Statistics and Its Interface Volume 12 (2019) 647–658

A multivariate coefficient of variation for functional data

Mirosław Krzyśko and Łukasz Smaga*

This paper considers an adaptation of the multivariate coefficient of variation to functional data. Similarly to the coefficient of variation and its multivariate generalizations, the functional multivariate coefficient of variation (FMCV) is useful in practical applications. Namely, it may be helpful for comparing the relative variation in different populations or the performance of different equipment characterized by univariate or multivariate functional data. Some theoretical properties of the new functional data analysis method are discussed. Using the basis function representation of the data, it is shown that the FMCV reduces to the multivariate coefficient of variation of a vector of coefficients of that representation. This enables effective computation of the FMCV. The performance of classical and robust estimators of the FMCV is compared in a finite sample setting using simulation studies. The new methods are illustrated on electrocardiography (ECG) data. These data are divided into two groups: normal and abnormal (representative of some cardiac pathology). The variability in the abnormal group is shown to be significantly greater than that in the normal group.

AMS 2000 subject classifications: Primary 62H05; secondary 62M99.
Keywords and phrases: Dispersion measure, Functional data analysis, Multivariate coefficient of variation, Robust estimation, Variability measure.

1. INTRODUCTION

The coefficient of variation (CV), being the ratio of the standard deviation to the population mean, is a widely used measure of relative variation. The CV is a dimensionless quantity and may be expressed as a percentage. It is commonly used to compare the variability of several populations, even when they are characterized by variables expressed in different units or have significantly different means. In particular, the CV is often used to assess the performance or reproducibility of measurement techniques or equipment. Of course, the lower the CV, the greater the precision of the technique or equipment.

In many experiments, the populations are characterized by more than one variable. In such cases, computing the CV for each variable is a common practice, although this ignores the correlation between them, and does not summarize the variability of the multivariate data into a single index. The known multivariate extensions of the CV have received less consideration in the literature. This may be because generalizing the univariate CV to the multivariate setting is not straightforward. Constructions of the multivariate CV are based on different approaches, which reduce to the CV in the univariate case. However, when the number of variables is greater than one, the multivariate CVs are not generally equal to each other and hence do not measure the same quantity. Thus there is no one universal definition of multivariate CV. Different constructions of the multivariate CV are briefly reviewed in the next paragraph.

Let $\mathbf{X} = (X_1, \ldots, X_p)^\top$ be a p-dimensional random vector with mean vector $\mathbf{u} \neq \mathbf{0}_p$ and covariance matrix $\mathbf{\Sigma}$. The following definitions of the multivariate coefficient of variation (MCV) were introduced in [31, 40, 41, 2], respectively:

$$\mathrm{MCV_R} = \sqrt{\frac{(\det \mathbf{\Sigma})^{1/p}}{\mathbf{u}^\top \mathbf{u}}},\ \mathrm{MCV_{VV}} = \sqrt{\frac{\mathrm{tr}\mathbf{\Sigma}}{\mathbf{u}^\top \mathbf{u}}},$$
$$\mathrm{MCV_{VN}} = \sqrt{\frac{1}{\mathbf{u}^\top \mathbf{\Sigma}^{-1} \mathbf{u}}},\ \mathrm{MCV_{AZ}} = \sqrt{\frac{\mathbf{u}^\top \mathbf{\Sigma} \mathbf{u}}{(\mathbf{u}^\top \mathbf{u})^2}}.$$

$\mathrm{MCV_R}$ and $\mathrm{MCV_{VV}}$ are based on the generalized variance $\det \mathbf{\Sigma}$ and the total variance $\mathrm{tr}\mathbf{\Sigma}$ respectively. In the case of $\mathrm{MCV_{VN}}$, the Mahalanobis distance $\mathbf{u}^\top \mathbf{\Sigma}^{-1} \mathbf{u}$ appears to be a natural extension of the CV. Voinov and Nikulin [41] made this specific allusion to the MCV, since they claimed that this measure for variation between the mean vector $\mathbf{u}$ and a covariance matrix $\mathbf{\Sigma}$ increases in the sense of nonnegative definiteness as $\mathbf{u}^\top \mathbf{u}$ increases. Finally, $\mathrm{MCV_{AZ}}$ is derived based on a matrix generalizing the square of the CV (see also Section 2 for another interpretation of $\mathrm{MCV_{AZ}}$). For a more detailed review of the MCVs we refer to [1, 2], where many of their properties are stated.

In this paper, we extend the definition of the MCV to the functional data framework. Functional data analysis (FDA) is a branch of statistics concerned with (potentially multidimensional) functions, curves or surfaces, and has received much attention in the literature. In practice, each observation consists of pairs (t_j, x_j), $j = 1, \ldots, m$, where $t_1, \ldots, t_m$ are time or location points, and $x_1, \ldots, x_m$ are discrete measurements being observations of a (smooth) function X, i.e., $x_1 = X(t_1), \ldots, x_m = X(t_m)$. Such data appear in many

*Corresponding author.

scientific fields, for instance chemometrics, economics, engineering, genetics, medicine, meteorology and plant science. Particular examples include data on activity monitoring with accelerometers, electrical activity of the heart, GDP per capita, growth curves, and temperature and precipitation in a given location. In contrast to traditional multivariate analysis, the statistical methods of FDA are developed for the functions as such rather than the individual measurements. A broad perspective on FDA methods is presented in the monographs of [9, 18, 27, 28, 42], as well as in the review papers of [36, 39]. Since the aims of FDA are mainly the same as for other statistical analyses, there are functional versions of many standard statistical methods, including analysis of variance [15, 16], canonical correlation analysis [20], classification [7, 23, 32], cluster analysis [5, 10], outlier detection [3], principal component analysis [4, 8], regression analysis [12, 21, 25] and repeated measures analysis [24, 35].

