REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000–000
https://doi.org/00.00000/revstat.v00i0.000

Review of Methods for Functional One-Way Analysis of Variance

Authors: Łukasz Smaga ✉
– Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poland
ls@amu.edu.pl

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- There are many statistical methods for the functional analysis of variance. They are based on different concepts for example aggregating pointwise statistics, basis function expansion, graphical envelope, principal components analysis, and random projections. This paper concerns different aspects of such strategies. The paper covers, in particular, one-way and multi-way analyses, univariate and multivariate functional data, independent and repeated measurements, and fully and partially observed functional data. Theoretical constructions and properties are accompanied by summaries of simulation results and illustrated on real data examples. Additionally, the R code for performing the analyses in the examples is available on GitHub.

Keywords:

- *aggregating pointwise statistics; basis function expansion; graphical methods; partially observed functional data; principal components analysis; random projections.*

AMS Subject Classification:

- 62M99, 62G10.

✉ Corresponding author

1. INTRODUCTION

Functional data analysis (FDA) is a branch of statistics that analyzes observations treated as functions, curves, or surfaces. To represent the data in such a way, one needs only to measure some variable over time or space, which is a scenario encountered in many fields, such as brain imaging data, medical measurements over time, biological development, meteorology, etc. Then the discrete data observed at so-called design time points can be transformed into functional data. A time series is treated as the realization of some random process and it is represented as a function, which is one functional observation. Such a representation allows us to avoid many problems of classical multivariate statistical methods, for example, the curse of dimensionality and missing data. Therefore, numerous methods have been developed for classification, clustering, dimension reduction, regression, and statistical hypothesis testing for functional data. We refer to the following monographs, where the methodology, real data examples, and computational aspects are studied: Ferraty and Vieu (2006), Horváth and Kokoszka (2012), Ramsay and Silverman (2005), and Zhang (2013).

In this paper, we review various aspects and strategies of the functional analysis of variance (FANOVA), which is popular analysis of FDA. The methods are based on different approaches, for example, dimension reduction (basis expansion, functional principal components), random projections on multivariate data, and aggregating pointwise statistics. In Section 2, we consider the two sample problem for functional data, which presents the use of the functional principal component analysis. Section 3 is the longest in the present paper since it concerns a popular issue of one-way analysis of variance for univariate functional data. In separate subsections, we present the test procedures based on: basis function expansion, random projections, aggregating pointwise test statistics, graphical approach, partially observed functional data, and repeated measurements. In Section 4, we consider the less popular but important problem of multivariate functional analysis of variance. Section 5 concludes the paper. In the supplement, we present the simulation studies based on real data examples, which show a small comparison of the tests considered and a way to do this for a particular data set. Note that these methods are suitable not just for FANOVA, but for many other problems in FDA. Thus, for each method, we also refer to another problems of FDA, where they are used. Of course, this work does not exhaust the topic, it is a subjective presentation of selected methods. For additional information and further considerations, we refer to the given literature. To maintain the clarity of the presentation, some technical aspects (e.g., assumptions) may have been omitted.

For each case, we present the theory and the real data example illustrating the use of the methods considered. Moreover, on GitHub `https://github.com/ls-git-17/fda_methods_fanova`, there is another supplement to this paper, which contains the R code for performing these examples and simulation studies in the supplement (R Core Team, 2024). Since the implementation of some of the tests presented here is not available in free packages, we give it in this repository.

1.1. General assumptions and definitions

Let us now give a few general assumptions, notations, and comments, which will appear in many places in this paper. First, as we mentioned above the functional data are observed discretely. Assume that we measure some variable (e.g., temperature) in a given location (e.g., in Poznań) and some time points (called design time points; e.g., every day of a year). Then, the discrete functional data are the pairs (t_i, x_i), $i = 1, \ldots, n$, where t_i is the ith design time point (e.g., the ith day in a year) and x_i is the value of some variable (e.g., temperature) observed in time point t_i (i.e., $x_i = x(t_i)$ for some function x). Second, one of the aspects, that has a significant impact on the power of the test procedures, is the amount of correlation of functional data. Writing the correlation of functional data, we think about the correlation between observations at any two different design time points. Finally, we usually assume that the functional data are random processes belonging to the Hilbert space $L_2^d(\mathcal{I})$ of d-dimensional vectors of square-integrable functions defined on the interval $\mathcal{I} = [a, b]$, $a, b \in \mathbb{R}$. This space is endowed with the following inner product:

$$\langle \mathbf{f}, \mathbf{g} \rangle_d = \int_{\mathcal{I}} \mathbf{f}^\top(t)\mathbf{g}(t)dt = \sum_{i=1}^{d} \int_{\mathcal{I}} f_i(t)g_i(t)dt = \sum_{i=1}^{d} \langle f_i, g_i \rangle_1$$

for $\mathbf{f} = (f_1, \ldots, f_d)^\top, \mathbf{g} = (g_1, \ldots, g_d)^\top \in L_2^d(\mathcal{I})$. Thus, the norm in this space is as follows:

$$\|\mathbf{f}\|_d = \sqrt{\langle \mathbf{f}, \mathbf{f} \rangle_d} = \sqrt{\sum_{i=1}^{d} \int_{\mathcal{I}} f_i^2(t)dt}.$$

When $d = 1$, we consider the univariate functional data, while for $d > 1$, we have the multivariate functional data.

1.2. Real data example

In this section, we describe the example of the real data set, which we will use for illustrating most of the FDA methods considered in this paper. To reach a wider audience, we use the Canadian weather data set, which is not very difficult and it is well-known and freely available in the R package `fda` Ramsay (2023). In Figure 1, we present the map of Canada with the points representing 35 weather stations. They are divided into three groups Eastern, Western, and Northern Canada. The number of stations in these groups is 15, 15, and 5, respectively. In each of these stations, different weather parameters are recorded. Here, we consider two of them, namely temperature (in Celsius degrees) and precipitation (in millimeters). In the data set, there are average daily temperature and rainfall (rounded to 0.1mm) for each day of the year. Their trajectories are presented in Figure 2. Since the variables are measured for each day of the year, we can model them are functional variables. Thus, we have three samples of univariate functional variable (temperature or precipitation) or multivariate functional features (temperature and precipitation), which are observed in 365 design time points. Figure 2 also presents the sample mean functions (defined later) for each group separately. They are the estimates of the central tendency of functional variables, and we will use them to compare the three groups under a central tendency. We can observe

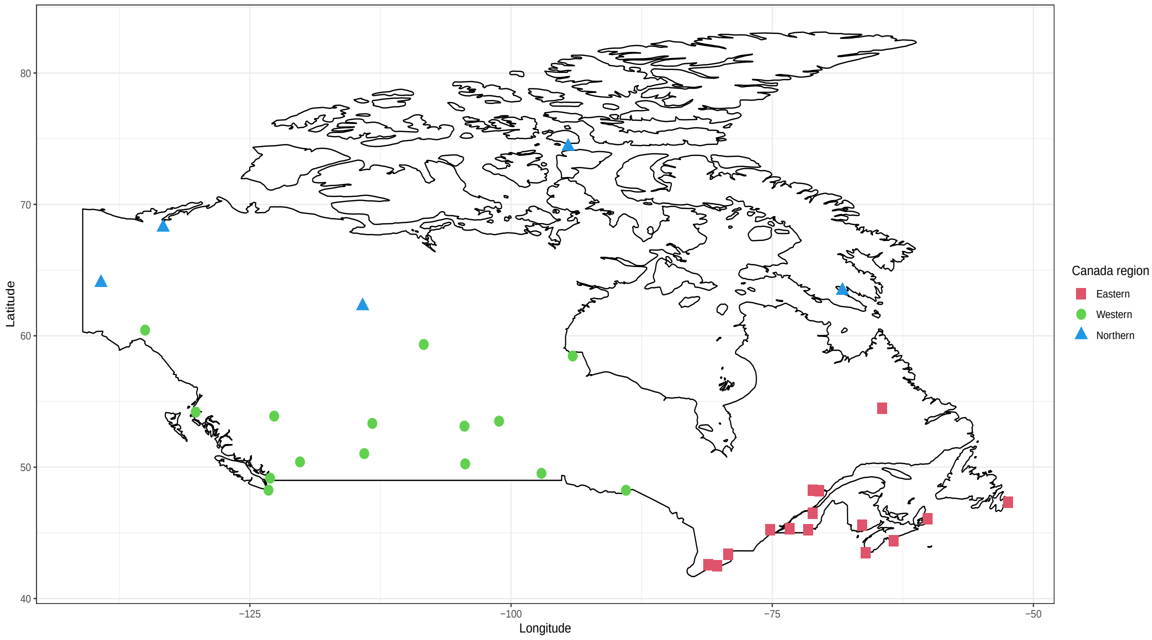

Figure 1: Map of Canada with noted weather stations in three different regions.

that the Eastern and Western weather stations seem to have a similar average temperature, while the temperature in Northern Canada is much smaller. The precipitation seems to be different in different areas of Canada. In the following sections, we will try to verify these observations using appropriate statistical methods for functional data.

2. TWO SAMPLE PROBLEM

In this section, we present the two sample problem for functional data and the test for it proposed by Horváth et al. (2013). This test is a nice example of the functional method based on the functional principal components (FPCs), which are very important in the FDA. For example, the functional principal components are used for testing the stationarity of functional time series (Horváth et al., 2014), functional regression (Nie et al., 2018), clustering functional data (Wu et al., 2022), and classifying them (Chatterjee et al., 2023). The FPCs are also applied to many practical problems. For review and some recent applications, we refer to Ullah and Finch (2013) and Karuppusami et al. (2022), respectively. For the two sample problems, there are however many other methods for example those presented in Ghiglietti et al. (2017); Hall and Van Keilegom (2007); Horváth et al. (2009); Jiang et al. (2019); Paparoditis and Sapatinas (2016); Zhang et al. (2010a,b,b).

Let $X_{11},\dots,X_{1n_1}$ and $X_{21},\dots,X_{2n_2}$ be two independent samples consisting of independent stochastic processes, where $n_1,n_2\in\mathbb{N}$. We assume that $X_{ij}\in L_2^1(\mathcal{I})$ and they satisfy the model:

$$X_{ij}(t)=\mu_i(t)+\varepsilon_{ij}(t),\ i=1,2,\ j=1,\dots,n_i,\ t\in\mathcal{I},$$

where μ_i are the mean functions and for each $i=1,2$, ε_{ij} are independent and identically distributed with $E(\varepsilon_{ij}(t))=0$ and $E\|\varepsilon_{ij}\|_1^4<\infty$. We are interested in testing the following

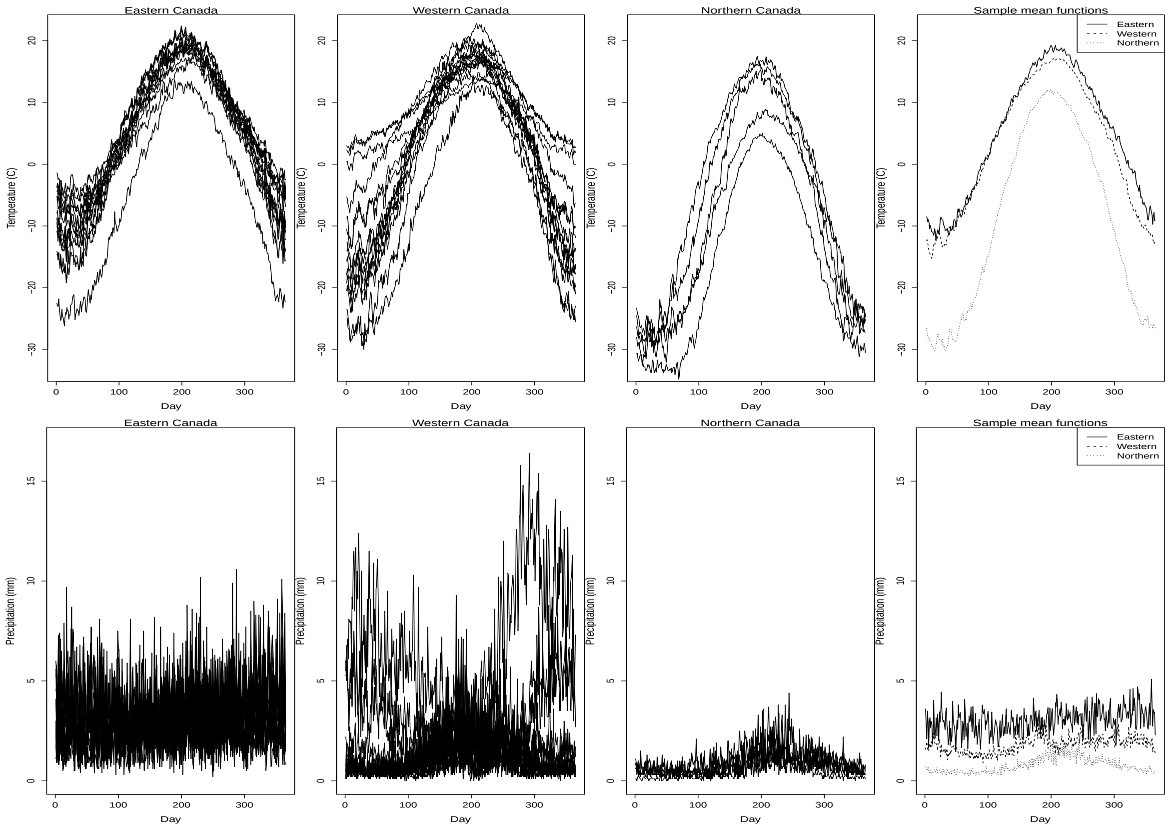

Figure 2: Trajectories of the temperature and precipitation noted in Canadian weather stations (first three columns) and the sample mean functions (fourth column).

