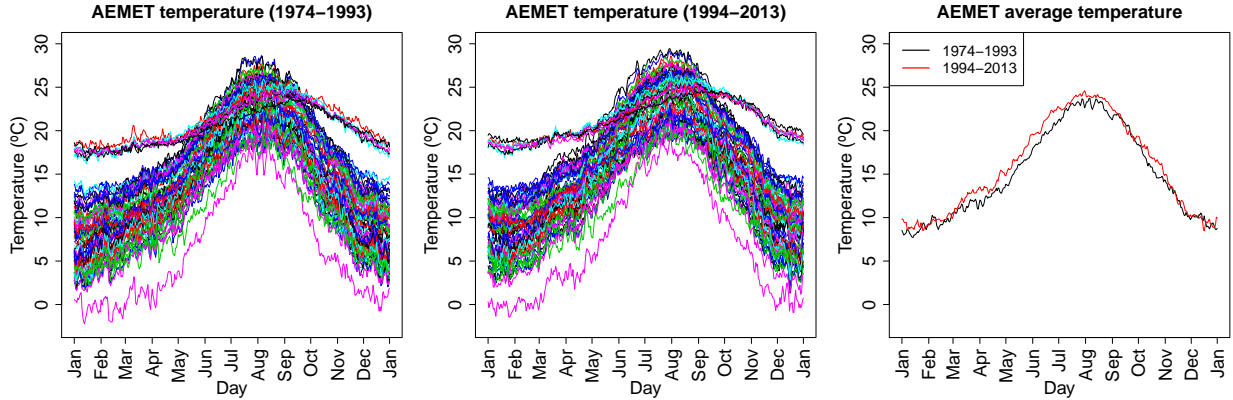


Figure 1: From left to right: samples of  $\mathcal{Y}$  and  $\mathcal{X}$ , and sample means of  $\mathcal{X}$  and  $\mathcal{Y}$ , indicating an increment of the average temperatures in the period 1994–2013 with respect to 1974–1983.

The “AEMET temperatures dataset” was constructed from the raw daily temperatures, along the span 1974–2013, of  $n = 73$  weather stations from the Meteorological State Agency of Spain (AEMET). We considered a partition of this dataset in two 20-year periods, 1974–1993 and 1994–2013, and computed the daily average temperature in each period. The aim of this partition is to explain the temperatures in the latter period ( $\mathcal{Y}$ ) from the ones in the former ( $\mathcal{X}$ ). Therefore, the response and predictor are valued in  $\mathbb{H}_1 = \mathbb{H}_2 = L^2([0, 365])$ . The functional observations are recorded in 365 equispaced grid points in the interval  $[0.5, 364.5]$  and are significantly rougher than in the previous application since no presmoothing is applied. The selected stations are the same as in the `aemet` dataset of the `fda.usc` package (Febrero-Bande and Oviedo de la Fuente, 2012) and were selected over a larger set of stations due to their consistent records and permanent locations over the 40-years period.

The PCvM test based on the data-driven  $\tilde{p} = 4$  and  $q = 3$  yielded  $p$ -value = 0.2538 when testing the GoF of the FLMFR. Hence, the sample shows no significant evidences against the FLMFR for any

sensible significance level. In addition,  $\hat{\beta}$  in Figure 2 (right) reveals several interesting insights: (i) the FLMFR mainly focuses on capturing positive correlation (positive values of  $\hat{\beta}$ ; marked in red) within a  $\pm 90$ -days band (in dashed lines) about a given time of the year, effectively corresponding to half a year; (ii) the predominance of positive values, together with the fact that almost all records are positive, points towards a general temperature increment on the 1994–2013 span with respect to 1974–1993; (iii) some of the visible temperature increments in the lower right panel of Figure 1, such as in Apr–May and Oct–Nov, are identified with the horizontal bands spanning the same periods on the right plot of Figure 2, for which there are almost no negative values of  $\hat{\beta}$ . We remark that the possible spatial dependence of the data was not taken into account in the analysis.