The techniques of FDA help to avoid many problems of classical multivariate analysis. First of all, FDA avoids the curse of dimensionality. When the number of time points exceeds the number of time series, most statistical methods do not give satisfactory results due to overparameterization. To avoid this problem, dimension reduction techniques, such as principal component analysis, are commonly used. However, in this case, some information about the spatial and temporal structure of the data may be lost. In the case of functional data, this problem can be avoided, because the time series are replaced with a set of curves independent of the number of time points. Secondly, FDA methods easily deal with the problem of missing data. Unfortunately, most of classical data analysis methods require complete time series. One solution is simply to delete a time series that has missing values from the data set, but this usually leads to information loss. Another possibility is to use one of many methods of predicting missing data, but then the results will depend on the method used. In contrast, in the case of functional data, the problem of missing observations is solved by expressing time series in the form of a set of curves. Moreover, the time points do not have to be evenly distributed in individual time series.

Applications of the functional multivariate coefficient of variation may be similar to those of the MCV, as for example comparing the variability in different populations or the performance of different equipment. However, the functional multivariate coefficient of variation is applicable, when the experimental units are characterized by multivariate functional data, for which the MCV can not be used, since it is designed for classical multivariate data only.

It appears not to be easy to generalize the MCVs to functional data. It is difficult to find functional counterparts of the elements appearing in the definitions of $\mathrm{MCV_R}$ and $\mathrm{MCV_{VN}}$. Moreover, it is easy to see that they require the non-singularity of $\mathbf{\Sigma}$, which may limit their applicability. On the other hand, $\mathrm{MCV_{VV}}$ does not take into account the correlation structure of the data, which is important in functional data analysis. Fortunately, $\mathrm{MCV_{AZ}}$ is free of these restrictions, and it can be quite easily extended to functional data by using its alternative form as given in [2], as we shall see in Section 2. We refer to $\mathrm{MCV_{AZ}}$ as the multivariate coefficient of variation of Albert-Zhang type. We show some theoretical properties of the functional version of $\mathrm{MCV_{AZ}}$, and we present it in a simple form using the basis function representation of the functional observations. Classical and robust estimators of the $\mathrm{MCV_{AZ}}$ for functional data are also proposed, and their performance is investigated by simulation studies (Section 3), which indicate the advantage of using robust estimates in the presence of outlying observations or under non-normal data. In Section 4, we illustrate the new methods on an electrocardiography (ECG) data set. Some conclusions are given in Section 5.

2. FUNCTIONAL MULTIVARIATE COEFFICIENT OF VARIATION

In this section we define the multivariate coefficient of variation for functional data by adapting the multivariate coefficient of variation of Albert-Zhang type ($\mathrm{MCV_{AZ}}$). We also consider classical and robust estimation methods.

Let us first note the following alternative form of the $\mathrm{MCV_{AZ}}$, which is established in Property 3 in the Appendix of [2]:

$$\mathrm{MCV_{AZ}} = \frac{\sqrt{\mathrm{Var}(\mathbf{u}_*^\top \mathbf{X})}}{\|\mathbf{u}\|}, \tag{1}$$

where $\mathbf{u}_* = \mathbf{u}/\|\mathbf{u}\|$. The form in Eq. (1) makes the definition of the $\mathrm{MCV_{AZ}}$ more explicit and similar to that of the univariate coefficient of variation: namely, it is given as the ratio of a standard deviation to a mean value. Thus, $\mathrm{MCV_{AZ}}$ is the univariate coefficient of variation for the random variable $\mathbf{u}_*^\top \mathbf{X}$. In the following, we use the expression in Eq. (1) for $\mathrm{MCV_{AZ}}$ to define the coefficient of variation for multivariate functional data.

Let $\mathbf{X}(t) = (X_1(t), \ldots, X_p(t))^\top$, $t \in [a,b]$, $a, b \in \mathbb{R}$ be a p-dimensional random process with mean function $\boldsymbol{\mu}(t) = (\mu_1(t), \ldots, \mu_p(t))^\top \neq \mathbf{0}_p$. We assume that the process $\mathbf{X}(t)$, $t \in [a,b]$ belongs to the Hilbert space $L_2^p[a,b]$ of p-dimensional vectors of square integrable functions on $[a,b]$. Let $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the inner product and the norm in the space $L_2^p[a,b]$.

Definition 2.1. *The functional multivariate coefficient of variation (FMCV) for the random process* $\mathbf{X}(t)$, $t \in [a,b]$ *is defined as follows:*

$$\mathrm{FMCV} = \frac{\sqrt{\mathrm{Var}(\langle \boldsymbol{\mu}_*, \mathbf{X} \rangle)}}{\|\boldsymbol{\mu}\|}, \tag{2}$$

where $\boldsymbol{\mu}_*(t) = \boldsymbol{\mu}(t)/\|\boldsymbol{\mu}\|$, $t \in [a,b]$.