null hypothesis:

$$H_0 : \mu_1 = \mu_2 \text{ in } L_2^1(\mathcal{I}).$$

The test statistic is based on the natural estimators of μ_i, i.e., the sample mean functions

$$\bar{X}_i(t) = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}(t), \; t \in \mathcal{I}$$

which are unbiased estimators for μ_i. The difference between these two sample mean functions is then projected onto the space determined by the leading eigenfunctions of the covariance operator $Z = (1-\theta)C_1 + \theta C_2$, where $n_1/(n_1+n_2) \to \theta \in [0, 1]$ and C_i is the covariance operator of X_{i1}, i.e., $C_i : L_2^1(\mathcal{I}) \to L_2^1(\mathcal{I})$ and $(C_i f)(\cdot) = \int_{\mathcal{I}} Cov(X_{i1}(\cdot), X_{i1}(t)) f(t)dt$ for $f \in L_2^1(\mathcal{I})$ (see, for example, Horváth and Kokoszka (2012) Chapter 2 for more detail). We assume that the eigenvalues of Z satisfy $\tau_1 > \tau_2 > \cdots > \tau_p > 0$ and $\varphi_1, \dots, \varphi_p$ are the corresponding eigenfunctions. Of course, they are not known in practice, so we use the eigenvalues $\hat{\tau}_k$ and eigenvectors $\hat{\varphi}_k$ of the empirical covariance function of the form:

$$\hat{Z}(s, t) = \sum_{i=1}^{2} \frac{n_{3-i}}{n_1 + n_2} \frac{1}{n_i} \sum_{j=1}^{n_i} (X_{ij}(s) - \bar{X}_i(s))(X_{ij}(t) - \bar{X}_i(t)).$$

Now, we project the difference $\bar{X}_1 - \bar{X}_2$ into the linear space spanned by $\hat{\varphi}_1, \dots, \hat{\varphi}_p$: $\hat{a}_k = \langle \bar{X}_1 - \bar{X}_2, \hat{\varphi}_k \rangle_1$ for $k = 1, \dots, p$. This means that instead of working with infinitely dimensional curve $\bar{X}_1 - \bar{X}_2$, we can work with p-dimensional vector $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_p)^\top$. Based on it, we construct the following test statistic:

$$T = \frac{n_1 n_2}{n_1 + n_2} \sum_{k=1}^{p} \frac{\hat{a}_k^2}{\hat{\tau}_k}.$$

Theorem 2.1 (Horváth et al. (2013)). *Under H_0 and the assumptions given above, we have $T \xrightarrow{d} \chi^2(p)$, where $\chi^2(p)$ denotes a chi-square random variable with p degrees of freedom.*

By Theorem 2.1, we have the following critical region of the asymptotic test based on T: $\{(x_{ij}) : T > \chi^2(1-\alpha, p)\}$, where $\chi^2(\beta, p)$ is the β-quantile of a chi-square distribution with p degrees of freedom and $\alpha \in (0,1)$ is the significance level. If $\mu_1 - \mu_2$ is not orthogonal to the linear span of $\varphi_1, \dots, \varphi_p$, then this test is consistent. Simulation studies conducted in Chapter 5 of Horváth and Kokoszka (2012) imply that this test controls the type I error level and has sensible power for moderate and large sample sizes. This is confirmed in the supplement.

As we noted, the test presented in this section is based on functional principal components. Let us now explain why. Namely, the construction of the test is based on the eigenfunctions $\hat{\varphi}_1, \dots, \hat{\varphi}_p$ of the sample covariance operator $\hat{Z}$, which are defined as the empirical functional principal components of the two samples we have. They are estimators of the functional principal components defined as the eigenfunctions of covariance operator Z (see, for example, Horváth and Kokoszka (2012) Chapter 3 for more detail). They play an important role in many methods of functional data analysis, as they reduce the infinite dimension of functional data to the finite dimension of the multivariate vectors of so-called scores (here $\hat{a}_k$). In this section, we can see how this technique can be used for the two sample problem for functional data.

There is a question of how to choose the number p. Due to the functional principal components application used here, the method based on the cumulative percentage of total variance (CPV) explained is very popular and effective. Let

$$CPV(p) = \frac{\sum_{k=1}^{p} \hat{\tau}_k}{\sum_{k=1}^{n_1+n_2} \hat{\tau}_k}.$$

One can use p for which $CPV(p)$ exceeds a desired level, e.g., 85% or 95%. Of course, there are also other methods for example those using information criteria or cross-validation.

To illustrate the above two-sample test, we consider three comparisons of mean temperature: Eastern vs. Western, Eastern vs. Northern, and Western vs. Northern. All of them are two sample problems for univariate functional data. The number of functional principal components used in the test was chosen to have $CPV(p) \geq 95\%$. The results are presented in Table 1. As we expected, there are no significant differences in average temperature in Eastern and Western Canada, while Northern Canada is significantly colder than the two other regions. Let us also observe that the number of chosen functional principal components is rather small while satisfying the criterion $CPV(p) \geq 95\%$.

	Eastern-Western	Eastern-Northern	Western-Northern
Test statistic	1.2020	76.8028	26.1506
p-value	0.5483	1.1e-16	2.09e-6
p	2	3	2

Table 1: Results of two sample test for comparing temperature mean functions in pairs of regions in Canada.

3. UNIVARIATE ANALYSIS OF VARIANCE

In this section, we focus on the analysis of variance (ANOVA) for univariate functional data ($d = 1$), which was explored by many authors. We will present various methods in different scenarios. We mainly consider the case of independent observations, but in Section 3.3, we also present the results for the repeated measures ANOVA for functional data.

3.1. ANOVA for independent functional samples

The one-way ANOVA problem for functional data can be formulated in the following way. Let $X_{i1}, \dots, X_{in_i}$ denote i-th sample of independent and identically distributed stochastic processes $SP(\mu_i, \gamma)$ for $i = 1, \dots, k$, where $SP(\mu, \gamma)$ denotes a stochastic process with mean function μ and covariance function γ. The k samples are assumed to be independent. Moreover, we usually suppose that $X_{ij} \in L_2^1(\mathcal{I})$. Let $n = n_1 + \cdots + n_k$ be the total sample size. It is interesting to test the following equality:

$$H_0 : \mu_1 = \cdots = \mu_k \text{ in } L_2^1(\mathcal{I}), \tag{3.1}$$

against $H_1 : \neg H_0$. We can observe that the null hypothesis considered in Section 2 is the special case of (3.1), so almost every test presented in the present section can be applied to the two sample problem for functional data.

3.1.1. Basis function expansion test

We start with the test procedure, which has some common characteristics with that presented in Section 2. Namely, we present the tests based on a basis function representation (expansion) of the functional data proposed in Górecki and Smaga (2015). The two-sample test uses the expansion of functional observation based on functional principal components. However, here we use the coefficients (scores) in the other way than in Section 2. A lot of methods based on basis function expansion can be found in Ramsay and Silverman (2005), for example, canonical correlation and discriminant analysis for functional data. It is also popular in scalar-function regression model (Collazos et al., 2016; Smaga and Matsui, 2018). Moreover, it can be used for the partially observed functional data (Kraus, 2019), which is described in Section 3.2.

Let $\{\varphi_l\}_{l=0}^{\infty}$ be an orthonormal basis of $L_2^1(\mathcal{I})$. Note that the basis is fixed here, while the functional principal components basis is data-dependent. The examples of fixed bases used in the FDA are the Fourier, B-spline, and Legendre. Then, the functional observations can be expanded as follows:

$$X_{ij}(t) = \sum_{l=0}^{\infty} c_{ijl}\varphi_l(t),\ t \in \mathcal{I},$$

where c_{ijl} are random variables with finite variance. Of course, in practice, we are not able to use this infinite representation. However, by taking sufficiently many elements of this expansion, we can approximate arbitrarily well any function in $L_2^1(\mathcal{I})$ (see Ramsay and Silverman (2005)). For this reason, we assume that the observations X_{ij} belong to a finite subspace of $L_2^1(\mathcal{I})$ and they can be expressed as follows:

$$X_{ij}(t) = \sum_{l=0}^{K} c_{ijl}\varphi_l(t),\ t \in \mathcal{I} \tag{3.2}$$

for sufficiently large K. The coefficients c_{ijl} can be estimated, for example, by the least squares method. The optimum K in the sense of smoothness can be selected for each X_{ij} using, for instance, the Bayesian Information Criterion. Then from the values of K corresponding to all processes, a modal value is selected as the common value for all X_{ij}. For more details about estimating coefficients c_{ijl} and choosing K see Krzyśko and Waszak (2013).

The coefficients (scores) c_{ijl} contain some information about the original functional data. For this reason, the tests for (3.1) can be constructed based on them. The first idea is to apply known tests for multivariate data to vectors of coefficients $\mathbf{c}_{ij} = (c_{ij0}, c_{ij1}, \ldots, c_{ijK})^\top$, which represent observations X_{ij}. For the ANOVA problem for functional data, this was done for example by Górecki and Smaga (2015), but this approach was not performing well. Thus, we need a more sophisticated method. Before, we describe it, let us note that for other problems with functional data, the approach presented in this paragraph can give very good results (see, for example, Krzyśko and Waszak (2013)).

The test statistic of the classical ANOVA F-test for real random variables can be adopted for functional data framework in the following way:

$$F = \frac{\frac{1}{k-1}\sum_{i=1}^{k} n_i ||\bar{X}_i - \bar{X}||_1^2}{\frac{1}{n-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i} ||X_{ij} - \bar{X}_i||_1^2}, \tag{3.3}$$

where $\bar{X} = n^{-1}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}$ and $\bar{X}_i = n_i^{-1}\sum_{j=1}^{n_i} X_{ij}$. Using the basis expansion (3.2), the test statistic (3.3) can be expressed just by the coefficients $\mathbf{c}_{ij}$, and it does not depend on the basis functions φ_l. This is shown in the following

Proposition 1 (Górecki and Smaga (2015))**.** The statistic F given by (3.3) is equal to

$$\frac{\frac{1}{k-1}(a-b)}{\frac{1}{n-k}(c-a)},$$

where

$$a = \sum_{i=1}^{k} \frac{1}{n_i} \sum_{m=1}^{n_i}\sum_{s=1}^{n_i} \mathbf{c}_{im}^\top \mathbf{c}_{is},\ b = \frac{1}{n}\sum_{i=1}^{k}\sum_{m=1}^{n_i}\sum_{t=1}^{k}\sum_{v=1}^{n_t} \mathbf{c}_{im}^\top \mathbf{c}_{tv},\ c = \sum_{i=1}^{k}\sum_{j=1}^{n_i} \mathbf{c}_{ij}^\top \mathbf{c}_{ij}.$$

It can also be seen that any permutation of all X_{ij} leaves the values of the sums b and c unchanged. Moreover, it is known that the permutation test is exact under exchangeability, i.e., in our framework, when all data have the same distribution. Thus, Górecki and Smaga (2015) proposed to use the permutation method to construct a test based on (3.3). The resulting test was denoted by FP. The permutation test has the following steps:

1. Compute the value of the test statistic F for the original sample data.
2. Create a new sample from the original data in the following way: From all observations X_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$, select randomly without replacement n_1 observations for the first new group, then from the remainder of the observations choose randomly without replacement n_2 observations for the second new group, and so on.
3. Obtain the value of the test statistic based on the new sample data.
4. Repeat steps 2–3 B times. Let $F_1, \ldots, F_B$ denote the obtained values of the test statistic.
5. The p-value of the permutation test is defined as $(1/B)\sum_{i=1}^{B} I(F_i \geq F)$, where I is the indicator.

This test controls the type I error very well and has comparable power to many other test procedures while being the most powerful for some cases, e.g., the number of observations and number of time points are small or under less correlated functional data (see simulation study in Mrkvička et al. (2020) and in the supplement).

We illustrate the FANOVA test based on basis expansion in comparing the three groups of weather stations in Canada. As a functional variable, we take the temperature once again. To perform the test, we consider the Fourier basis and choose the number of K of basis functions using BIC from the set $\{3, 5, \ldots, 21\}$. Here, we consider just an odd number of basis functions, which is dictated by the appropriate function in the `fda` package. This criterion suggests $K = 7$. Figure 3 presents the raw temperature data and their smoothed version. We can observe that the smoothing eliminated much of the noise in the data and the smoothed data seem to appropriately represent the original ones. The value of the test statistic is equal to 16.35013, while the p-value equals zero. Thus, we reject the null hypothesis (3.1) and conclude that there are significant differences in mean functions in the three regions of Canada.