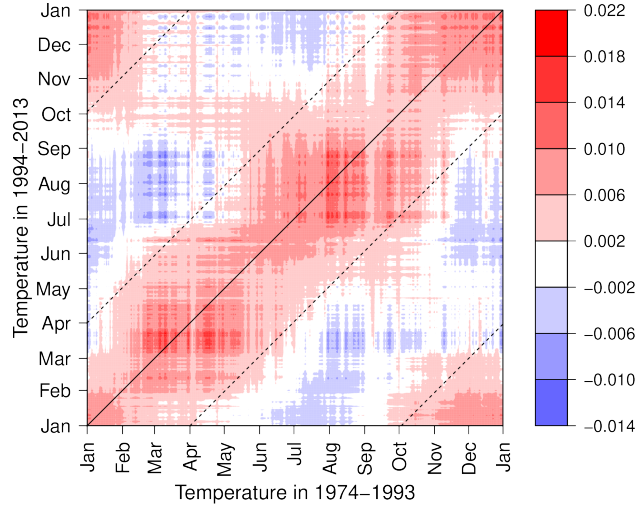


Figure 2: FPCR-L1S estimator  $\hat{\beta}$  for the AEMET temperatures dataset. Note how  $\hat{\beta}$  reflects the smoothness of the data, inherited by the FPC.

One may wonder whether  $\hat{\beta}$  (Figure 2) is associated to a simpler FLMFR. The answer appears to be negative, as the following attempted simplifications evidence: (i)  $\mathcal{H}_0 : \beta = 0$  was rejected by the KMSZ, PSS, LZS, and PCvM tests with null  $p$ -values; (ii)  $\mathcal{H}_0 : \beta = \hat{b}$ , where  $\hat{b}$  stands for the average value of the  $\hat{\beta}$  surface, was rejected by the PCvM test with null  $p$ -value; (iii)  $\mathcal{H}_0 : \beta(s, t) = \mathbb{1}_{\{s=t\}}$ , the stationary-temperature hypothesis, was rejected by the PCvM test with  $p$ -value =  $10^{-4}$ ; (iv)  $\mathcal{H}_0 : \beta(s, t) = \tilde{\beta}(s, t)$ , where  $\tilde{\beta}$  is constructed by averaging along the periodic diagonals of  $\hat{\beta}$ , was rejected by the PCvM test with null  $p$ -value. Interestingly, the third analysis is congruent with the outcome of the projected ANOVA (Cuesta-Albertos and Febrero-Bande, 2010) which, when ran using `fda.usc`'s implementation with 30 projections, rejected the equality of the mean group curves with null  $p$ -value. As a conclusion, this data application not only reveals that there are no evidences against the FLMFR in the studied data, but also provides yet another evidence, within a short-term and a localized region, of a significant climate change. Analogous results were obtained presmoothing with a local linear estimator featuring a cross-validated bandwidth.

## 6 Conclusions

We have developed a GoF test for assessing the composite null hypothesis of the FLMFR. Our statistic: (i) is based on a characterization of the null hypothesis in terms of finite-dimensional directions; (ii) can be regarded as a weighted quadratic norm of the coefficients of the residuals in a truncated basis of  $\mathbb{H}_2$ ; (iii) neatly extends a previous proposal for the FLMSR. Furthermore, together with a novel estimator for the FLMFR and the use of several convenient computational procedures, we can achieve an expedient bootstrap calibration of the test statistic. Empirical results show that, in the studied scenarios, the test calibrates adequately the composite null hypothesis and detects a

variety of linear and nonlinear alternatives. In addition, it is competitive against previous proposals for testing the significance of the functional predictor.

As noted, the PCvM statistic only depends on the functional residuals. Hence, the formulated test could be extended to alternative (possibly non-linear) regression models, provided that reliable estimators exist for them. Evident extensions are the testing of the FLMFR in the presence of several functional covariates and the testing of the functional linear model with functional response and scalar predictor.

## Supporting information

An extra real data application is provided in the supporting information.

## Acknowledgements

The authors gratefully thank Prof. Manuel Febrero-Bande for discussions at early stages of the project and for providing access to the dataset of raw AEMET temperatures. The first author acknowledges support from grants PGC2018-097284-B-I00 and IJCI-2017-32005 from the Spanish Ministry of Economy and Competitiveness (co-funded with FEDER funds). The second author acknowledges support from grant PGC2018-099549-B-I00 from the same agency. The first and fourth authors acknowledge financial support from grant MTM2016-76969-P also from the same agency. The third author acknowledges support through the Severo Ochoa Program from the Government of the Principality of Asturias (grant PA-20-PF-BP19-053). The authors gratefully acknowledge the computing resources of the Supercomputing Center of Galicia (CESGA). Comments by an Associate Editor and a referee are gratefully acknowledged.