The above definition is analogous to the (classical) multivariate coefficient of variation of Albert-Zhang type given in the form of Eq. (1). Namely, the inner product $\mathbf{u}_*^\top \mathbf{X}$ in Euclidean space $\mathbb{R}^p$ is naturally replaced by the inner product $\langle \boldsymbol{\mu}_*, \mathbf{X} \rangle$ in the Hilbert space $L_2^p[a,b]$, and similarly, the norm $\|\mathbf{u}\|$ is replaced by $\|\boldsymbol{\mu}\|$. Of course, the FMCV is well-defined if $\mathrm{Var}(\langle \boldsymbol{\mu}_*, \mathbf{X} \rangle)$ exists. A condition guaranteeing its existence is the square integrability of the components of the process $\mathbf{X}(t)$, $t \in [a,b]$, as is stated in the following result. Moreover, this theorem states that the FMCV is the CV of the random variable $\langle \boldsymbol{\mu}_*, \mathbf{X} \rangle$, similarly to the multivariate coefficient of variation of Albert-Zhang type. The proof is outlined in Appendix A.

Theorem 2.1. *If $X_i(t)$, $t \in [a,b]$, $i = 1, \ldots, p$, are square integrable, i.e., $\mathrm{E}\|X_i\|^2 = \mathrm{E}\int_a^b X_i^2(t)\, dt < \infty$, then $\mathrm{Var}(\langle \boldsymbol{\mu}_*, \mathbf{X} \rangle)$ exists. Furthermore, the FMCV defined in Eq. (2) is the CV of the random variable $\langle \boldsymbol{\mu}_*, \mathbf{X} \rangle$.*

Now, we derive a simpler form of the FMCV defined in Eq. (2) using the basis function representation of the process $\mathbf{X}(t)$, $t \in [a,b]$. Let $\mathbf{X}(t)$ belong to a finite dimensional subspace $\mathcal{L}_2^p[a,b]$ of $L_2^p[a,b]$, where the components of $\mathbf{X}(t)$ can be represented by a finite number of basis functions (see, for example [28, 36]), i.e.,

$$X_k(t) = \sum_{l=1}^{B_k} \alpha_{kl} \varphi_{kl}(t), \tag{3}$$

where $k = 1, \ldots, p$, $t \in [a,b]$, $B_k \in \mathbb{N}$, α_{kl} are random variables with finite variance and $\{\varphi_{kl}\}_{l=1}^\infty$, $k = 1, \ldots, p$ are bases in the space $L_2^1[a,b]$. The Eq. (3) can be expressed in the following matrix notation:

$$\mathbf{X}(t) = \boldsymbol{\Phi}(t)\boldsymbol{\alpha}, \tag{4}$$

where

$$\boldsymbol{\Phi}(t) = \mathrm{diag}\left(\boldsymbol{\varphi}_1^\top(t), \ldots, \boldsymbol{\varphi}_p^\top(t)\right)$$

is the block diagonal matrix of $\boldsymbol{\varphi}_k^\top(t) = (\varphi_{k1}(t), \ldots, \varphi_{kB_k}(t))$, $k = 1, \ldots, p$ and $\boldsymbol{\alpha} = (\alpha_{11}, \ldots, \alpha_{1B_1}, \ldots, \alpha_{p1}, \ldots, \alpha_{pB_p})^\top$. By Eq. (4), for $t \in [a,b]$, we have

$$\boldsymbol{\mu}(t) = \mathrm{E}(\mathbf{X}(t)) = \boldsymbol{\Phi}(t)\mathrm{E}(\boldsymbol{\alpha}) = \boldsymbol{\Phi}(t)\mathbf{a}. \tag{5}$$

From Eq. (4) and Eq. (5), it follows that

$$\begin{aligned} \langle \boldsymbol{\mu}_*, \mathbf{X} \rangle &= \int_a^b \boldsymbol{\mu}_*(t)^\top \mathbf{X}(t)\, dt \\ &= \int_a^b \frac{\mathbf{a}^\top}{\|\boldsymbol{\mu}\|} \boldsymbol{\Phi}(t)^\top \boldsymbol{\Phi}(t)\boldsymbol{\alpha}\, dt = \frac{\mathbf{a}^\top \mathbf{J}_{\boldsymbol{\Phi}} \boldsymbol{\alpha}}{\|\boldsymbol{\mu}\|}, \end{aligned} \tag{6}$$

$$\begin{aligned} \|\boldsymbol{\mu}\| &= \sqrt{\int_a^b \boldsymbol{\mu}(t)^\top \boldsymbol{\mu}(t)\, dt} \\ &= \sqrt{\int_a^b \mathbf{a}^\top \boldsymbol{\Phi}(t)^\top \boldsymbol{\Phi}(t)\mathbf{a}\, dt} \\ &= \sqrt{\mathbf{a}^\top \mathbf{J}_{\boldsymbol{\Phi}} \mathbf{a}} = \|\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\mathbf{a}\|, \end{aligned} \tag{7}$$

if the matrix $\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}$ exists, where $\mathbf{J}_{\boldsymbol{\Phi}} = \mathrm{diag}(\mathbf{J}_{\boldsymbol{\varphi}_1}, \ldots, \mathbf{J}_{\boldsymbol{\varphi}_p})$ and $\mathbf{J}_{\boldsymbol{\varphi}_k} = \int_a^b \boldsymbol{\varphi}_k(t)\boldsymbol{\varphi}_k^\top(t)\, dt$ is the $B_k \times B_k$ cross product matrix corresponding to the basis $\{\varphi_{kl}\}_{l=1}^\infty$, $k = 1, \ldots, p$. For an orthonormal basis, for instance the Fourier basis, the cross product matrix is equal to the identity matrix. A formula for $\mathbf{J}_{\boldsymbol{\Phi}}$ for a B-spline basis is given, for example, in [19]. The approximation of the cross product matrix for these as well as other bases is also implemented in the function inprod() in the R package fda [29, 30].