3.1.2. Tests based on random projections

In this section, we present the projection-based tests proposed by Cuesta-Albertos and Febrero-Bande (2010). In this method, we also project the functional data, but differently than in the previous sections. We will present this method for the one-way classification, but it can be adopted for various other experimental designs. Some other examples of the use of random projections are for functional linear regression (Cuesta-Albertos et al., 2019), outlier detection (Navarro-Esteban and Cuesta-Albertos, 2021), and testing statistical hypotheses (Meléndez et al., 2021).

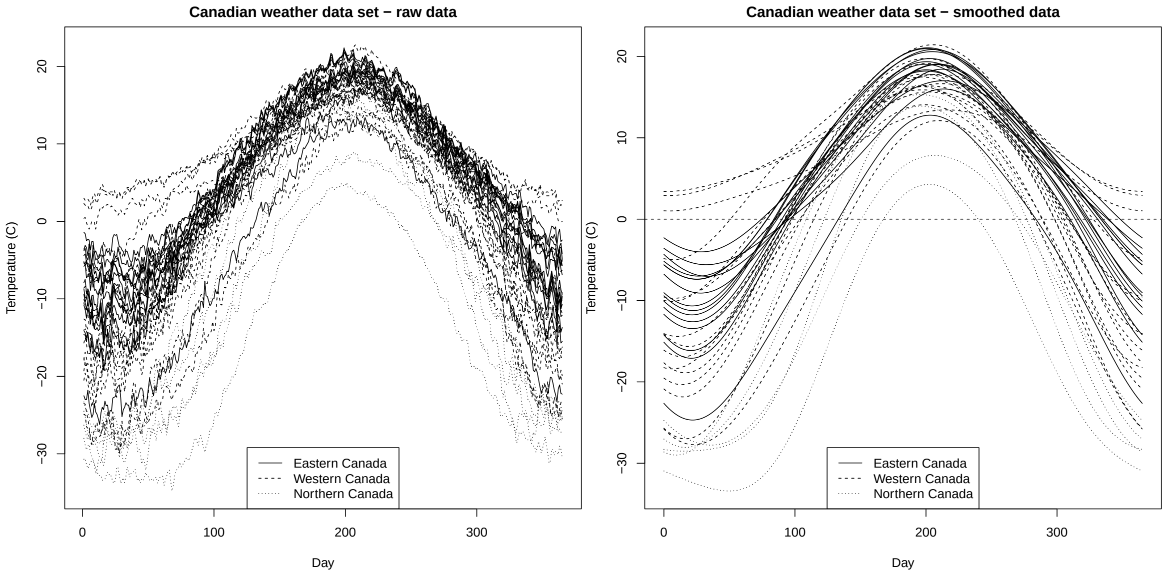

Figure 3: In the left panel, the raw Canada temperature data are presented, while in the right panel, we have the smoothed data in the Fourier basis with seven basis functions.

Let ξ be a Gaussian distribution on $L_2^1(\mathcal{I})$, whose all one-dimensional projections are nondegenerate. Using the distribution ξ, we select randomly a function v from $L_2^1(\mathcal{I})$. When (3.1) holds, then for every $v \in L_2^1(\mathcal{I})$, the following null hypothesis

$$H_0^v : \langle \mu_1, v\rangle_1 = \cdots = \langle \mu_k, v\rangle_1$$

also holds. On the other hand, we also have the following result.

Theorem 3.1 (Cuesta-Albertos and Febrero-Bande (2010)). *Under the above assumptions, if $\mu_i \neq \mu_j$ for some $i \neq j$, then $\xi(v \in L_2^1(\mathcal{I}) : \langle \mu_1, v\rangle_1 = \cdots = \langle \mu_k, v\rangle_1) = 0$.*

Theorem 3.1 implies that if H_0 fails, then for ξ-almost every $v \in L_2^1(\mathcal{I})$, H_0^v also fails. Therefore, we have a kind of equivalence between the null hypotheses H_0 and H_0^v. This justifies the following projection-based test for H_0:

1. Select, with Gaussian distribution, functions $v_m \in L_2^1(\mathcal{I})$, $m = 1, \ldots, l$, $l \geq 1$.
2. Compute
$$\pi_{ij}^m = \frac{\langle X_{ij}, v_m\rangle_1}{\|v_m\|_1} = \frac{\int_{\mathcal{I}} X_{ij}(t) v_m(t) dt}{\sqrt{\int_{\mathcal{I}} v_m^2(t) dt}}$$
for $i = 1, \ldots, k$, $j = 1, \ldots, n_i$, $m = 1, \ldots, l$.
3. For each $m = 1, \ldots, l$, apply the appropriate one-way ANOVA test for π_{ij}^m, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$. Let $p_1, \ldots, p_l$ denote the obtained p-values.
4. Compute the final p-value for H_0 as $\inf\left\{lp_{(m)}/m : m = 1, \ldots, l\right\}$, where $p_{(1)} \leq \cdots \leq p_{(l)}$ are the ordered p-values obtained in step 3.

In step 1, one can use for example a Gaussian white noise or a standard Brownian motion. The simulation results indicate that the projection-based tests with Gaussian white noise (respectively standard Brownian motion) are the most powerful when the functional data are higher (respectively less) correlated.

In step 2, we generate l random projections of all the functional data. For each of these l random projections, we apply some ANOVA test separately in step 3. Thus, we perform l tests on the real data. Moreover, in step 4, we correct the obtained p-values to control the false discovery rate (FDR) using the procedure by Benjamini and Hochberg (1995). The reason for these steps is that we want to eliminate the two drawbacks of the projection method: (1) the loss of information, since a function is replaced by just one real number; (2) some random instability in the procedure, because it may happen that running the procedure twice, we obtain two different decisions.

To perform step 3, we can choose different tests. The first choice can be the standard F test, but then we may need to inspect the projection data in terms of normality or equality of variances. The other possibilities can also be applied. For instance, in the R package `fdANOVA` (Górecki and Smaga, 2019), the ANOVA-type statistic (ATS) proposed by Brunner et al. (1997), and the Wald-type permutation statistic (WTPS) by Pauly et al. (2015) are implemented.

The last thing we should discuss is the number of projections l. It is a hyperparameter and has to be independent concerning the data. The simulation results of Cuesta-Albertos and Febrero-Bande (2010) and Górecki and Smaga (2017) suggest to chose l near 30. However, if needed the greater values of l may also be used. The simulation studies imply that although the greater l results in a more conservative projection-based test (i.e., the empirical sizes are much smaller than the significance level α), the power of the test generally increases with l (see also the supplement).

The random projection test is illustrated for the same problem as in Section 3.1.1. We use different versions of it. Namely, we take different Gaussian processes (Gaussian white noise, standard Brownian motion) and three final tests applied to random projections (standard ANOVA test, ANOVA-type statistic, WTPS). For each of these six tests, we use the number l of random projections ranging from 5 to 30. The p-values of all these tests are presented in Figure 4. We can observe that all the tests reject the null hypothesis about the equality of temperature mean functions. It seems that the ATS test is slightly less stable for different l than the other tests.

3.1.3. Tests based on aggregating pointwise test statistics

In the earlier sections, we have presented the methods, which reduced the dimension of the data by projecting them into finite dimensional spaces. However, some methods apply directly to the discrete functional observations. In particular, such methods were developed in Zhang (2013), Zhang and Liang (2013), and Zhang et al. (2019). The idea of these methods is to appropriately aggregate the values of the pointwise test statistics, which are constructed based on statistics of classical test procedures. The natural applications of tests based on aggregating pointwise test statistics are the test procedures for verifying the equality of

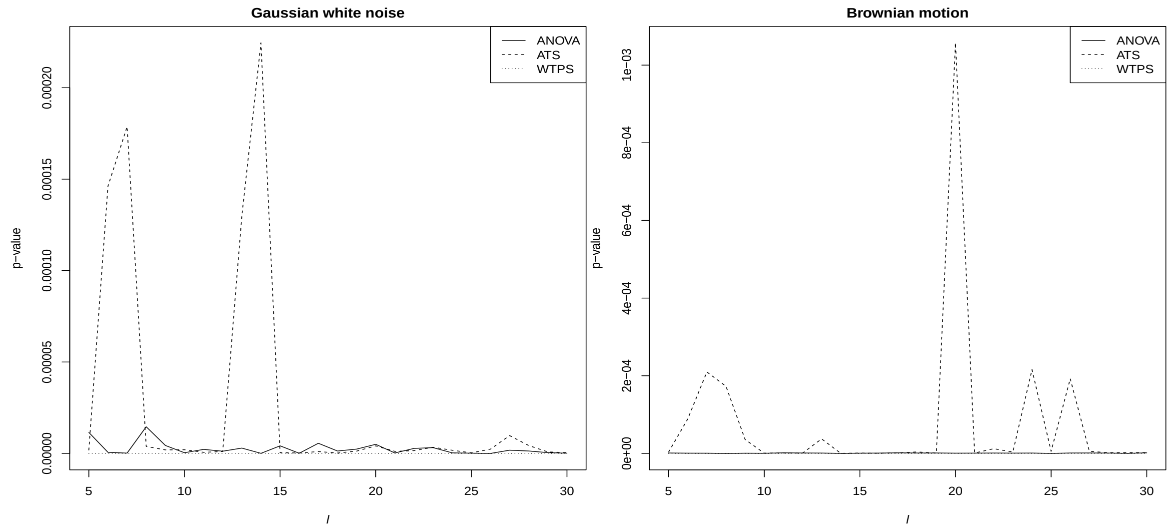

Figure 4: P-values of random projection test for comparing temperature mean functions in three regions in Canada. l denotes the number of projections used.

covariance functions Guo et al. (2018, 2019). However, they can be also used in the functional regression model to test the significance of variables and their selection Smaga (2021).

For $t \in \mathcal{I}$, let

$$\mathrm{SSR}_n(t) = \sum_{i=1}^{k} n_i(\bar{X}_i(t) - \bar{X}(t))^2,\ \mathrm{SSE}_n(t) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}(t) - \bar{X}_i(t))^2$$

denote the pointwise between-subject and within-subject variations respectively. In practice, we obtain values of the about functions in the design time points. To infer based on them, one needs to aggregate them into one real number (more precisely random variable). In the FDA, the first idea for this is integration, which is very popular. Let us first consider the L^2-norm-based test Zhang (2013), which uses the following test statistic:

$$S_n = \int_{\mathcal{I}} \mathrm{SSR}_n(t)dt.$$

Under (3.1), the distribution of S_n has similar shape to $\beta\chi_d^2$, i.e., the distribution of S_n is nonnegative and generally skewed. The parameters β and d are determined by matching the first two moments of S_n and $\beta\chi_d^2$. This method is called the Box-type approximation or the two-cumulant approximation. We have $\beta = \mathrm{tr}(\gamma^{\otimes 2})/\mathrm{tr}(\gamma)$, $d = (k-1)\kappa$, $\kappa = \mathrm{tr}^2(\gamma)/\mathrm{tr}(\gamma^{\otimes 2})$, $\gamma^{\otimes 2}(s,t) = \int_{\mathcal{I}} \gamma(s,u)\gamma(u,t)du$, and $\mathrm{tr}(\gamma) = \int_{\mathcal{I}} \gamma(t,t)dt$ denotes the trace of $\gamma(s,t)$. These quantities are unknown in practice, thus we have to estimate them. The pooled sample covariance function

$$\hat{\gamma}(s,t) = \frac{1}{n-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}(s) - \bar{X}_i(s))(X_{ij}(t) - \bar{X}_i(t)) \tag{3.4}$$

is an unbiased estimator of $\gamma(s,t)$. In the naive method, we just use the plugin method, which results in L^2N tests with the following critical region: $\{(x_{ij}) : S_n > \hat{\beta}\chi^2(1-\alpha,\hat{d})\}$. The bias-reduced method (resulting in L^2B test), we have $\hat{\beta} = \widehat{\mathrm{tr}(\gamma^{\otimes 2})}/\mathrm{tr}(\hat{\gamma})$, $\hat{d} = (k-1)\hat{\kappa}$

and $\hat{\kappa} = \widehat{\mathrm{tr}^2(\gamma)}/\widehat{\mathrm{tr}(\gamma^{\otimes 2})}$, where

$$\widehat{\mathrm{tr}(\gamma^{\otimes 2})} = \frac{(n-k)^2}{(n-k-1)(n-k+2)}\left(\mathrm{tr}(\hat{\gamma}^{\otimes 2}) - \frac{\mathrm{tr}^2(\hat{\gamma})}{n-k}\right),$$
$$\widehat{\mathrm{tr}^2(\gamma)} = \frac{(n-k)(n-k+1)}{(n-k-1)(n-k+2)}\left(\mathrm{tr}^2(\hat{\gamma}) - \frac{2\mathrm{tr}(\hat{\gamma}^{\otimes 2})}{n-k+1}\right).$$

The L^2-norm-based tests take into account just the between-subject variability, while considering also the within-subject variability may result in the use of more information from the data. The first of such procedures is the F-type test Zhang (2013), which uses the test statistic

$$F_n = \frac{\int_{\mathcal{I}} \mathrm{SSR}_n(t)dt/(k-1)}{\int_{\mathcal{I}} \mathrm{SSE}_n(t)dt/(n-k)}.$$

Under (3.1), the distribution of F_n can be approximated by the F-distribution F_{d_1,d_2}, where d_1 is the same as for the L^2-norm-based test and $d_2 = (n-k)\kappa$. To estimate the unknown parameter κ, we can use the same methods as above, i.e., the naive and bias-reduced ones, which result in the FN and FB tests, respectively.