## A Proofs

*Proof of Lemma 1.* We proceed by proving equivalences by pairs. First of all, the equivalence of  $i$  and  $ii$  can be derived straightforwardly by the definition of  $m(x) = \mathbb{E}[\mathcal{Y}|\mathcal{X} = x]$  under  $\mathcal{H}_0$ . The equivalence of  $ii$  and  $iii$  follows by applying Lemma 2.1(A) in Patilea and Sánchez-Sellero (2020) pointwisely to each point  $t$  of  $\mathcal{Z}(t) := (\mathcal{Y} - \langle \mathcal{X}, \beta \rangle)(t)$ . The implication  $iii \implies iv$  is trivial from the linearity of the inner product and the conditional expectation. On the other hand, the converse implication follows by taking  $\gamma_{\mathcal{Y}}$  in the orthonormal basis of  $\mathbb{H}_2$ ,  $\{\Phi_k\}_{k=1}^{\infty}$ . Then,  $\mathcal{Z}(t) = \sum_{k=1}^{\infty} z_k \Phi_k(t)$  a.s. with  $\mathbb{E}[z_k | \langle \mathcal{X}, \gamma_{\mathcal{X}} \rangle_{\mathbb{H}_1} = u] = 0$  for all  $k \geq 1$  and a.e.  $u \in \mathbb{R}$ , from where  $iv$  follows. Finally, the equivalence between  $iv$  and  $v$  arises due to the equivalence between the (real-valued and real-conditioned) conditional expectation and the integrated regression function (see, e.g., page 615 in Stute (1997)).  $\square$

*Proof of Lemma 2.* By applying pointwisely Lemma 2.1(A) in Patilea and Sánchez-Sellero (2020),  $iii$  is equivalent to  $iii'$ , which replaces “ $\forall \gamma_{\mathcal{X}} \in \mathbb{S}_{\mathbb{H}_1}$ ” by “ $\forall \gamma_{\mathcal{X}} \in \mathbb{S}_{\mathbb{H}_1, \{\Psi_j\}_{j=1}^{\infty}}^{p-1}$  and for all  $p \geq 1$ ”. As in the proof of Lemma 1,  $iii' \implies iv'$  trivially, and the converse follows by similar arguments. The equivalence between  $iv'$  and  $v'$  is also analogous.  $\square$

*Proof of Lemma 3.* Let  $\bar{\mathbf{x}} := \mathbf{x}/\|\mathbf{x}\| \in \mathbb{S}^{q-1}$  and  $\bar{\mathbf{y}} := \mathbf{y}/\|\mathbf{y}\| \in \mathbb{S}^{q-1}$  for  $\mathbf{x} \neq \mathbf{0}$  and  $\mathbf{y} \neq \mathbf{0}$  (otherwise the result is trivial). Consider then the tangent-normal decomposition  $\boldsymbol{\omega} = t\bar{\mathbf{x}} + (1-t^2)^{1/2}\mathbf{B}_{\bar{\mathbf{x}}}\boldsymbol{\xi}$  (as given, e.g., in Lemma 2 in García-Portugués et al. (2013)), where  $t \in [-1, 1]$ ,  $\boldsymbol{\xi} \in \mathbb{S}^{q-2}$ , and  $\mathbf{B}_{\bar{\mathbf{x}}}$  is a  $q \times (q-1)$  semi-orthogonal matrix such that  $\mathbf{B}_{\bar{\mathbf{x}}}^{\prime}\mathbf{B}_{\bar{\mathbf{x}}} = \mathbf{I}_{q-1}$  and  $\mathbf{B}_{\bar{\mathbf{x}}}\mathbf{B}_{\bar{\mathbf{x}}}^{\prime} = \mathbf{I}_q - \bar{\mathbf{x}}\bar{\mathbf{x}}^{\prime}$ . Then, the integral can be rewritten as

$$\|\mathbf{x}\|\|\mathbf{y}\| \int_{\mathbb{S}^{q-1}} (\bar{\mathbf{x}}^{\prime}\boldsymbol{\omega})(\bar{\mathbf{y}}^{\prime}\boldsymbol{\omega}) d\boldsymbol{\omega} = \|\mathbf{x}\|\|\mathbf{y}\| \int_{\mathbb{S}^{q-2}} \int_{-1}^1 t \left( \bar{\mathbf{y}}^{\prime}\bar{\mathbf{x}} + (1-t^2)^{1/2}\mathbf{B}_{\bar{\mathbf{x}}}\boldsymbol{\xi} \right) (1-t^2)^{(q-3)/2} dt d\boldsymbol{\xi}$$