For random vector $\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\boldsymbol{\alpha}$, we have $\mathrm{E}(\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\boldsymbol{\alpha}) = \mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\mathbf{a}$ and $\mathrm{Cov}(\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\boldsymbol{\alpha}) = \mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}$, where $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}} = \mathrm{Cov}(\boldsymbol{\alpha})$. By Eq. (2), (6) and (7), we obtain

$$\mathrm{FMCV} = \frac{\sqrt{\mathrm{Var}\left(\frac{\mathbf{a}^\top \mathbf{J}_{\boldsymbol{\Phi}} \boldsymbol{\alpha}}{\|\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\mathbf{a}\|}\right)}}{\|\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\mathbf{a}\|} = \sqrt{\frac{\mathbf{a}^\top \mathbf{J}_{\boldsymbol{\Phi}} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}} \mathbf{J}_{\boldsymbol{\Phi}} \mathbf{a}}{(\mathbf{a}^\top \mathbf{J}_{\boldsymbol{\Phi}} \mathbf{a})^2}}. \tag{8}$$

Thus, we have proved the following result.

Theorem 2.2. *Under the above assumptions and notation, the functional multivariate coefficient of variation for the random process $\mathbf{X}(t)$, $t \in [a,b]$, defined in Eq. (2), is the multivariate coefficient of variation of Albert-Zhang type for the random vector $\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\boldsymbol{\alpha}$, if the matrix $\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}$ exists.*

By Theorem 2.2, the FMCV reduces to the multivariate coefficient of variation of Albert-Zhang type for the $(B_1 + \cdots + B_p)$-dimensional random vector $\mathbf{J}_{\boldsymbol{\Phi}}^{1/2}\boldsymbol{\alpha}$. Although the FMCV is defined for univariate and multivariate functional data, we note that even when $p = 1$, the FMCV reduces to $\mathrm{MCV}_{\mathrm{AZ}}$ (not to the CV), since B_1 is usually greater than one.

The above definitions and assumptions are population-based. In practice, given a random functional sample, we have to estimate the unknown vector $\boldsymbol{\alpha}$ in Eq. (4), as well as its parameters $\mathbf{a}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$, appearing in the expression for the FMCV given in Eq. (8).

Let $\mathbf{x}_1(t), \ldots, \mathbf{x}_n(t)$, $t \in [a,b]$ be a random sample containing realizations of the process $\mathbf{X}(t)$. These observations are represented similarly as in Eq. (4), i.e.,

$$\mathbf{x}_i(t) = \boldsymbol{\Phi}(t)\boldsymbol{\alpha}_i,$$

where $t \in [a,b]$ and $i = 1, \ldots, n$. Then, the vectors $\boldsymbol{\alpha}_i$, $i = 1, \ldots, n$ can be estimated by the least squares method or the roughness penalty approach (see, for example, [28]). The expansion lengths B_k in Eq. (3) can be selected deterministically or with the use of information criteria such as the Akaike and Bayesian information criteria. Shmueli [34]

showed that the Akaike criterion gives the best prediction, while the Bayesian criterion gives the best fit.

Using the estimators of $\boldsymbol{\alpha}_i$, say $\hat{\boldsymbol{\alpha}}_i$, $i = 1, \ldots, n$, we can estimate the mean vector $\mathbf{a}$ and the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$. The classical estimators are the sample mean and the sample covariance matrix, i.e.,

$$\hat{\mathbf{a}} = \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{\alpha}}_i, \ \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\alpha}} = \frac{1}{n}\sum_{i=1}^{n} (\hat{\boldsymbol{\alpha}}_i - \hat{\mathbf{a}})(\hat{\boldsymbol{\alpha}}_i - \hat{\mathbf{a}})^\top. \tag{9}$$

However, these estimators may break down when the data contain outliers. Thus, many authors recommend the use of robust estimators of location and scatter in the presence of outlying observations [1, 2, 37, 38]. There are many robust estimators worthy of consideration. Similarly to [1], we shall refer mainly to the two most commonly used, namely the minimum covariance determinant (MCD) estimator [33] and the S-estimator [6].

For a given breakdown point α, the MCD estimator is based on a subset of $\{\hat{\boldsymbol{\alpha}}_1, \ldots, \hat{\boldsymbol{\alpha}}_n\}$ of size $h = \lfloor n(1-\alpha) \rfloor$ minimizing the generalized variance (i.e., the determinant of the covariance matrix) among all possible subsets of size h. Then the MCD estimators of $\mathbf{a}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$ are the sample mean and the sample covariance matrix (multiplied by a consistency factor) computed from this subset. The location and scatter S-estimators are the vector $\mathbf{a}_n$ and the positive definite symmetric matrix $\boldsymbol{\Sigma}_n$ respectively which minimizes $\det(\boldsymbol{\Sigma}_n)$ subject to

$$\frac{1}{n}\sum_{i=1}^{n} \rho\left(\sqrt{(\hat{\boldsymbol{\alpha}}_i - \mathbf{a}_n)^\top \boldsymbol{\Sigma}_n^{-1} (\hat{\boldsymbol{\alpha}}_i - \mathbf{a}_n)}\right) = b_0,$$

where $\rho : \mathbb{R} \to [0, \infty)$ is a given non-decreasing and symmetric function (e.g., Tukey's biweight) and b_0 is a constant needed to ensure consistency of the estimator.