In the above tests, SSR_n and SSE_n were aggregated separately. The next idea is to do this together in the sense of aggregating the values of the pointwise F-test statistic for (3.1), defined as

$$F_n(t) = \frac{\mathrm{SSR}_n(t)}{k-1} \Big/ \frac{\mathrm{SSE}_n(t)}{n-k}.$$

The first test of this type is the globalizing pointwise F-test (the GPF test; Zhang and Liang (2013)), which is based on the test statistic

$$T_n = \int_{\mathcal{I}} F_n(t)dt.$$

Under (3.1), the distribution of T_n can be approximated by $\hat{\beta}_w \chi^2_{\hat{d}_w}$, where

$$\hat{\beta}_w = (n-k-2)\mathrm{tr}(\hat{\gamma}_w^{\otimes 2})/((k-1)(n-k)(b-a)),$$

$$\hat{d}_w = (k-1)(n-k)^2(b-a)^2/((n-k-2)^2\mathrm{tr}(\hat{\gamma}_w^{\otimes 2}))$$

and $\hat{\gamma}_w(s,t) = \hat{\gamma}(s,t)/\sqrt{\hat{\gamma}(s,s)\hat{\gamma}(t,t)}$, where $\hat{\gamma}(s,t)$ is given by (3.4).

In the aforementioned test procedures, the integration was used to aggregate the values of the pointwise test statistics. The other idea for this is taking the supremum. This is less common, but it can result in a very powerful procedure. Zhang et al. (2019) proposed the bootstrap tests based on the following max-type test statistic:

$$F_{\max} = \sup_{t \in \mathcal{I}} \{F_n(t)\}.$$

Note that the $F_{\max}$ method was already introduced earlier in neuroscience literature by Winkler et al. (2014). To construct a test based on $F_{\max}$, the group-wise or pooled bootstrap methods were used, since the distribution of $F_{\max}$ seems to be very complicated. These procedures are similar to the permutation test presented in Section 3.1.1, but selecting new data in step 2 is made with replacement. In the group-wise bootstrap, we draw from each sample separately, while in the pooled bootstrap, we chose from all observations. Such

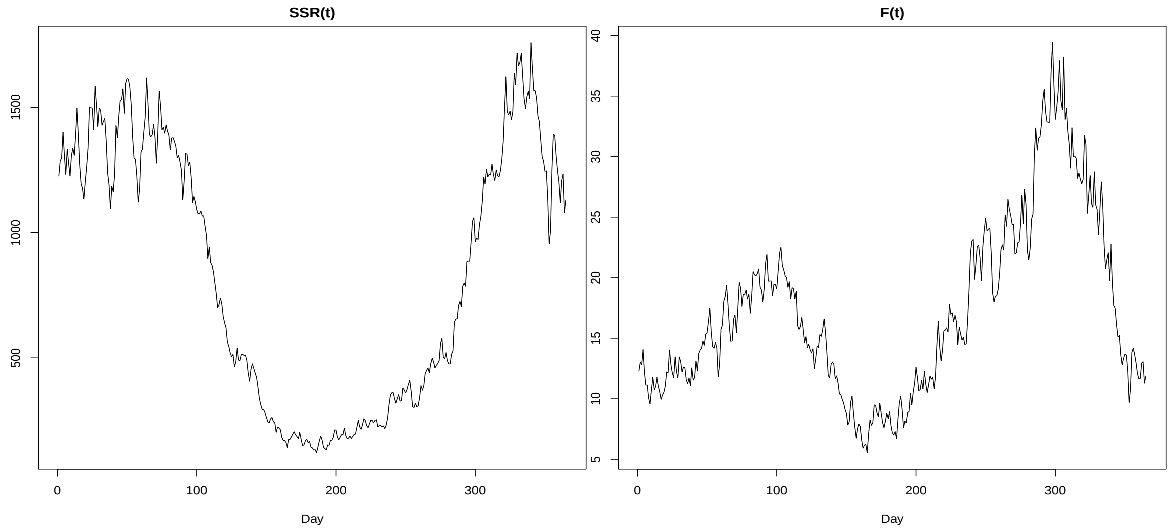

Figure 5: Pointwise test statistics.

bootstrap methods are also recommended for tests based on S_n, F_n, and T_n statistics, when the k samples are not Gaussian and when the sample sizes are small. These bootstrap tests are denoted by L^2b, Fb, and $GPFb$, respectively.

Note that these tests were extended for testing the linear hypotheses and for weighted functional data by Smaga and Zhang (2019, 2020), respectively. In practice, the values of the pointwise test statistics SSR_n, SSE_n, and F_n are calculated in design time points, and then integral and supremum are calculated by sum (or mean) and maximum, respectively.

The simulation results conducted in many papers suggest that the $F_{\max}$ test is very powerful for moderatory and highly correlated functional data. On the other hand, for less correlated functional data, the GPF test seems to be the best, and it is followed by the S_n and F_n tests. See also the supplement for simulation study based on real data example considered in the next paragraph.

Let us now apply the tests based on aggregating pointwise test statistics to the FANOVA problem for temperature in Canada. The pointwise test statistics are presented in Figure 5. They suggest that the temperature in the three regions of Canada is the most different in the first and last months of the year. This is some additional information, which we do not obtain from the earlier tests considered. The significant differences in the three groups of weather stations are confirmed by all tests considered in the present section (see Table 2).

	L^2N	L^2B	Lb	FN	FB	Fb	GPF	$F_{\max}$
Test stat.	303808.1	303808.1	303808.1	16.20986	16.20986	16.20986	17.23131	39.42033
p-value	1.51e-10	2.69e-11	0	2.02e-07	1.25e-07	0	4.78e-13	0
$\hat{\beta}$ or $\hat{d}_1$	6235.936	5773.932		3.005511	3.120999		0.271331	
$\hat{d}$ or $\hat{d}_2$	3.005511	3.120999		48.08817	49.93598		3.931244	

Table 2: Results of tests based on aggregating pointwise test statistics for comparing temperature mean functions in three regions in Canada.

3.1.4. Graphical functional ANOVA test

In this section, we present the one-way graphical functional ANOVA test proposed by Mrkvička et al. (2020). The results of this paper are a continuation of the work done in Myllymäki et al. (2017) and Mrkvička et al. (2017), where a global envelope test was used in multivariate settings. The test is based on permutations and ranks of the discrete functional data, which distinguishes this method from the previously presented ones. Moreover, we obtain a graphical interpretation, which automatically identifies which groups are responsible for the potential rejection and also it identifies which parts of the functions are responsible for the rejection. This may be very important for the interpretation of the result of the test. This innovative graphical approach can also be applied to other problems related to functional data analysis. For example, in Mrkvička et al. (2021), there are considered nonparametric graphical tests of significance in functional general linear models.

We assume that all functions X_{ij} are discretized in the same design time points $\mathbf{t} = (t_1, \ldots, t_m)$. When this is not true, we can apply a smoothing technique (see, for example, Zhang (2013)) and then discretize smoothed X_{ij} in a common grid of design time points.

To test the null hypothesis H_0 in (3.1), one can use the vectors consisting of the sample mean functions calculated in the design time points in $\mathbf{t}$. Namely, let $\bar{\mathbf{X}} = (\bar{X}_1(\mathbf{t}), \ldots, \bar{X}_k(\mathbf{t}))^\top$, where $\bar{X}_i(\mathbf{t}) = (\bar{X}_i(t_1), \ldots, \bar{X}_i(t_m))$, $i = 1, \ldots, k$. The length of the vector $\bar{\mathbf{X}}$ is equal to $k \cdot m$ and it can be quite large. The one-way graphical functional ANOVA test is as follows:

1. Create a permutation sample from the raw functions X_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$.
2. Compute the vector $\bar{\mathbf{X}}$ for the permutation sample.
3. Repeat steps 1 and 2 B times to obtain the vectors $\bar{\mathbf{X}}_1, \ldots, \bar{\mathbf{X}}_B$.
4. Apply the global rank envelope test to $\bar{\mathbf{X}}_1, \ldots, \bar{\mathbf{X}}_B$.

Thus, the above procedure is the application of the global rank envelope test to the vector $\bar{\mathbf{X}}$, which represents functional data. Let us now describe the global rank envelope test. Assume that $\mathbf{V} = (V_1, \ldots, V_d)$ is a multivariate vector. Let $\mathbf{V}_1 = (V_{11}, \ldots, V_{1d}), \ldots, \mathbf{V}_B = (V_{B1}, \ldots, V_{Bd})$ be the vectors generated from $\mathbf{V}$ by permutation of its components under the null hypothesis, and $\mathbf{V}_1 = \mathbf{V}$. The test is based on ranks of the components of vectors $\mathbf{V}_i$. For $m \in \{1, \ldots, d\}$, let R_{im} be the rank of an element V_{im} among $V_{1m}, \ldots, V_{Bm}$ (the elements of vectors $\mathbf{V}_1, \ldots, \mathbf{V}_B$ for the mth coordinate) such that the lowest ranks correspond to the most extreme values of the statistics, $i = 1, \ldots, B$. Moreover, consider the vectors of pointwise ordered ranks:

$$\mathbf{R}_i = (R_{i[1]}, \ldots, R_{i[d]}),\ i = 1, \ldots, B,$$

where

$$\{R_{i[1]}, \ldots, R_{i[d]}\} = \{R_{i1}, \ldots, R_{id}\},\ R_{i[m]} \le R_{i[m']},\ m \le m'.$$

Then, the extreme rank length ordering for the vector $\mathbf{R}_i$ is defined as follows:

$$R_i^{erl} = \frac{1}{B} \sum_{j=1}^{B} I(\mathbf{R}_j \prec \mathbf{R}_i),$$

where $\mathbf{R}_j \prec \mathbf{R}_i \;\Leftrightarrow\; \exists l \leq d : R_{j[m]} = R_{i[m]} \forall m < l,\; R_{j[l]} < R_{i[l]}$. Using this order, we can define the following p-value based on the extreme rank length ordering:

$$p^{erl} = 1 - \frac{1}{B}\sum_{i=1}^{B} I(\mathbf{R}_1 \prec \mathbf{R}_i).$$

Mrkvička et al. (2020) proposed a new graphical envelope for graphical interpretation. Assume that all the vectors $\mathbf{V}_i$ follow the same joint distribution. Then, one can construct rank envelopes with level $1-\alpha$ as sets $\{\mathbf{V}_l^\alpha, \mathbf{V}_u^\alpha\}$ such that $P(\exists m \in \{1,\dots,d\} : V_m \notin [V_{l,m}^\alpha, V_{u,m}^\alpha]) \leq \alpha$. The contruction is as follows: Let $R_{(\alpha)}^{erl}$ be the largest value in the set $\{R_1^{erl},\dots,R_B^{erl}\}$ for which

$$\frac{1}{B}\sum_{i=1}^{B} I(R_i^{erl} < R_{(\alpha)}^{erl}) \leq \alpha.$$

Then, we can define $I_\alpha = \{i \in \{1,\dots,s\} : R_i^{erl} \geq R_{(\alpha)}^{erl}\}$, i.e., the index set of vectors whose extreme rank length measure is larger than or equal to the critical value $R_{(\alpha)}^{erl}$. Finally, the global extreme rank length envelope is defined as follows:

$$V_{l,m}^\alpha = \min_{i\in I_\alpha} V_{im},\;\; V_{u,m}^\alpha = \max_{i\in I_\alpha} V_{im}.$$

It can be shown that the inference based on the p^{erl} and the global envelope specified by $V_{l,m}^\alpha$ and $V_{u,m}^\alpha$ are equivalent. For the one-way graphical functional ANOVA test, it is formulated in the following

Theorem 3.2 (Mrkvička et al. (2020))**.** *Consider a one-way graphical functional analysis of variance test with test vector $\bar{\mathbf{X}}$. Let $\bar{X}_{l,m}^\alpha$ and $\bar{X}_{u,m}^\alpha$ define the $100(1-\alpha)\%$ global extreme rank length envelope. Then, assuming that there are no pointwise ties with probability 1 in the stochastic process $SP(\mu,\gamma)$, it holds that:*

1. *$\bar{X}_{1,m} < \bar{X}_{l,m}^\alpha$ or $\bar{X}_{1,m} > \bar{X}_{u,m}^\alpha$ for some m if and only if $p^{erl} \leq \alpha$, in which case the null hypothesis is rejected;*
2. *$\bar{X}_{l,m}^\alpha \leq \bar{X}_{1,m} \leq \bar{X}_{u,m}^\alpha$ for all m if and only if $p^{erl} > \alpha$, in which case the null hypothesis holds.*

Mrkvička et al. (2020) conducted the simulation study, where they showed that their test controls the type I error appropriately, and can be a good competitor to other test procedures in terms of power (see also the supplement). However, it has an advantage in interpretation, which in particular, we show in the following paragraph for the Canadian weather data set.