$$= \|\mathbf{x}\| \|\mathbf{y}\| \bar{\mathbf{x}}' \bar{\mathbf{y}} \int_{\mathbb{S}^{q-2}} d\boldsymbol{\xi} \times \int_{-1}^1 t^2 (1-t^2)^{(q-3)/2} dt,$$

where symmetry simplifies the first integral. The result follows from  $\int_{\mathbb{S}^{q-2}} d\boldsymbol{\xi} = 2\pi^{(q-1)/2} / \Gamma((q-1)/2)$  and  $\int_{-1}^1 t^2 (1-t^2)^{(q-3)/2} dt = \sqrt{\pi} \Gamma((q-1)/2) / (2\Gamma(q/2+1))$ .  $\square$

*Proof of Lemma 4.* Assume  $n \geq 2$ , as if  $n = 1$  trivially  $\mathbf{A}_\bullet = 2\pi^{p/2} / \Gamma(p/2) > 0$ . From the cases described in (16), and since the coefficients  $\{\mathbf{x}_{i,p}\}_{i=1}^n$  are pairwise distinct, it follows that

$$\mathbf{A}_\bullet = \frac{\pi^{p/2-1}}{\Gamma(p/2)} \left\{ \sum_{r=1}^n \mathbf{A}_r + \pi \mathbf{I}_n \right\}, \quad (\mathbf{A}_r)_{ij} := \begin{cases} \pi, & \text{if } i = r \text{ or } j = r, \\ (15), & \text{otherwise,} \end{cases}$$

where  $\mathbf{e}_r$  stands for the  $r$ -th canonical vector in  $\mathbb{R}^n$ . The matrices  $\mathbf{A}_r$  have a clear block structure. For example, if  $r = n$ , then  $\mathbf{A}_r = (\mathbf{B}_r, \pi \mathbf{1}_{n-1}; \pi \mathbf{1}'_{n-1}, \pi)$ , where  $\mathbf{1}_{n-1}$  is a vector of  $n-1$  ones and  $(\mathbf{B}_r)_{k\ell} := A_{o_k o_\ell r}^{(\angle)}$ , with indexes  $o_i := i + \mathbb{1}_{\{i \leq r\}}$ ,  $i = 1, \dots, n-1$ . Analogous block expressions follow for  $r < n$ , yet more cumbersome since  $\mathbf{B}_r$  is split into four blocks. In any case, for any  $r = 1, \dots, n$ , given  $\mathbf{v} \in \mathbb{R}^n$ , then  $\mathbf{v}' \mathbf{A}_r \mathbf{v} = \mathbf{v}'_{-r} \mathbf{B}_r \mathbf{v}_{-r} + \pi \left[ v_r^2 + 2v_r \sum_{\substack{j=1 \\ j \neq r}}^n v_j \right]$  and, as a consequence,  $\mathbf{v}' \left\{ \sum_{r=1}^n \mathbf{A}_r + \pi \mathbf{I}_n \right\} \mathbf{v} = \sum_{r=1}^n \mathbf{v}'_{-r} \mathbf{B}_r \mathbf{v}_{-r} + 2\pi \left( \sum_{j=1}^n v_j \right)^2$ . Therefore, the sum is positive for any  $\mathbf{v} \neq \mathbf{0}$  if the matrices  $\mathbf{B}_r$ ,  $r = 1, \dots, n$ , are positive semi-definite.

Set  $\mathbf{y}_k := (\mathbf{x}_{k,p} - \mathbf{x}_{r,p}) / \|\mathbf{x}_{k,p} - \mathbf{x}_{r,p}\| \in \mathbb{S}^{p-1}$  for  $k = 1, \dots, n$ ,  $k \neq r$ , and  $p \geq 1$ . From (15),  $(\mathbf{B}_r)_{k\ell} = \psi(\cos^{-1}(\mathbf{y}'_k \mathbf{y}_\ell))$ , with  $\psi(\theta) = \pi - \theta$ ,  $\theta \in [0, \pi]$ . Define  $\tilde{\psi}(\theta) := \psi(\theta) / (2\pi) - 1/4$ . If  $p \geq 2$ , from the asymptotic distribution of the Ajne's statistic (Prentice, 1978, page 172),