Finally, the (classical and robust) estimators of the FMCV are obtained by replacing the parameters $\mathbf{a}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$ in Eq. (8) by their estimators $\hat{\mathbf{a}}$ and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\alpha}}$. In the next section, we test their behavior on finite samples by simulation experiments.

3. SIMULATION STUDIES

In this section, simulation experiments are conducted to measure the finite sample performance of the estimators of the FMCV given by Eq. (8). We consider a classical estimator based on the sample mean and the sample covariance matrix given by Eq. (9), as well as robust estimators based on the MCD and S estimators of the parameters $\mathbf{a}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$.

3.1 Simulation design

We consider the functional sample $\mathbf{x}_1(t), \ldots, \mathbf{x}_n(t)$ of size $n = 100, 200, 300$ containing realizations of the random process $\mathbf{X}(t)$, $t \in [0, 1]$ of dimension $p = 5$. These observations are generated in the following discretized way:

$$\mathbf{x}_i(t_j) = \boldsymbol{\Phi}(t_j)\boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_{ij},$$

where $i = 1, \ldots, n$, t_j, $j = 1, \ldots, 50$ are equally spaced design time points in $[0, 1]$, the matrix $\boldsymbol{\Phi}(t)$ is as in Section 2 with $B_k = 5$, $k = 1, \ldots, p$, $\boldsymbol{\alpha}_i$ are $5p$-dimensional random vectors, and $\boldsymbol{\epsilon}_{ij} = (\epsilon_{ij1}, \ldots, \epsilon_{ijp})^\top$ are measurement errors such that $\epsilon_{ijk} \sim N(0, 0.025 r_{ik})$ and r_{ik} is the range of the k-th row of the matrix

$$(\boldsymbol{\Phi}(t_1)\boldsymbol{\alpha}_i \ldots \boldsymbol{\Phi}(t_{50})\boldsymbol{\alpha}_i),$$

$k = 1, \ldots, p$. For data generation and for evaluation of the estimators, we consider two commonly used bases, namely the Fourier and B-spline bases.

The vectors $\boldsymbol{\alpha}_i$, $i = 1, \ldots, n$ were generated from a multivariate normal distribution or multivariate t-distribution with five degrees of freedom, with mean vector $\mathbf{a}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$, or a distribution of $\mathbf{Z}\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{1/2} + \mathbf{a}$, where $\mathbf{Z} = (\mathbf{Y} - \mathrm{E}(\mathbf{Y}))\mathrm{Cov}(\mathbf{Y})^{-1/2}$ and $\mathbf{Y}$ follows a mixture of two independent multivariate normal distributions $N_{5p}(\mathbf{1}_{5p}, \mathbf{I}_{5p})$ (with probability 0.3) and $N_{5p}(2\mathbf{1}_{5p}, 2\mathbf{I}_{5p})$ (with probability 0.7). Here $\mathbf{1}_{5p} = (1, \ldots, 1)^\top$ and $\mathbf{I}_{5p}$ is the identity matrix. Similarly to [1], we set $\mathbf{a} = \mathbf{a}_1 := a\mathbf{e}_1$ or $\mathbf{a} = \mathbf{a}_2 := (a/(5p)^{1/2})\mathbf{1}_{5p}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}} = (1-\rho)\mathbf{I}_{5p} + \rho\mathbf{1}_{5p}\mathbf{1}_{5p}^\top$, where a is chosen to obtain a given value of the FMCV, $\rho = 0, 0.5, 0.8$ and $\mathbf{e}_1 = (1, 0, \ldots, 0)^\top$. Note that $\mathbf{a}_2$ is an eigenvector of $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$. We set FMCV $= 0.1, 0.5, 0.9$. Moreover, to obtain uncontaminated and contaminated functional data, $\varepsilon\%$ of the observations are generated with the covariance matrix equal to $10\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$, where $\varepsilon = 0, 10, 20, 30, 40, 50$. The fact that this method of data generation results in contaminated functional data (if $\varepsilon > 0$) was confirmed by an outliergram, that is, by the functional outlier detection method of [3] detecting outlying observations by connecting two functional depths. The R code of this method is available at http://halweb.uc3m.es/esp/Personal/personas/aarribas/esp/public.html. Figure 1 on page 651 shows sample realizations of simulated functional data.

For each combination of the above simulation parameters, 1000 samples were generated, and the classical, MCD and S estimators (with breakdown point equal to 0.5) of the FMCV were evaluated on these samples. Based on the values of the estimators, the estimated mean squared error (MSE) was computed to assess the performance of the estimators, namely

$$\mathrm{MSE} = \frac{1}{1000}\sum_{i=1}^{1000} (\widehat{\mathrm{FMCV}}_i - \mathrm{FMCV})^2,$$

where $\widehat{\mathrm{FMCV}}_i$ is the value of the estimator of the FMCV obtained in the i-th sample, $i = 1, \ldots, 1000$. The resulting MSEs are given in Tables 1–3 and Tables 1–7 in the Supplementary Materials.

The simulation experiments as well as the real data example in Section 4 were conducted in the R computing environment [30]. The R code reproducing the simulation results, the results of the real data example of Section 4, etc., are given in the Supplementary Materials.