When we apply the graphical functional ANOVA test to the temperature in the three groups of weather stations in Canada, we obtain the p-value equal to 0.002, and thus we reject the null hypothesis (3.1) similarly as before. However, additionally, we obtain a graphical interpretation of the results. It is presented in Figure 6. We can see there the graphical envelope for group means. For Eastern and Northern weather stations, we observe the values of sample mean functions outside the envelope for some design time points. This suggests the areas, where significant differences appear. For Eastern stations, this appears for the

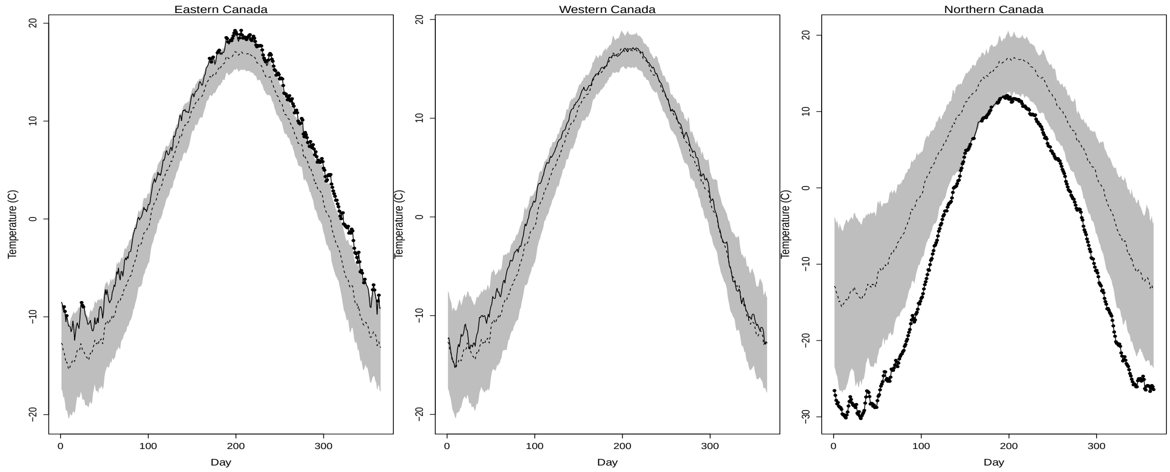

Figure 6: The one-way graphical functional ANOVA test for equality of means of the temperature in Canada in the three groups using the group means.

second half of the year, while for the Northern ones, this is true for almost the whole year. It seems that there are no significant differences between the Western stations. Therefore, the graphical functional ANOVA test can give more interpretation, which is important from a practical point of view.

3.2. ANOVA for partially observed functional data

In this section, we present the tests for partially observed functional data proposed by Kraus (2019). In the FDA, it is usually assumed that all functional data are completely, densely, or sparsely observed in the same domain. Some applications have brought attention to situations where each functional observation may be observed only on a subset of the domain while no information about it is available on the complement. In such a way, we obtain partially observed functional data. A nice example from medical research Kraus (2019) is as follows: In the experiment, the evolution of the heart rate of humans was measured in the period from 8 PM to 2 AM. Unfortunately, many of the participants of the experiment were switching off the measuring device even for a few hours during this period of time. Thus, the corresponding curves are not fully known. The proposed test procedures are the adaptation of the tests presented in Sections 3.1.3 and 2. We refer to Kraus (2019) for references to other statistical methods for partially observed functional data.

We assume that the values of the functions X_{ij} are available on a subset O_{ij} of $\mathcal{I}$, with no information on the complement of O_{ij}. The observation sets may be non-random or random, but they are assumed to be mutually independent and independent of X_{ij} and to consist of a finite union of intervals. Moreover, Kraus (2019) considered the heteroscedastic case, i.e., the covariance functions in groups are γ_i and do not have to be equal.

The group mean functions μ_i and the common mean function μ are estimated as follows:

$$\hat{\mu}_i(t) = \frac{J_i(t)}{N_i(t)} \sum_{j=1}^{n_i} O_{ij}(t) X_{ij}(t),\ \hat{\mu}(t) = \sum_{i=1}^{k} \hat{w}_i(t)\hat{\mu}_i(t),$$

where $i = 1, \dots, k$, $t \in \mathcal{I}$, $O_{ij}(t)$ is the indicator that the value of $X_{ij}(t)$ is observed, $N_i(t) = \sum_{j=1}^{n_i} O_{ij}(t)$ is the number of available observations at time t in the ith sample, $J_i(t) = I(N_i(t) > 0)$,

$$\hat{w}_i(t) = \frac{N_i(t)/\hat{g}_i^2}{\sum_{m=1}^{k} N_m(t)/\hat{g}_m^2}$$

is weight corresponding to the ith group, $\hat{g}_i^2 = \mathrm{tr}(\hat{\gamma}_i)$, and $\hat{\gamma}_i$ is the estimator of the covariance function in the ith group of the form:

$$\hat{\gamma}_i(s,t) = \frac{J_i(s,t)}{M_i(s,t)} \sum_{j=1}^{n_i} U_{ij}(s,t)(X_{ij}(s) - \hat{\mu}_{ist}(s))(X_{ij}(t) - \hat{\mu}_{ist}(t)),$$

$U_{ij}(s,t) = O_{ij}(s)O_{ij}(t)$, $M_i(s,t) = \sum_{j=1}^{n_i} U_{ij}(s,t)$, $J_i(s,t) = I(M_i(s,t) > 0)$, and

$$\hat{\mu}_{ist}(s) = \frac{J_i(s,t)}{M_i(s,t)} \sum_{j=1}^{n_i} U_{ij}(s,t) X_{ij}(s).$$

The first test procedure is similar to the F-type test considered in Section 3.1.3. Namely, the test statistic is as follows:

$$T_{L^2} = \sum_{i=1}^{k} \frac{\int_{\mathcal{I}} N_i(t)^{1/2}(\hat{\mu}_i(t) - \hat{\mu}(t))^2 dt}{\hat{g}_i^2},$$

and we reject the null hypothesis for large values of this test statistic.

The second test is similar to the test based on the functional principal components presented in Section 2. Let $\hat{\varphi}_1, \dots, \hat{\varphi}_p$ be orthonormal linearly independent functions in $L_2^1(\mathcal{I})$, which can be deterministic (as in Section 3.1.1) or data dependent (as in Section 2). Then, the scores of the standardized contrast processes concerning the basis functions $\hat{\varphi}_l$ are defined as follows:

$$Q_{il} = \langle N_i(\cdot)(\hat{\mu}_i(\cdot) - \hat{\mu}(\cdot)), \hat{\varphi}_l \rangle_1 / (\hat{g}_i n_i^{1/2}),$$

$i = 1, \dots, k$, $l = 1, \dots, p$. The score test statistic is given in the following way:

$$T_Q = \mathbf{Q}^\top \hat{\mathbf{V}}^- \mathbf{Q},$$

where $\mathbf{Q} = (Q_{11}, \dots, Q_{1l}, \dots, Q_{k1}, \dots, Q_{kl})^\top$ and $\hat{\mathbf{V}}^-$ is the pseudoinverse of the estimated $(kl) \times (kl)$ covariance matrix of $\mathbf{Q}$, whose entries are as follows:

$$\hat{\mathbf{V}}_{jl,fg} = \int_{\mathcal{I}^2} \hat{\pi}_j(s)^{1/2} \hat{\varphi}_l(s) \hat{v}_{jf}(s,t) \hat{\varphi}_g(t) \hat{\pi}_f(t)^{1/2} ds dt,$$

where $j, f = 1, \dots, k$, $l, g = 1, \dots, l$, $\hat{\pi}_j(t) = N_j(t)/n_j$,

$$\hat{v}_{jf}(s,t) = \frac{\sum_{i=1}^{k} (\delta_{ji} - N_j(s)^{1/2} \hat{w}_i(s) N_i(s)^{-1/2}) \hat{\kappa}_i(s,t) (\delta_{fi} - N_f(t)^{1/2} \hat{w}_i(t) N_i(t)^{-1/2})}{\hat{g}_j^2},$$

δ_{ji} is the Kronecker delta, and $\hat{\kappa}_i$ is an estimator of kernel of some operator (see Section 2.1 in Kraus (2019)).

The asymptotic null distribution of T_{L^2} and T_Q is given in the following theorem.

Theorem 3.3 (Kraus (2019))**.** *For $i = 1, \dots, k$, assume that $n_i \to \infty$, $n_i/n \to a_i > 0$, $E\|X_{i1}\|_1^2 < \infty$ and all eigenvalues of γ_i have multiplicity 1. Let $\pi_{ij}(t) = P(O_{ij}(t) = 1)$ and $\bar{\pi}_i(t) = n_i^{-1}\sum_{j=1}^{n_i}\pi_{ij}(t)$, and there be a function $\pi_i(t)$ such that $\inf_{t\in\mathcal{I}}\pi_i(t) > 0$ and $\sup_{t\in\mathcal{I}}|\bar{\pi}_i(t) - \pi_i(t)| \to 0$ for $n_i \to \infty$. Then, under the null hypothesis:*

1. *T_{L^2} is asymptotically distributed as $\sum_{l=1}^{\infty}\tau_l C_l$, where C_l are independent random variables of the chi-square distribution with one degree of freedom and τ_l can be consistently estimated by the eigenvalues of $\hat{\mathbf{V}}$.*
2. *Assume that there exist linearly independent non-random functions $\varphi_1, \dots, \varphi_p$ such that $\|\hat{\varphi}_l - \varphi_l\| \xrightarrow{p} 0$ for $l = 1, \dots, p$. Then, T_Q is asymptotically chi-square distributed with $(k-1)p$ degrees of freedom.*

By Theorem 3.3, the asymptotic tests based on T_{L^2} and T_Q can be contructed. However, the computational problems appear. Namely, for discrete functional data, the computation of $\hat{\mathbf{V}}$ is very demanding in terms of computer memory, making the computation almost impossible or time-consuming. For this reason, Kraus (2019) proposed the following bootstrap-based test procedures:

1. Calculate $\hat{\mu}_i$ and $\hat{\mu}$.
2. Calculate T_{L^2} and the score vector $\mathbf{Q}$.
3. Set $X_{ij0} = X_{ij} - \hat{\mu}_i + \hat{\mu}$.
4. For $b = 1, \dots, B$:
   - For each $i = 1, \dots, k$, sample with replacement from fragments $X_{i10}, \dots, X_{in_i0}$ to get $X_{i10}^*, \dots, X_{in_i0}^*$.
   - Calculate $T_{L^2,b}^*$ and score vector $\mathbf{Q}_b^*$ from $X_{i10}^*, \dots, X_{in_i0}^*$, $i = 1, \dots, k$.
5. Approximate the p-value of the F-type test using T_{L^2} and $T_{L^2,b}^*$, $b = 1, \dots, B$.
6. Calculate the empirical covariance matrix $\hat{\mathbf{V}}^*$ of $\mathbf{Q}_b^*$, $b = 1, \dots, B$ and $T_Q = \mathbf{Q}^\top\hat{\mathbf{V}}^{*-}\mathbf{Q}$.
7. Approximate the p-value of the projection test using T_Q and the $\chi^2((k-1)p)$ distribution.

The simulation results of Kraus (2019) indicate that both bootstrap tests control the type I error level and have sensible power, but the T_Q test may be too liberal. This is confirmed in the supplement.

Unfortunately, we do not have access to the heart rate data set. Thus, we will illustrate the ANOVA test for partially observed functional data using the artificially modified temperature data. Namely, we use one of the schemas from simulation studies in Kraus (2019) to

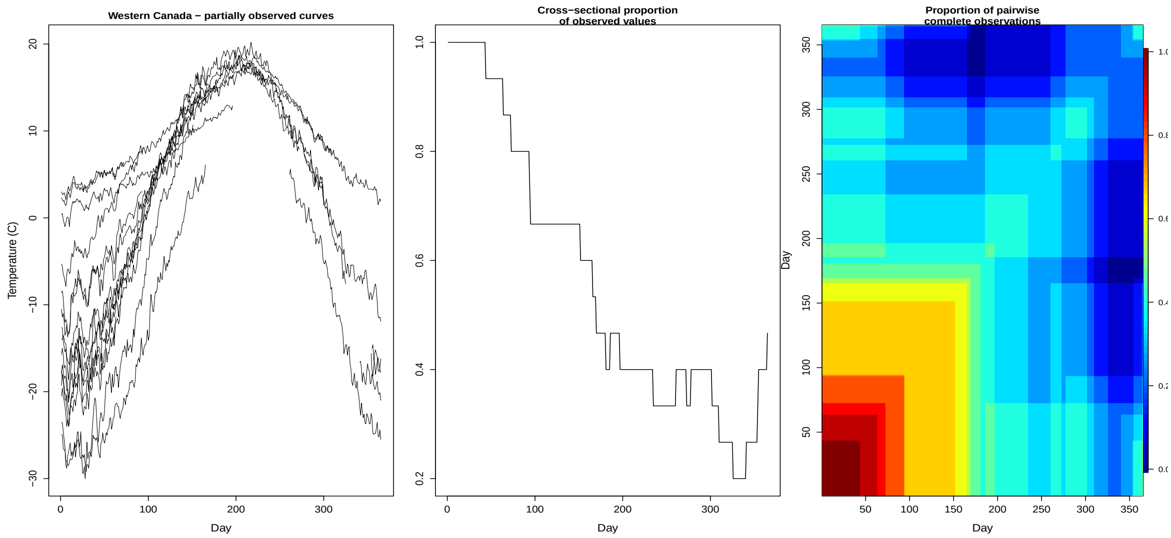

Figure 7: Partially observed temperature data and the summary of missingness.

remove a part of the trajectories of the temperature. Figure 7 presents the obtained partially observed temperature data for Western stations. In this particular case, for many stations, the temperature data are available just for the first half of the year. Having so many gaps in the trajectories, it is not possible to effectively apply the test procedures presented in previous sections. Thus, we use the ANOVA test for partially observed functional data. The values of test statistics are $T_{L^2} = 35.55663$ and $T_Q = 199.4411$. The corresponding p-values of both tests are equal to zero. Thus, we reject the equality of mean functions in the three groups, obtaining the same decision as for the fully observed functional data.