$$\tilde{\psi}(\theta) = \sum_{k=1}^{\infty} \frac{4(k-1) + p}{p-2} b_{2k-1}^2 C_{2k-1}^{(p-2)/2}(\cos \theta), \quad b_{2k-1} = \frac{2^{p-2} \Gamma(p/2) \Gamma(k-1+p/2) (2k-2)!}{(-1)^{k-1} \pi (k-1)! (2k+p-3)!},$$

where  $C_k^\alpha$  denotes the Gegenbauer polynomial of index  $\alpha$  and order  $k$  (when  $p = 2$ , we use implicitly that  $\lim_{\alpha \rightarrow 0} C_k^\alpha(\cos \theta) / \alpha = (2/k) \cos(k\theta)$ ). Therefore, the Gegenbauer coefficients of  $\tilde{\psi}$  are non-negative (positive if odd; null if even) and, due to the properties of the Gegenbauer polynomials, so do are the coefficients of  $\psi$ . Then, the characterization by Schoenberg (1942) entails that  $\psi$  is definite positive. This implies that, for any collection of points  $\mathbf{z}_1, \dots, \mathbf{z}_m \in \mathbb{S}^{p-1}$ , for any  $m \geq 2$  and  $p \geq 2$ , the matrix  $(\psi(\cos^{-1}(\mathbf{z}'_k \mathbf{z}_\ell)))_{k,\ell=1,\dots,m}$  is positive semi-definite. When  $p = 1$ , recall that  $y_k \in \{-1, +1\}$  and  $(\mathbf{B}_r)_{k\ell} = \pi \delta_{y_k y_\ell}$ , so  $\mathbf{B}_r$  can be rearranged as  $(\pi \mathbf{1}_{n_- \times n_-}, \mathbf{0}_{n_- \times n_+}; \mathbf{0}_{n_+ \times n_-}, \pi \mathbf{1}_{n_+ \times n_+})$ , where  $n_\pm$  denotes the number of  $y_k$ 's equal to  $\pm 1$ . Trivially,  $(\mathbf{B}_r)_{k\ell}$  is rank 2 with non-null eigenvalues  $n_+ \pi$  and  $n_- \pi$ . As a consequence,  $\mathbf{B}_r$  is positive semi-definite for all  $r = 1, \dots, n$ ,  $n \geq 2$ ,  $p \geq 1$ .  $\square$

## References

- Aguilera, A. M. and Aguilera-Morillo, M. C. (2013). Penalized PCA approaches for B-spline expansions of smooth functional data. *Appl. Math. Comput.*, 219(14):7805–7819.
- Aneiros-Pérez, G. and Vieu, P. (2013). Testing linearity in semi-parametric functional data analysis. *Comput. Statist.*, 28(2):413–434.
- Benatia, D., Carrasco, M., and Florens, J. P. (2017). Functional linear regression with functional response. *J. Econometrics*, 201(2):269–291.
- Bickel, P. J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.*, 1(6):1071–1095.

- Bücher, A., Dette, H., and Wiecek, G. (2011). Testing model assumptions in functional regression models. *J. Multivariate Anal.*, 102(10):1472–1488.
- Cardot, H., Ferraty, F., Mas, A., and Sarda, P. (2003). Testing hypotheses in the functional linear model. *Scand. J. Statist.*, 30(1):241–255.
- Cardot, H., Ferraty, F., and Sarda, P. (1999). Functional linear model. *Statist. Prob. Lett.*, 45(1):11–22.
- Chiou, J. M. and Müller, H. G. (2007). Diagnostics for functional regression via residual processes. *Comput. Statist. Data Anal.*, 51(10):4849–4863.
- Chiou, J. M., Müller, H. G., Wang, J. L., and Carey, J. R. (2003). A functional multiplicative effects model for longitudinal data, with application to reproductive histories of female medflies. *Statist. Sinica*, 13(4):1119–1133.
- Crambes, C., Kneip, A., and Sarda, P. (2009). Smoothing splines estimators for functional linear regression. *Ann. Statist.*, 37(1):35–72.
- Crambes, C. and Mas, A. (2013). Asymptotics of prediction in functional linear regression with functional outputs. *Bernoulli*, 19(5B):2627–2651.
- Cuesta-Albertos, J. A. and Febrero-Bande, M. (2010). Multiway ANOVA for functional data. *TEST*, 19(3):537–557.
- Cuesta-Albertos, J. A., García-Portugués, E., Febrero-Bande, M., and González-Manteiga, W. (2019). Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes. *Ann. Statist.*, 47(1):439–467.
- Cuevas, A., Febrero, M., and Fraiman, R. (2002). Linear functional regression: the case of fixed design and functional response. *Canad. J. Statist.*, 30(2):285–300.
- Delsol, L., Ferraty, F., and Vieu, P. (2011). Structural test in regression on functional variables. *J. Multivariate Anal.*, 102(3):422–447.
- Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.*, 1(2):279–290.
- Escanciano, J. C. (2006). A consistent diagnostic test for regression models using projections. *Econometric Theory*, 22(6):1030–1051.
- Febrero-Bande, M. and Oviedo de la Fuente, M. (2012). Statistical computing in functional data analysis: The R package fda.usc. *J. Stat. Softw.*, 51(4):1–28.
- Ferraty, F., Keilegom, I. V., and Vieu, P. (2012). Regression when both response and predictor are functions. *J. Multivariate Anal.*, 109:10–28.
- Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2011). Kernel regression with functional response. *Electron. J. Statist.*, 5:159–171.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. Springer Series in Statistics. Springer, New York.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.*, 33(1):1–22.
- Gabrys, R., Horváth, L., and Kokoszka, P. (2010). Tests for error correlation in the functional linear model. *J. Amer. Statist. Assoc.*, 105(491):1113–1125.