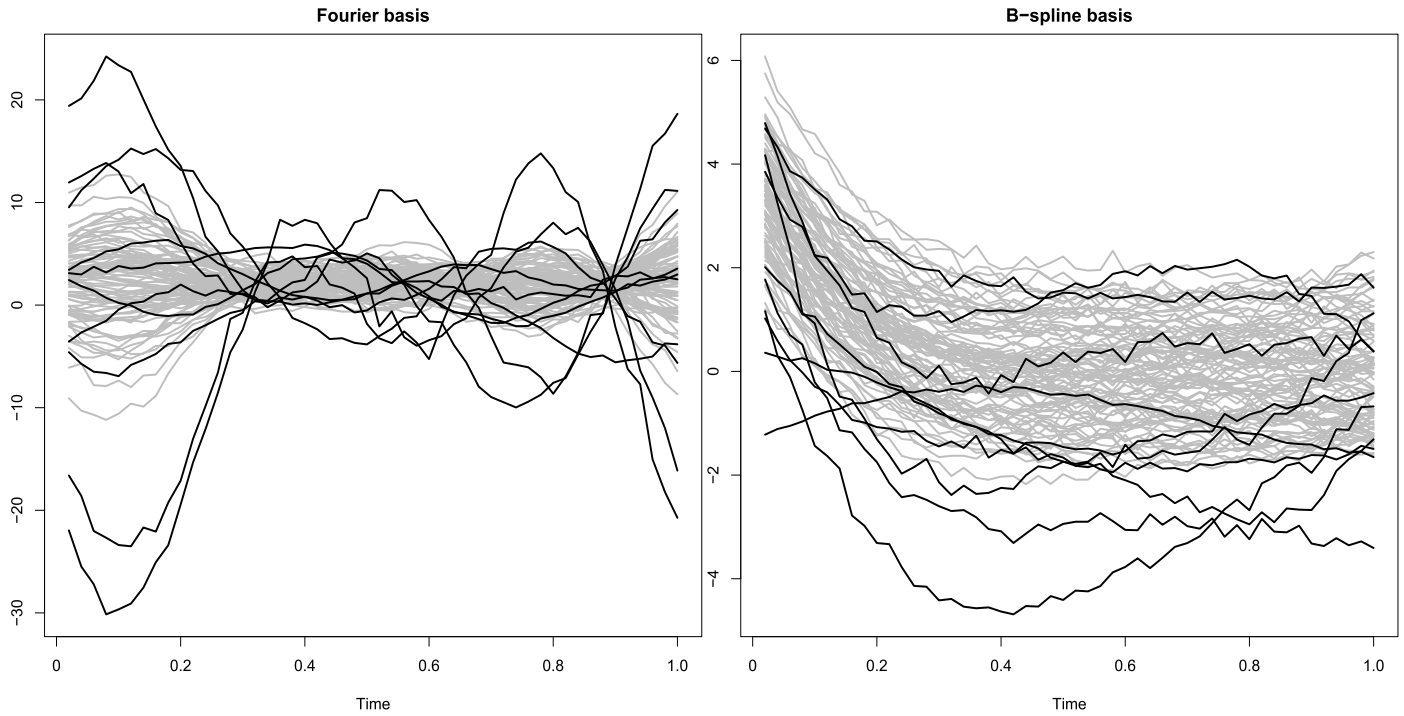

Figure 1. Sample realizations of the first functional variables of simulated data with 10% outlying observations. Uncontaminated and contaminated data are depicted in gray and black respectively.

Table 1. MSEs of classical and robust (MCD and S) estimators of the FMCV in the case of normal distribution, Fourier basis, $n = 100$, $p = 5$ *and* $\mathbf{a} = \mathbf{a}_1$.

	class.	MCD	S	class.	MCD	S	class.	MCD	S
ε	FMCV $= 0.1, \rho = 0$			FMCV $= 0.1, \rho = 0.5$			FMCV $= 0.1, \rho = 0.8$		
0	0.0000	0.0001	0.0001	0.0001	0.0002	0.0001	0.0003	0.0005	0.0003
10	0.0016	0.0001	0.0001	0.0019	0.0002	0.0002	0.0025	0.0005	0.0004
20	0.0047	0.0001	0.0004	0.0053	0.0002	0.0006	0.0068	0.0006	0.0009
30	0.0085	0.0002	0.0017	0.0095	0.0003	0.0019	0.0118	0.0007	0.0025
40	0.0130	0.0009	0.0064	0.0147	0.0011	0.0069	0.0181	0.0016	0.0080
50	0.0179	0.0059	0.0133	0.0204	0.0065	0.0140	0.0251	0.0088	0.0161
ε	FMCV $= 0.5, \rho = 0$			FMCV $= 0.5, \rho = 0.5$			FMCV $= 0.5, \rho = 0.8$		
0	0.0019	0.0029	0.0019	0.0209	0.0314	0.0212	0.0656	0.0953	0.0664
10	0.0332	0.0031	0.0025	0.1278	0.0347	0.0304	0.2922	0.1007	0.0912
20	0.0940	0.0032	0.0084	0.2919	0.0368	0.0519	0.6255	0.1054	0.1340
30	0.1630	0.0036	0.0331	0.4902	0.0412	0.1177	1.0005	0.1209	0.2661
40	0.2345	0.0168	0.1345	0.7307	0.0936	0.3261	1.4984	0.2156	0.6466
50	0.3033	0.1191	0.2693	0.9976	0.3683	0.6583	1.9943	0.8076	1.2419
ε	FMCV $= 0.9, \rho = 0$			FMCV $= 0.9, \rho = 0.5$			FMCV $= 0.9, \rho = 0.8$		
0	0.0131	0.0163	0.0130	0.1197	0.1654	0.1216	0.3854	0.5333	0.3876
10	0.0706	0.0154	0.0095	0.6283	0.1763	0.1662	1.5364	0.5636	0.5306
20	0.1838	0.0160	0.0171	1.1743	0.1909	0.2820	2.7146	0.5771	0.7428
30	0.2867	0.0136	0.0640	1.7589	0.2138	0.5742	4.2450	0.6646	1.4568
40	0.3755	0.0335	0.2904	2.3889	0.4412	1.4175	5.8112	1.1368	3.2821
50	0.4462	0.2220	0.5432	2.9514	1.3361	2.5026	7.3496	3.3350	5.6680