3.3. ANOVA for functional repeated measurements

In this section, we present the results of the repeated measures ANOVA for functional data. For two sample and one-way ANOVA problems, the tests for this problem were considered in Martínez-Camblor and Corral (2011), Smaga (2019a,b, 2020), and Kuryło and Smaga (2024). They are based on the ideas of test procedures presented in Sections 3.1.2 and 3.1.3. On the other hand, Acal and Aguilera (2023) considered the two-way ANOVA for repeated functional observations using the basis function expansion test as in Section 3.1.1. Finally, let us mention that Ditzhaus and Gaigall (2022) developed method for testing marginal homogeneity in paired design in Hilbert spaces.

The functional repeated measures analysis problem can be formulated as follows: We have n subjects subjected to $\ell \geq 2$ (possibly) different conditions. The results of the experiments are functional observations. Thus, we have a functional sample consisting of independent stochastic processes $X_1, \dots, X_n$ defined on the interval $[0, \ell]$. Assume that they satisfy the following model proposed by Martínez-Camblor and Corral (2011):

$$X_j(t) = \mu(t) + e_j(t),\ j = 1, \dots, n,\ t \in [0, \ell], \tag{3.5}$$

where μ is a fixed mean function, and e_j is a random process with zero mean function and covariance function γ. Here, $t \in [0,1]$ corresponds to the first experimental condition, $t \in [1,2]$ to the second, and so on. Thus, in this model, we ignore the possible time periods between repetitions of the experiment, but this does not mean that they do not exist. It is interesting to test the following null hypothesis:

$$H_0 : \mu(t) = \mu(t+1) = \cdots = \mu(t+(\ell-1)),\ t \in [0,1]. \tag{3.6}$$

Rejecting H_0 determines the presence of significant differences in the mean functions corresponding to the experimental conditions.

Let us start with the tests that aggregate the pointwise test statistics (see Section 3.1.3). Martínez-Camblor and Corral (2011) and Smaga (2019b) proposed the tests based on the following test statistic of the L^2-norm type:

$$\mathcal{C}_n(\ell) = \int_0^1 \mathrm{SSA}(t)dt,$$

where

$$\mathrm{SSA}(t) = n\sum_{i=1}^{\ell}(\bar{X}_{i\cdot}(t) - \bar{X}(t))^2,\ t \in [0,1],$$

is the pointwise sum of squares due to the hypothesis,

$$\bar{X}_{i\cdot}(t) = n^{-1}\sum_{j=1}^{n} X_j(t+(i-1)),\ \bar{X}(t) = N^{-1}\sum_{i=1}^{\ell}\sum_{j=1}^{n} X_j(t+(i-1)),$$

$N = n\ell$, $i = 1,\dots,\ell$. On the other hand, Kuryło and Smaga (2024) proposed to use the following test statistics:

$$\mathcal{D}_n(\ell) = \int_0^1 F(t)dt, \quad \mathcal{E}_n(\ell) = \sup_{t\in[0,1]} F(t),$$

where

$$F(t) = \frac{\mathrm{SSA}(t)/(\ell-1)}{\mathrm{SSR}(t)/((\ell-1)(n-1))},\ t \in [0,1],$$

is the pointwise counterpart of the F-type test statistic of classical test for real variables,

$$\mathrm{SSR}(t) = \sum_{i=1}^{\ell}\sum_{j=1}^{n}(X_j(t+(i-1)) - \bar{X}_{i\cdot}(t) - \bar{X}_{\cdot j}(t) + \bar{X}(t))^2$$

and $\bar{X}_{\cdot j}(t) = (1/\ell)\sum_{i=1}^{\ell} X_j(t+(i-1))$, $j = 1,\dots,n$. For the paired two-sample problem for functional data, these test statistics were reduced to those considered by Smaga (2020).

To approximate the null distribution of test statistics, Kuryło and Smaga (2024) used different resampling procedures. As the first method, they consider the permutation approach proposed by Martínez-Camblor and Corral (2011) and Smaga (2020). In the permutation procedure ($\mathcal{P}_1$ for short), a permutation sample was created as follows: For each $j = 1,\dots,n$ separately, randomly permute the observations $X_j(t), X_j(t+1),\dots,X_j(t+(\ell-1))$, $t \in [0,1]$ corresponding to ℓ experimental conditions, and use them to form the permuted observation $X_j^b(t)$, $t \in [0,\ell]$.

In the second permutation approach ($\mathcal{P}_2$ for short), permutation of the data is performed as follows: We draw $X_1^b(t), \ldots, X_n^b(t)$ for $t \in [0,1]$ randomly without replacement from the set

$$\mathcal{A} = \{X_1(t), \ldots, X_n(t), X_1(t+1), \ldots, X_n(t+1), \ldots, X_1(t+(\ell-1)), \ldots, X_n(t+(\ell-1))\}$$

for $t \in [0,1]$ containing all N observations; after that we draw $X_1^b(t), \ldots, X_n^b(t)$ for $t \in [1,2]$ randomly without replacement from the remaining elements in $\mathcal{A}$, and so on.

The first bootstrap approach is the nonparametric bootstrap approach ($\mathcal{B}_1$ for short). In this approach, to select independent bootstrap samples $X_1^b(t), \ldots, X_n^b(t)$, $t \in [0,\ell]$, one draws with replacement from the original sample $X_1(t), \ldots, X_n(t)$, $t \in [0,\ell]$. Moreover, SSA for bootstrap samples is calculated as follows:

$$\mathrm{SSA}^b(t) = n \sum_{i=1}^{\ell} (\bar{X}_{i\cdot}^b(t) - \bar{X}_{i\cdot}(t) - \bar{X}^b(t) + \bar{X}(t))^2,$$

where $t \in [0,1]$, and $\bar{X}_{i\cdot}^b$ and $\bar{X}^b$ are the appropriate sample means computed on the bootstrap sample.

In the second nonparametric bootstrap approach ($\mathcal{B}_2$ for short), we first have to center the observations, i.e., $X_{1,c}(t) = X_1(t) - \bar{X}_\bullet(t), \ldots, X_{n,c}(t) = X_n(t) - \bar{X}_\bullet(t)$ for $t \in [0,\ell]$, where $\bar{X}_\bullet(t) = n^{-1}\sum_{j=1}^n X_j(t)$ is the sample mean function for the original sample $X_1, \ldots, X_n$. Second, $X_1^b(t), \ldots, X_n^b(t)$ for $t \in [0,1]$ are randomly drawn with replacement from $X_{1,c}(t), \ldots, X_{n,c}(t)$, $t \in [0,1]$; independently, $X_1^b(t), \ldots, X_n^b(t)$ for $t \in [1,2]$ are randomly drawn with replacement from $X_{1,c}(t), \ldots, X_{n,c}(t)$, $t \in [1,2]$, and so on.

In the parametric bootstrap approach ($\mathcal{B}_3$ for short), we generate the bootstrap samples from the Gaussian process with zero mean function and covariance function equal to

$$\hat{\gamma}(s,t) = \frac{1}{n-1}\sum_{j=1}^n (X_j(s) - \bar{X}_\bullet(s))(X_j(t) - \bar{X}_\bullet(t)),\ s,t \in [0,\ell].$$

The simulation studies suggest that the best type I error control and power are achieved by the $\mathcal{E}_n(\ell)$-based $\mathcal{B}_1$ and $\mathcal{B}_3$ tests. However, the tests based on the test statistic $\mathcal{D}_n(\ell)$ (especially permutation tests) also have good finite sample properties.

The other solution to test the null hypothesis (3.6) is the projection method similar to that presented in Section 3.1.2 (Smaga, 2019a). Let $\mu_i(t) = \mu(t+(i-1))$, $t \in [0,1]$, $i = 1, \ldots, l$. Then, we can proceed similarly as in Section 3.1.2 by identifying $X_{ij}(t)$ with $X_i(t+(j-1))$ and taking $\mathcal{I} = [0,1]$ and $k = \ell$. Of course, to perform step 3, we can apply different test procedures for repeated measures ANOVA problem, which may be preceded by a projection data inspection, for example, in terms of normality. In simulations, Smaga (2019a) used a standard repeated measures ANOVA test. The main conclusions were similar to those in Section 3.1.2.

Since in this section, we consider the repeated measures ANOVA, we use another example than Canadian weather data. We show the data example investigated in Kuryło and Smaga (2024). It uses the DTI data available in the R package `refund` (Goldsmith et al., 2022). The DTI is a magnetic resonance imaging technique providing different measures of

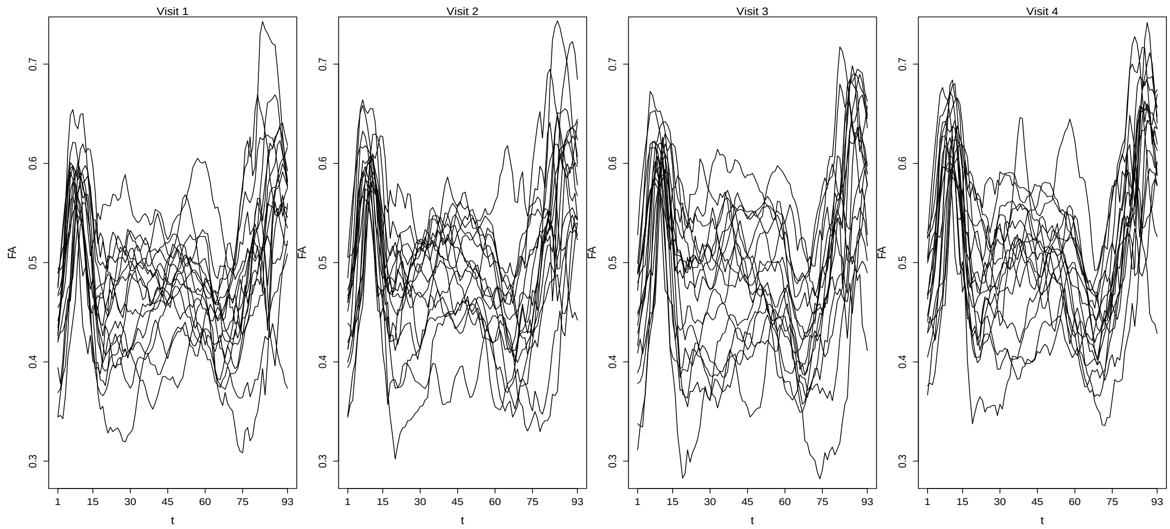

Figure 8: Trajectories for the FA profiles at four visits.

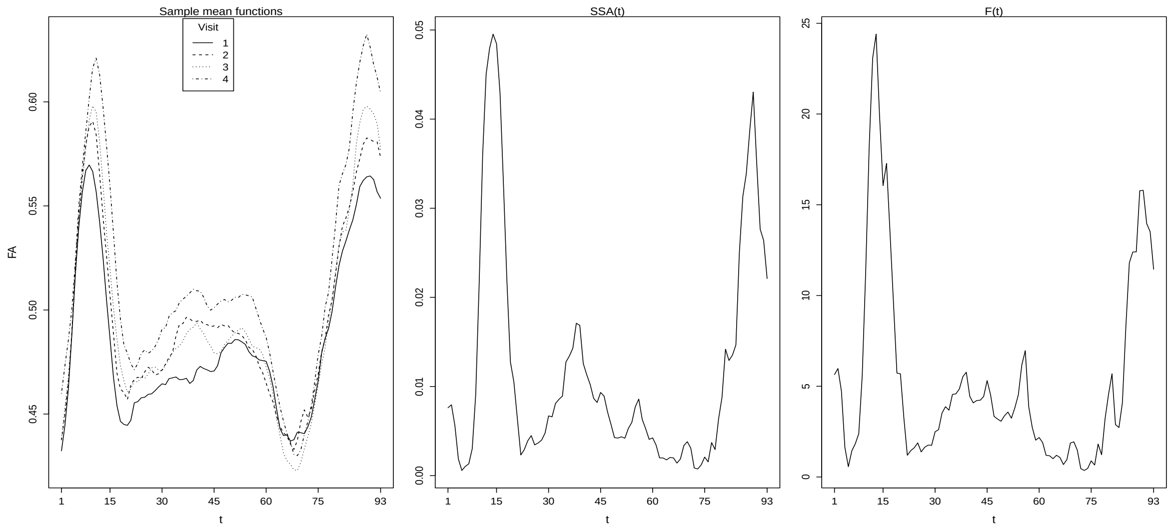

Figure 9: Sample mean functions of the FA at different visits and pointwise test statistics.

water diffusivity along brain white matter tracts. In the experiment, DTI brain scans were recorded for multiple sclerosis (MS) patients at several visits. The aim was to assess the effect of neurodegeneration on disability. Using DTI, the fractional anisotropy (FA) tract profiles for the corpus callosum (CCA) were determined. For each patient and each visit, the FA profiles are measured in 93 design time points. Thus, we treat these data as functional data on $[1, 93]$.