- García-Portugués, E. and Álvarez-Liébaná, J. (2020). *goffda: Goodness-of-fit tests for functional data*. R package version 0.0.7.
- García-Portugués, E., Crujeiras, R. M., and González-Manteiga, W. (2013). Kernel density estimation for directional-linear data. *J. Multivariate Anal.*, 121:152–175.
- García-Portugués, E., González-Manteiga, W., and Febrero-Bande, M. (2014). A goodness-of-fit test for the functional linear model with scalar response. *J. Comp. Graph. Stat.*, 23(3):761–778.
- González-Manteiga, W. and Crujeiras, R. M. (2013). An updated review of Goodness-of-Fit tests for regression models. *TEST*, 22(3):361–411.
- Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.*, 21(4):1926–1947.
- He, G., Müller, H. G., Wang, J. L., and Yang, W. (2010). Functional linear regression via canonical analysis. *Bernoulli*, 16(3):705–729.
- Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer Series in Statistics. Springer, New York.
- Horváth, L. and Reeder, R. (2013). A test of significance in functional quadratic regression. *Bernoulli*, 19(5A):2130–2151.
- Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester.
- Imaizumi, M. and Kato, K. (2018). PCA-based estimation for functional linear regression with functional responses. *J. Multivariate Anal.*, 163:15–36.
- Kokoszka, P., Maslova, I., Sojka, J., and Zhu, L. (2008). Testing for lack of dependence in the functional linear model. *Canad. J. Statist.*, 36(2):207–222.
- Kong, D., Staicu, A. M., and Maity, A. (2016). Classical testing in functional linear models. *J. Nonparametr. Stat.*, 28(4):813–830.
- Lee, C. E., Zhang, X., and Shao, X. (2020). Testing conditional mean independence for functional data. *Biometrika*, 107(2):331–346.
- Lian, H. (2011). Convergence of functional k-nearest neighbor regression estimate with functional responses. *Electron. J. Statist.*, 5:31–40.
- McLean, M. W., Hooker, G., and Ruppert, D. (2015). Restricted likelihood ratio tests for linearity in scalar-on-function regression. *Stat. Comput.*, 25(5):997–1008.
- Panaretos, V. M. and Tavakoli, S. (2013). Fourier analysis of stationary time series in function space. *Ann. Statist.*, 41(2):568–603.
- Park, J. and Qian, J. (2012). Functional regression of continuous state distribution. *J. Econometrics*, 167(2):397–412.
- Park, T., Shao, X., and Yao, S. (2015). Partial martingale difference correlation. *Electron. J. Statist.*, 9(1):1492–1517.
- Patilea, V. and Sánchez-Sellero, C. (2020). Testing for lack-of-fit in functional regression models against general alternatives. *J. Statist. Plann. Inf.*, 209(516):229–251.