Table 2. MSEs of classical and robust (MCD and S) estimators of the FMCV in the case of t-distribution with five degrees of freedom, Fourier basis, $n = 100$, $p = 5$ and $\mathbf{a} = \mathbf{a}_1$.

	class.	MCD	S	class.	MCD	S	class.	MCD	S
ε	FMCV $= 0.1, \rho = 0$			FMCV $= 0.1, \rho = 0.5$			FMCV $= 0.1, \rho = 0.8$		
0	0.0001	0.0004	0.0002	0.0002	0.0004	0.0002	0.0004	0.0006	0.0003
10	0.0017	0.0003	0.0001	0.0020	0.0004	0.0001	0.0028	0.0005	0.0003
20	0.0050	0.0002	0.0002	0.0057	0.0003	0.0002	0.0072	0.0004	0.0004
30	0.0086	0.0001	0.0008	0.0109	0.0002	0.0010	0.0143	0.0005	0.0014
40	0.0137	0.0005	0.0025	0.0156	0.0007	0.0028	0.0185	0.0011	0.0035
50	0.0189	0.0019	0.0056	0.0210	0.0024	0.0062	0.0264	0.0034	0.0074
ε	FMCV $= 0.5, \rho = 0$			FMCV $= 0.5, \rho = 0.5$			FMCV $= 0.5, \rho = 0.8$		
0	0.0044	0.0121	0.0062	0.0249	0.0201	0.0153	0.0719	0.0461	0.0435
10	0.0517	0.0098	0.0030	0.1420	0.0188	0.0176	0.3039	0.0481	0.0557
20	0.1282	0.0073	0.0039	0.3428	0.0217	0.0299	0.6786	0.0588	0.0858
30	0.2123	0.0045	0.0167	0.5727	0.0294	0.0717	1.0843	0.0849	0.1740
40	0.3010	0.0097	0.0543	0.8068	0.0631	0.1579	1.5770	0.1653	0.3360
50	0.4019	0.0359	0.1160	1.1104	0.1627	0.3023	2.1642	0.3614	0.6079
ε	FMCV $= 0.9, \rho = 0$			FMCV $= 0.9, \rho = 0.5$			FMCV $= 0.9, \rho = 0.8$		
0	0.0206	0.0535	0.0296	0.1373	0.0833	0.0775	0.4147	0.2098	0.2328
10	0.1608	0.0454	0.0157	0.6510	0.0797	0.0961	1.5298	0.2339	0.3127
20	0.3397	0.0364	0.0125	1.3905	0.1019	0.1729	3.0649	0.3039	0.5154
30	0.5174	0.0222	0.0347	1.9681	0.1479	0.3667	4.3468	0.4605	0.9566
40	0.6139	0.0222	0.1200	2.6334	0.3042	0.7510	5.9736	0.8762	1.8222
50	0.7742	0.0638	0.2467	3.3383	0.6651	1.2988	7.6180	1.7064	2.9605

Table 3. MSEs of classical and robust (MCD and S) estimators of the FMCV in the case of mixture of normal distributions, Fourier basis, $n = 100$, $p = 5$ and $\mathbf{a} = \mathbf{a}_1$.

	class.	MCD	S	class.	MCD	S	class.	MCD	S
ε	FMCV $= 0.1, \rho = 0$			FMCV $= 0.1, \rho = 0.5$			FMCV $= 0.1, \rho = 0.8$		
0	0.0001	0.0001	0.0001	0.0001	0.0006	0.0003	0.0003	0.0019	0.0011
10	0.0017	0.0001	0.0001	0.0019	0.0004	0.0002	0.0025	0.0013	0.0008
20	0.0047	0.0001	0.0004	0.0053	0.0003	0.0003	0.0064	0.0009	0.0008
30	0.0087	0.0002	0.0012	0.0097	0.0003	0.0004	0.0118	0.0008	0.0010
40	0.0134	0.0006	0.0035	0.0152	0.0003	0.0007	0.0183	0.0011	0.0010
50	0.0176	0.0024	0.0088	0.0202	0.0008	0.0038	0.0247	0.0031	0.0077
ε	FMCV $= 0.5, \rho = 0$			FMCV $= 0.5, \rho = 0.5$			FMCV $= 0.5, \rho = 0.8$		
0	0.0021	0.0043	0.0024	0.0216	0.0734	0.0248	0.0653	0.2783	0.1070
10	0.0379	0.0038	0.0027	0.1285	0.0469	0.0285	0.2856	0.1852	0.1222
20	0.1020	0.0037	0.0078	0.2947	0.0377	0.0641	0.6031	0.1437	0.2408
30	0.1759	0.0051	0.0247	0.5047	0.0616	0.3980	1.0405	0.2095	1.1201
40	0.2545	0.0128	0.0732	0.7491	0.3798	1.7084	1.5105	1.0244	3.3714
50	0.3208	0.0492	0.1862	0.9737	1.5144	2.7948	1.9882	2.7528	4.8398
ε	FMCV $= 0.9, \rho = 0$			FMCV $= 0.9, \rho = 0.5$			FMCV $= 0.9, \rho = 0.8$		
0	0.0132	0.0296	0.0151	0.1190	0.4660	0.2736	0.3770	1.2859	0.8901
10	0.0870	0.0229	0.0113	0.5964	0.3671	0.3287	1.5149	1.0949	1.0272
20	0.2114	0.0178	0.0172	1.1848	0.3196	0.6659	2.8603	0.9864	1.8050
30	0.3298	0.0150	0.0484	1.8255	0.4887	2.3572	4.3818	1.3322	4.8912
40	0.4291	0.0259	0.1609	2.4682	1.8147	4.9669	5.8772	3.8394	7.8594
50	0.5041	0.0935	0.3871	2.9989	3.0804	6.0332	7.3744	4.7974	9.2819