For illustrative purposes, Kuryło and Smaga (2024) selected the observations for $n = 17$ patients with $\ell = 4$ successive visits. They are presented in Figure 8. It is of interest to check whether the mean functions of the FA at different visits are the same or not. The sample mean functions as well as the pointwise test statistics are presented in Figure 9. The sample mean functions suggest that the mean functions increase at subsequent visits. The p-values of the ANOVA for functional repeated measurements are presented in Table 3 and Figure 10. We observe that almost all tests reject the equality of FA profiles for various visits. The

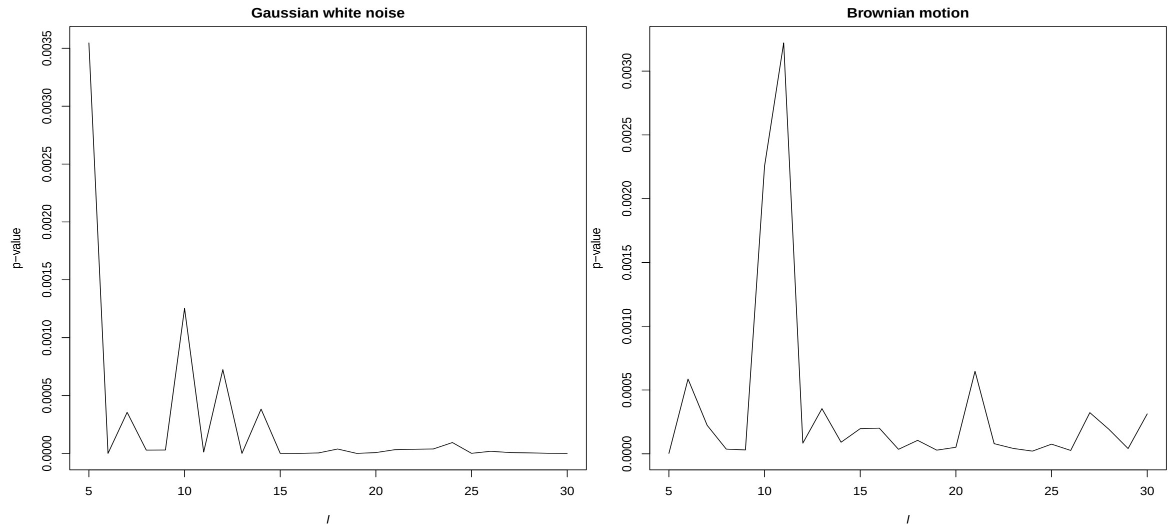

Figure 10: P-values of the random projection tests by Smaga (2019a) for the DTI data set.

exceptions are the $\mathcal{P}_2$ and $\mathcal{B}_2$ $\mathcal{C}_{17}(4)$ tests, which have larger p-values due to their very conservative character.

	p-values				
Test statistic	$\mathcal{P}_1$	$\mathcal{P}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_3$
$\mathcal{C}_{17}(4) = 1.096$	0	0.321	0	0.317	0
$\mathcal{D}_{17}(4) = 494.1$	0	0.001	0	0	0
$\mathcal{E}_{17}(4) = 24.4$	0	0	0	0	0

Table 3: Results of the aggregating the pointwise test statistics procedures by Kuryło and Smaga (2024) for the DTI data set.

4. MULTIVARIATE ANALYSIS OF VARIANCE

In the previous sections, we have studied the ANOVA problem for univariate functional data, i.e., we have measured one functional variable, e.g., the temperature in given locations. However, we can measure many functional features for one experimental unit, e.g., the temperature, precipitation, atmospheric pressure, and wind speed in given locations. Then, we have multivariate functional data and the MANOVA for them. The multivariate analysis of variance for functional data was first considered in Górecki and Smaga (2017), while in Qiu et al. (2021), the two sample tests were considered. Other statistical problems are also considered for multivariate functional data. To show just a few examples, we refer to the following papers: Tokushige et al. (2007) and Jacques and Preda (2014) - clustering; Górecki et al. (2018) - principal component analysis, discriminant coordinates, canonical correlation analysis; Górecki et al. (2019) - variable selection in multivariate functional data classification; Górecki et al. (2020) - independence testing. Some recent applications of multivariate functional data analysis are as follows: comparative study of countries' competitiveness (Krzyśko et al., 2022); analysis of data observed on different (dimensional) domains (Happ and Greven, 2018); analysis of air pollution in the South of France (Bouveyron et al., 2022).

Let $\mathbf{X}_{ij} = (X_{ij1}, \dots, X_{ijd})^\top$, $i = 1, \dots, k$, $j = 1, \dots, n_i$ denote k groups of vectors of stochastic processes belonging to the space $L_2^d(\mathcal{I})$. Let $SP_d(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ denote a d-dimensional stochastic process with mean vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^\top$ and covariance function $\boldsymbol{\Gamma}$. Assuming that the $\mathbf{X}_{ij}$ are i.i.d. $SP_d(\boldsymbol{\mu}_i, \boldsymbol{\Gamma})$, $i = 1, \dots, k$, $j = 1, \dots, n_i$, we are interested in testing the equality of k mean vectors

$$H_0 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k \text{ in } L_2^d(\mathcal{I}). \tag{4.1}$$

For testing H_0, Górecki and Smaga (2017) proposed the permutation tests based on a basis function representation of the components of the stochastic processes $\mathbf{X}_{ij}$, $i = 1, \dots, k$, $j = 1, \dots, n_i$, similarly to the test presented in Section 3.1.1. We can assume that

$$X_{ijm}(t) = \sum_{l=0}^{K_m} \alpha_{ijml} \varphi_l(t), \ t \in \mathcal{I}, \ m = 1, \dots, d, \tag{4.2}$$

where $\{\varphi_l\}_{l=0}^\infty$ is an orthonormal basis of the space $L_2^1(\mathcal{I})$, and α_{ijml}, $l = 0, \dots, K_m$, are random variables and $\mathrm{Var}(\alpha_{ijml}) < \infty$. We can estimate the coefficients in the same way as described in Section 3.1.1. Let

$$\begin{aligned} KM &= \max\{K_1, \dots, K_d\}, \\ \boldsymbol{\alpha}_{ijm} &= (\alpha_{ijm0}, \dots, \alpha_{ijmK_m}, 0, \dots, 0) \in \mathbb{R}^{KM+1}, \\ \boldsymbol{\varphi} &= (\varphi_0, \dots, \varphi_{KM})^\top. \end{aligned}$$

Then

$$\mathbf{X}_{ij}(t) = \begin{pmatrix} \boldsymbol{\alpha}_{ij1} \\ \vdots \\ \boldsymbol{\alpha}_{ijd} \end{pmatrix} \boldsymbol{\varphi}(t) = \boldsymbol{\alpha}_{ij} \boldsymbol{\varphi}(t), \ t \in \mathcal{I}. \tag{4.3}$$

Let

$$\begin{aligned} \mathbf{E} &= \sum_{i=1}^k \sum_{j=1}^{n_i} \int_{\mathcal{I}} \left(\mathbf{X}_{ij}(t) - \bar{\mathbf{X}}_i(t)\right) \left(\mathbf{X}_{ij}(t) - \bar{\mathbf{X}}_i(t)\right)^\top dt, \\ \mathbf{H} &= \sum_{i=1}^k n_i \int_{\mathcal{I}} \left(\bar{\mathbf{X}}_i(t) - \bar{\mathbf{X}}(t)\right) \left(\bar{\mathbf{X}}_i(t) - \bar{\mathbf{X}}(t)\right)^\top dt, \end{aligned}$$

where $\bar{\mathbf{X}}_i(t) = n_i^{-1} \sum_{j=1}^{n_i} \mathbf{X}_{ij}(t)$ and $\bar{\mathbf{X}}(t) = n^{-1} \sum_{i=1}^k \sum_{j=1}^{n_i} \mathbf{X}_{ij}(t)$ for $i = 1, \dots, k$, $t \in \mathcal{I}$. Similarly to the classical multivariate analysis of variance, the matrices $\mathbf{E}$ and $\mathbf{H}$ were used to construct test statistics for H_0. Using a basis function representation of the data, the matrices $\mathbf{E}$ and $\mathbf{H}$ can be expressed in more useful forms, as is presented in the following proposition.

Proposition 2 (Górecki and Smaga (2017))**.** If the components of stochastic processes $\mathbf{X}_{ij}$, $i = 1, \dots, k$, $j = 1, \dots, n_i$, are represented by a finite number of orthonormal basis functions, i.e., the equation (4.3) holds, then $\mathbf{E} = \mathbf{A} - \mathbf{B}$ and $\mathbf{H} = \mathbf{B} - \mathbf{C}$, where

$$\mathbf{A} = \sum_{i=1}^k \sum_{j=1}^{n_i} \boldsymbol{\alpha}_{ij} \boldsymbol{\alpha}_{ij}^\top, \ \mathbf{B} = \sum_{i=1}^k \frac{1}{n_i} \sum_{j=1}^{n_i} \sum_{m=1}^{n_i} \boldsymbol{\alpha}_{ij} \boldsymbol{\alpha}_{im}^\top, \ \mathbf{C} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} \sum_{t=1}^k \sum_{u=1}^{n_t} \boldsymbol{\alpha}_{ij} \boldsymbol{\alpha}_{tu}^\top.$$

By Proposition 2, the matrices $\mathbf{E}$ and $\mathbf{H}$ can be designated only based on the coefficient matrices $\boldsymbol{\alpha}_{ij}$, which is important for the practical implementation of the tests. For testing H_0, we considered the test statistics, which are constructed based on test statistics appearing in MANOVA, namely the Wilks' lambda statistic $\mathrm{W} = \det(\mathbf{E})/\det(\mathbf{E}+\mathbf{H})$, the Lawley-Hotelling trace statistic $\mathrm{LH} = \mathrm{trace}(\mathbf{HE}^{-1})$, the Pillai trace statistic $\mathrm{P} = \mathrm{trace}(\mathbf{H}(\mathbf{H}+\mathbf{E})^{-1})$ and the Roy's maximum root statistic $\mathrm{R} = \lambda_{\max}(\mathbf{HE}^{-1})$. ($\lambda_{\max}(\mathbf{M})$ denotes the maximum eigenvalue of a matrix $\mathbf{M}$.) We proposed the permutation tests based on these test statistics. We observed that any permutation of the stochastic processes $\mathbf{X}_{ij}$ leaves the values of the sums $\mathbf{A}$ and $\mathbf{C}$ unchanged. This observation results in quite fast implementation of the permutation tests.

Górecki and Smaga (2017) also proposed tests based on random projections, the idea for which is the same as that in Section 3.1.2. Let $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{id})^\top$, $\mu_{ij} \in L_2^1(\mathcal{I})$, $i = 1, \ldots, k$, $j = 1, \ldots, d$, and a vector $v_m \in L_2^1(\mathcal{I})$ is chosen randomly using the Gaussian distribution ξ for $m = 1, \ldots, d$. If the null hypothesis H_0 holds, then the null hypothesis

$$H_0^{\mathbf{V}} : \begin{pmatrix} \langle v_1, \mu_{11}\rangle_1 \\ \vdots \\ \langle v_d, \mu_{1d}\rangle_1 \end{pmatrix} = \cdots = \begin{pmatrix} \langle v_1, \mu_{l1}\rangle_1 \\ \vdots \\ \langle v_d, \mu_{ld}\rangle_1 \end{pmatrix}$$

also holds for every $\mathbf{V} = (v_1, \ldots, v_d)^\top \in L_2^d(\mathcal{I})$. Moreover, we have the following theorem.

Theorem 4.1 (Górecki and Smaga (2017)). *Under the same assumptions about ξ as in Section 3.1.2, if $\mu_{ij} \in L_2^1(\mathcal{I})$, $i = 1, \ldots, k$, $j = 1, \ldots, d$, and there exist r_1, r_2, s such that $\mu_{r_1 s} \neq \mu_{r_2 s}$, then $(\xi \times \cdots \times \xi)(\mathcal{A}) = 0$, where $\xi \times \cdots \times \xi$ is a product measure on the space $L_2^d(\mathcal{I})$ and the set $\mathcal{A}$ consists of all vectors $\mathbf{V} = (v_1, \ldots, v_d)^\top \in L_2^d(\mathcal{I})$ for which $H_0^{\mathbf{V}}$ is true.*

Theorem 4.1 shows that if H_0 fails, then for $(\xi \times \cdots \times \xi)$-almost every $\mathbf{V} \in L_2^d(\mathcal{I})$, the null hypothesis $H_0^{\mathbf{V}}$ also fails. Therefore, we concluded that a test for the multivariate analysis of variance for random vectors applied to test $H_0^{\mathbf{V}}$ can be used to test the multivariate analysis of variance problem for functional data. Thus, we can apply an analogous test procedure as in Section 3.1.2. To perform step 3, we can use the classical MANOVA tests, namely the Wilks' lambda test, the Lawley-Hotelling trace test, the Pillai trace test, and Roy's maximum root test, but other tests can also be used.