- Patilea, V., Sánchez-Sellero, C., and Saumard, M. (2016a). *fdapss: A companion R package for the paper “Testing the predictor effect on a functional response”*. R package version 1.0.0.
- Patilea, V., Sánchez-Sellero, C., and Saumard, M. (2016b). Testing the predictor effect on a functional response. *J. Amer. Statist. Assoc.*, 111(516):1684–1695.
- Preda, C. and Saporta, G. (2005). PLS regression on a stochastic process. *Comput. Stat. Data Anal.*, 48(1):149–158.
- Prentice, M. J. (1978). On invariant tests of uniformity for directions and orientations. *Ann. Statist.*, 6(1):169–176.
- Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer Series in Statistics. Springer, New York.
- Schoenberg, I. J. (1942). Positive definite functions on spheres. *Duke Math. J.*, 9(1):96–108.
- Shao, J. (1993). Linear model selection by cross-validation. *J. Amer. Statist. Assoc.*, 88(422):486–494.
- Staicu, A. M., Lahiri, S. N., and Carrol, R. J. (2015). Significance tests for functional data with complex dependence structure. *J. Statist. Plann. Inf.*, 156:1–13.
- Stute, W. (1997). Nonparametric model checks for regression. *Ann. Statist.*, 25(2):613–641.
- Stute, W., González-Manteiga, W., and Presedo-Quindimil, M. (1998). Bootstrap approximations in model checks for regression. *J. Amer. Statist. Assoc.*, 93(44):141–149.
- Tibshirani, R. J. (1996). Regression shrinkage and selection via the LASSO. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 58(1):267–288.
- Wang, H., Zhong, P. S., Cui, Y., and Li, Y. (2018). Unified empirical likelihood ratio tests for functional concurrent linear models and the phase transition from sparse to dense functional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 60(2):343–364.
- Yao, F., Müller, H. G., and Wang, J. L. (2005). Functional linear regression analysis for longitudinal data. *Ann. Statist.*, 33(6):2873–2903.
- Yasemin-Tekbudak, M., Alfaro-Córdoba, M., Maity, A., and Staicu, A. M. (2019). A comparison of testing methods in scalar-on-function regression. *AStA Adv. Stat. Anal.*, 103:411–436.
- Zhang, X. (2016). White noise testing and model diagnostic checking for functional time series. *J. Econometrics*, 194(1):76–95.
- Zhao, P. and Yu, B. (2006). On model selection consistency of Lasso. *J. Mach. Learn. Res.*, 7:2541–2563.



# Supporting information for “A goodness-of-fit test for the functional linear model with functional response”

Eduardo García-Portugués<sup>1,2</sup>, Javier Álvarez-Liébana<sup>3</sup>,  
Gonzalo Álvarez-Pérez<sup>4,6</sup>, and Wenceslao González-Manteiga<sup>5</sup>

## Abstract

This supporting information contains an extra real data application.

**Keywords:** Bootstrap; Cramér–von Mises statistic; Functional data; Regularization.

The “Ontario dataset”, constructed by the authors of Benatia et al. (2017), contains the hourly electricity consumption ( $\mathcal{Y}$ ; measured in gigawatts) and smoothed temperature ( $\mathcal{X}$ ; Celsius degrees) in the province of Ontario (Canada). More precisely, it features a set of  $n = 368$  daily curves on 2010–2014, where only summer months are taken into account, while weekends and holidays are discarded (hence, the  $i$ -th datum is not necessarily consecutive in time to the  $(i + 1)$ -th). The response is valued in  $\mathbb{H}_2 = L^2([0, 24])$  and discretized in 25 equispaced grid points. Each temperature curve is valued in  $\mathbb{H}_1 = L^2([-24, 48])$  and discretized in 73 equispaced grid points. The interval  $[-24, 48]$  accounts for a 3-days window that is considered since the past and future temperatures of a given day may influence the demand of energy on that day. Thus, the response is also regressed on 24 past and future hours. The raw temperature records are smoothed by a local polynomial regression on a weighted average of the temperatures of 41 Ontarian cities, producing the smoothed temperature, finally shifted so its minimum is set to  $0^\circ$ .

We check whether there exists a linear relation in the Ontario dataset. This is inspired by the data application in Benatia et al. (2017), where a FLMFR featuring several seasonal dummies is considered. Therefore, testing the GoF of the “canonical” FLMFR allows to evaluate if a seasonal-free simplified model succeeds in describing the daily electricity consumption from the temperature alone. Based on the data-driven selection of  $\tilde{p} = 7$  and  $q = 4$ , the PCvM test gave null  $p$ -value, rejecting emphatically the FLMFR. When testing for significance, the KMSZ, PSS, LZS, and PCvM tests clearly rejected with null  $p$ -values. Hence, a nontrivial and nonlinear functional relation between daily electricity consumption and the temperature is evidenced, and the seasonal-free version of Benatia et al. (2017)’s model is shown to be inadequate for modeling such relation.

A referee and Associate Editor pointed out that the presence of temporal dependence in the data violates the iid assumption of our GoF test. Indeed, the data construction inherited from Benatia et al. (2017) employs 3-days overlapping windows that notably increases the serial dependency of the functional records. In order to investigate if this dependency was the leading rejection cause of the FLMFR, we have run our test retaining the 3-days windows but ensuring there are no overlaps in the observations. That is, we have considered only the curves for *day 1* (includes days 0, 1, 2), *day 4* (days 3, 4, 5), *day 7* (days 6, 7, 8), etc., properly handling weekends and holidays. The results

---

<sup>1</sup>Department of Statistics, Carlos III University of Madrid (Spain).

<sup>2</sup>UC3M-Santander Big Data Institute, Carlos III University of Madrid (Spain).

<sup>3</sup>Department of Statistics and Operations Research and Mathematics Didactics, University of Oviedo (Spain).

<sup>4</sup>Department of Physics, University of Oviedo (Spain).

<sup>5</sup>Department of Statistics, Mathematical Analysis and Optimization, University of Santiago de Compostela (Spain).

<sup>6</sup>Corresponding author. e-mail: gonzaloalvarez@uniovi.es.

are the same as in the original application: the FLMFR is emphatically rejected (null  $p$ -values) for the three possible subsettings of non-overlapping data and for different estimators. The no effect hypothesis is also rejected with null  $p$ -values. From this analysis, we are confident that the rejections with the original data are not primarily driven by temporal dependence (though still present in the non-overlapping data, e.g., by annual periodicity), and that a reduction in the complexity of the model in Benatia et al. (2017) through a seasonal-free version is not possible.

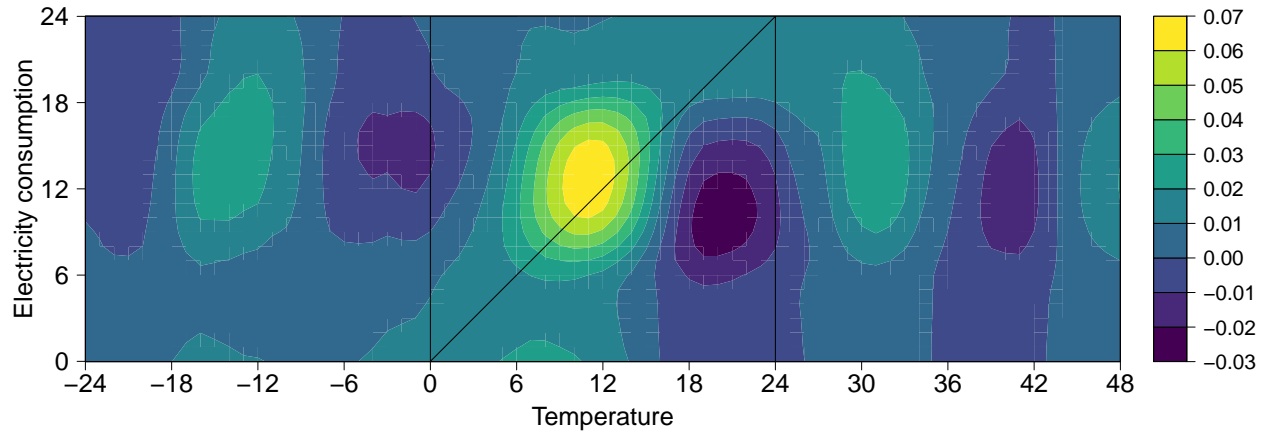


Figure 3: FPCR-L1S estimator  $\hat{\beta}$  for the Ontario dataset. Note how  $\hat{\beta}$  reflects the smoothness of the data, inherited by the FPC. The plot is coherent with Figure 11 in Benatia et al. (2017), yet ours is less centered at the diagonal, probably since no seasonal dummies were considered for fitting the FLMFR.

## References

Benatia, D., Carrasco, M., and Florens, J. P. (2017). Functional linear regression with functional response. *J. Econometrics*, 201(2):269–291.