Recently, Qiu et al. (2021) proposed another tests, but just for the two sample problem ($k = 2$). Their tests are based on aggregating the following pointwise test statistic, similar to Section 3.1.3:

$$T_n(t) = \frac{n_1 n_2}{n}(\bar{\mathbf{X}}_1(t) - \bar{\mathbf{X}}_2(t))^\top (\hat{\mathbf{\Gamma}}(t,t))^{-1}(\bar{\mathbf{X}}_1(t) - \bar{\mathbf{X}}_2(t)),$$

where for $s, t \in \mathcal{I}$

$$\hat{\mathbf{\Gamma}}(s,t) = \frac{1}{n-2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}(\mathbf{X}_{ij}(s) - \bar{\mathbf{X}}_i(s))(\mathbf{X}_{ij}(t) - \bar{\mathbf{X}}_i(t))^\top.$$

To obtain test statistics, the integration and supremum are used:

$$T_n = \int_{\mathcal{I}} T_n(t)dt, \; T_{n,\max} = \sup_{t\in\mathcal{I}} T_n(t).$$

The distribution of T_n is approximated by the similar procedure as in Section 3.1.3, while for $T_{n,\max}$, the nonparametric bootstrap is used.

Simulation results indicate that almost all of the above test procedures control the type I error level and have sensible power. The main exception is the projection test based on Roy's test, which is extremely liberal. Naturally, the other projection tests can have conservative behavior. These are confirmed in the supplement.

Finally, we present the application of the above test procedures to compare both temperature and precipitation in Canada in one analysis. We want to test the null hypothesis (4.1) for vectors of functional data for these two variables. First, we apply the FMANOVA test by Górecki and Smaga (2017) for three groups of Canadian weather stations. For the permutation tests based on a basis function representation, we use the Fourier basis with $K_1 = 7$ and $K_2 = 3$, which were chosen from the set $\{3, 5, \ldots, 31\}$ using BIC. The values of the test statistics are equal to $W = 0.4160014$, $LH = 1.210839$, $P = 0.6642867$, $R = 1.021993$, and the p-values of all tests equal zero. The p-values of the tests based on random projections are also close to zero. Therefore, all tests reject the null hypothesis (4.1) and the significant differences in vectors of mean functions for temperature and precipitation in Canada are detected.

In the end, let us consider the two sample problem for comparing the temperature and precipitation in Eastern and Western Canada. For this problem, the p-values of the tests by Qiu et al. (2021) based on T_n and $T_{n,\max}$ are equal to 0.0009174405 and 0, respectively. Thus, we reject the null hypothesis, which seems to be due to the differences in the precipitation (see Figure 2).

5. CONCLUSIONS

We have presented methodological frameworks, their properties, and practical applications of functional analysis of variance. This analysis has led us to explore various widely adopted approaches in the field of functional data analysis. These approaches are grounded in diverse concepts, such as the aggregation of pointwise statistics, the expansion of functional data using basis functions, the utilization of graphical envelopes, the application of principal components analysis, and the incorporation of random projections. These techniques find utility in a multitude of functional data analysis solutions, extending beyond the realm of functional analysis of variance as discussed in this paper. Furthermore, their versatility allows for their application to a wide range of issues within the field of functional data analysis.

SUPPLEMENT

The supplement presents the simulation studies based on real data examples.

ACKNOWLEDGMENTS

The author is grateful to the editor and the referee for their insightful comments that substantially improved the paper's quality. A part of calculations for the simulation studies was made at the Poznań Supercomputing and Networking Center (grant pl0253-02).

REFERENCES

Acal, C. and Aguilera, A. (2023). Basis expansion approaches for functional analysis of variance with repeated measures. *Advances in Data Analysis and Classification*, 17:291–321.

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

Bouveyron, C., Jacques, J., Schmutz, A., Simões, F., and Bottini, S. (2022). Co-clustering of multivariate functional data for the analysis of air pollution in the south of france. *Annals of Applied Statistics*, 16(3):1400–1422.

Brunner, E., Dette, H., and Munk, A. (1997). Box-type approximations in nonparametric factorial designs. *Journal of the American Statistical Association*, 92:1494–1502.

Chatterjee, A., Mazumder, S., and Das, K. (2023). Functional classwise principal component analysis: a classification framework for functional data analysis. *Data Mining and Knowledge Discovery*, 37:552–594.

Collazos, J., Dias, R., and Zambom, A. (2016). Consistent variable selection for functional regression models. *Journal of Multivariate Analysis*, 146:63–71.

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

Cuesta-Albertos, J., 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. *The Annals of Statistics*, 47:439–467.

Ditzhaus, M. and Gaigall, D. (2022). Testing marginal homogeneity in hilbert spaces with applications to stock market returns. *Test*, 31:749–770.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. Springer.

Ghiglietti, A., Ieva, F., and Paganoni, A. (2017). Statistical inference for stochastic processes: two-sample hypothesis tests. *Journal of Statistical Planning and Inference*, 180:49–68.

Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Di, C., Gellar, J., Harezlak, J., McLean, M., Swihart, B., Xiao, L., Crainiceanu, C., and Reiss, P. (2022). refund: Regression with functional data. R package version 0.1-28. `https://CRAN.R-project.org/package=refund`.

Górecki, T., Krzyśko, M., Waszak, Ł., and Wołyński, W. (2018). Selected statistical methods of data analysis for multivariate functional data. *Statistical Papers*, 59:153–182.

Górecki, T., Krzyśko, M., and Wołyński, W. (2019). Variable selection in multivariate functional data classification. *Statistics in Transition - new series*, 20:123–138.

Górecki, T., Krzyśko, M., and Wołyński, W. (2020). Independence test and canonical correlation analysis based on the alignment between kernel matrices for multivariate functional data. *Artificial Intelligence Review*, 53:475–499.

Górecki, T. and Smaga, Ł. (2015). A comparison of tests for the one-way ANOVA problem for functional data. *Computational Statistics*, 30:987–1010.

Górecki, T. and Smaga, Ł. (2017). Multivariate analysis of variance for functional data. *Journal of Applied Statistics*, 44:2172–2189.

Górecki, T. and Smaga, Ł. (2019). fdANOVA: An R software package for analysis of variance for univariate and multivariate functional data. *Computational Statistics*, 34:571–597.

Guo, J., Zhou, B., and Zhang, J. T. (2018). A supremum-norm based test for the equality of several covariance functions. *Computational Statistics & Data Analysis*, 124:15–26.

Guo, J., Zhou, B., and Zhang, J. T. (2019). New tests for equality of several covariance functions for functional data. *Journal of the American Statistical Association*, 114:1251–1263.

Hall, P. and Van Keilegom, I. (2007). Two-sample tests in functional data analysis starting from discrete data. *Statistica Sinica*, 17:1511–1531.

Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*, 113:649–659.

Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer.

Horváth, L., Kokoszka, P., and Reeder, R. (2013). Estimation of the mean of functional time series and a two-sample problem. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 75(1):103–122.

Horváth, L., Kokoszka, P., and Reimherr, M. (2009). Two sample inference in functional linear models. *Canadian Journal of Statistics*, 37:571–591.

Horváth, L., Kokoszka, P., and Rice, G. (2014). Testing stationarity of functional time series. *Journal of Econometrics*, 179(1):66–82.

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

Jiang, Q., Hušková, M., Meintanis, S., and Zhu, L. (2019). Asymptotics, finite-sample comparisons and applications for two-sample tests with functional data. *Journal of Multivariate Analysis*, 170:202–220.

Karuppusami, R., Antonisamy, B., and Premkumar, P. (2022). Functional principal component analysis for identifying the child growth pattern using longitudinal birth cohort data. *BMC Medical Research Methodology*, 22:76.

Kraus, D. (2019). Inferential procedures for partially observed functional data. *Journal of Multivariate Analysis*, 173:583–603.

Krzyśko, M., Nijkamp, P., Ratajczak, W., and Wołyński, W. (2022). Multidimensional economic indicators and multivariate functional principal components analysis (MFPCA) in a comparative study of countries' competitiveness. *Journal of Geographical Systems*, 24:49–65.

Krzyśko, M. and Waszak, Ł. (2013). Canonical correlation analysis for functional data. *Biometrical Letters*, 50:95–105.

Kuryło, K. and Smaga, Ł. (2024). Functional repeated measures analysis of variance and its application. *Statistics in Transition new series*, 25:185–204.

Martínez-Camblor, P. and Corral, N. (2011). Repeated measures analysis for functional data. *Computational Statistics & Data Analysis*, 55:3244–3256.

Meléndez, R., Giraldo, R., and Leiva, V. (2021). Sign, Wilcoxon and Mann-Whitney tests for functional data: An approach based on random projections. *Mathematics*, 9:44.

Mrkvička, T., Myllymäki, M., and Hahn, U. (2017). Multiple monte carlo testing, with applications in spatial point processes. *Statistics and Computing*, 27:1239–1255.

Mrkvička, T., Myllymäki, M., Jilek, M., and Hahn, U. (2020). A one-way ANOVA test for functional data with graphical interpretation. *Kybernetika*, 56:432–458.

Mrkvička, T., Roskovec, T., and Rost, M. (2021). A nonparametric graphical tests of significance in functional GLM. *Methodology and Computing in Applied Probability*, 23:593–612.

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., and Hahn, U. (2017). Global envelope tests for spatial processes. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 79:381–404.

Navarro-Esteban, P. and Cuesta-Albertos, J. (2021). High-dimensional outlier detection using random projections. *Test*, 30:908–934.

Nie, Y., Wang, L., Liu, B., and Cao, J. (2018). Supervised functional principal component analysis. *Statistics and Computing*, 28:713–723.

Paparoditis, E. and Sapatinas, T. (2016). Bootstrap-based testing of equality of mean functions or equality of covariance operators for functional data. *Biometrika*, 103:727–733.

Pauly, M., Brunner, E., and Konietschke, F. (2015). Asymptotic permutation tests in general factorial designs. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 77:461–473.

Qiu, Z., Chen, J., and Zhang, J.-T. (2021). Two-sample tests for multivariate functional data with applications. *Computational Statistics & Data Analysis*, 157:107160.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramsay, J. (2023). fda: Functional data analysis. R package version 6.1.4. `https://CRAN.R-project.org/package=fda`.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis, Second Edition*. Springer.

Smaga, Ł. (2019a). Projection-based repeated measures analysis for functional data. In Steland, A., Rafajłowicz, E., and Okhrin, O., editors, *Stochastic Models, Statistics and Their Applications*, volume 294 of *Springer Proceedings in Mathematics and Statistics*, pages 229–237.

Smaga, Ł. (2019b). Repeated measures analysis for functional data using Box-type approximation - with applications. *REVSTAT - Statistical Journal*, 17:523–549.

Smaga, Ł. (2020). A note on repeated measures analysis for functional data. *AStA Advances in Statistical Analysis*, 104:117–139.

Smaga, Ł. (2021). General linear hypothesis testing in functional response model. *Communications in Statistics - Theory and Methods*, 50:5068–5083.

Smaga, Ł. and Matsui, H. (2018). A note on variable selection in functional regression via random subspace method. *Statistical Methods and Applications*, 27:455–477.

Smaga, Ł. and Zhang, J.-T. (2019). Linear hypothesis testing with functional data. *Technometrics*, 61:99–110.

Smaga, Ł. and Zhang, J.-T. (2020). Linear hypothesis testing for weighted functional data with applications. *Scandinavian Journal of Statistics*, 47:493–515.

Tokushige, S., Yadohisa, H., and Inada, K. (2007). Crisp and fuzzy k-means clustering algorithms for multivariate functional data. *Computational Statistics*, 22:1–16.

Ullah, S. and Finch, C. (2013). Applications of functional data analysis: A systematic review. *BMC Medical Research Methodology*, 13:43.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., and Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92:381–397.

Wu, R., Wang, B., and Xu, A. (2022). Functional data clustering using principal curve methods. *Communications in Statistics - Theory and Methods*, 51:7264–7283.

Zhang, C., Peng, H., and Zhang, J. (2010a). Two samples tests for functional data. *Communications in Statistics - Theory and Methods*, 39:559–578.

Zhang, J. T. (2013). *Analysis of Variance for Functional Data*. Chapman & Hall, London.

Zhang, J.-T., Cheng, M.-Y., Wu, H.-T., and Zhou, B. (2019). A new test for functional one-way ANOVA with applications to ischemic heart screening. *Computational Statistics & Data Analysis*, 132:3–17.

Zhang, J. T. and Liang, X. (2013). One-way ANOVA for functional data via globalizing the pointwise F-test. *Scandinavian journal of Statistics*, 41:51–71.

Zhang, J.-T., Liang, X., and Xiao, S. (2010b). On the two-sample Behrens-Fisher problem for functional data. *Journal of Statistical Theory and Practice*, 4:571–587.