3.2 Simulation results

The simulation results presented in Tables 1–3 and Tables 1–7 in the Supplementary Materials. lead to the following conclusions concerning the finite sample behavior of the estimators of the FMCV under uncontaminated and contaminated data.

When the FMCV or the number of outliers increases, the MSEs of all estimators also increase. However, the increase in the MSE of the MCD estimator appears to be slower

than that of the other estimators when ε increases, except some cases for $\varepsilon \in \{40, 50\}$, FMCV $= 0.9$ and $\mathbf{a} = \mathbf{a}_2$ under mixture of normal distributions. Moreover, sometimes the MSE of the MCD (respectively S) estimator decreases with an increasing number of outliers up to 30% (respectively 10%); for example, under a t-distribution, $\mathbf{a} = \mathbf{a}_1$, FMCV $=$ 0.9 and $\rho = 0$.

Under uncontaminated data and normal distribution or mixture of normal distributions, the classical estimator is at least slightly better than the S-estimator, and much better than the MCD estimator. The situation is similar in the case of the t-distribution and small FMCV or lack of correlation. However, in the remaining cases, the robust estimators (especially the S-estimator) often perform better than the classical estimator.

Under contaminated data, the classical estimator breaks down, even when ε is small. This indicates, as was expected, the lack of robustness of the classical estimator based on the sample mean and the sample covariance matrix. On the other hand, the robust estimators are much more resistant to contamination. However, they may break down under mixture of normal distributions and FMCV $= 0.5, 0.9$, when ε is close to the breakdown point (see Table 6 in the Supplementary Materials). Fortunately, the MCD estimator overcomes this problem, when the number of observations increases (see Table 7 in the Supplementary Materials). The S estimator outperforms the MCD estimator when the number of outlying observations is small, i.e., $\varepsilon \leq 10$ or $\varepsilon \leq 20$, but the MCD estimator performs better when $\varepsilon \geq 20$ or $\varepsilon \geq 30$. The main reason for this is that the bias of the MCD estimator is greater than that of the S-estimator in the presence of intermediate outlying observations, but smaller than the bias of the S-estimator in case of severe contamination. The squared bias of the estimators behaves similarly to the MSE, and it is therefore omitted to save space.

The classical estimator usually performs worse under a t-distribution than under a normal distribution. For robust estimators, the same holds when $\mathbf{a} = \mathbf{a}_2$ and $\varepsilon = 0$, but in the other cases the reverse is true. Under mixture of normal distributions, the classical estimator usually performs worse than under normal distribution, when $\rho = 0$ or the Fourier basis is used, but the reverse is true in the other scenarios. On the other hand, the MSEs of the robust estimators are usually smaller under mixture of normal distributions than under normal distribution, when $\mathbf{a} = \mathbf{a}_1$, $\varepsilon > 0$, $\rho = 0$ or FMCV $= 0.1$, but the reverse is true in the other situations. The MSEs usually increase when the correlation increases in the case $\mathbf{a} = \mathbf{a}_1$. However, for $\mathbf{a} = \mathbf{a}_2$, the MSEs are quite stable for small FMCV, and for greater FMCV the MSEs of robust estimators may even decrease as the correlation increases, except the case of mixture of normal distributions, where the MSEs usually increase. These findings all indicate that the variability of the estimators of the FMCV depends on the direction of $\mathbf{a}$.

The MSEs are usually similar for different bases, but greater differences may be found when the FMCV is larger. Moreover, the MSE is often greater for the B-spline basis than for the Fourier basis. This may be because the cross product matrix in $\mathbf{J}_{\mathbf{\Phi}}$ corresponding to the B-spline basis is numerically approximated, while for the Fourier basis, it is equal to the identity matrix.

To summarize, the estimators of the FMCV perform very satisfactorily when the FMCV is smaller. However, for greater values, they may overestimate the FMCV. We also observed that it is necessary to use robust estimators when outlying observations or non-normal distribution of the data are suspected. The S-estimator performs better than the MCD estimator under uncontaminated or less contaminated functional data, while in the other cases the reverse is true.

4. REAL DATA APPLICATION

In this section, we demonstrate the applicability of the FMCV using an electrocardiogram (ECG) data set. We also complement the simulation results of Section 3 with some additional observations. We consider the ECG data set originating from [26] and available in the R package mfds [17].

Electrocardiography is a diagnostic procedure that monitors the electrical activity of the heart with the intention of diagnostic cardiac pathologies. An electrocardiogram is generated by placing one or more electrodes at standardized locations on the body, and recording the electrical potential difference observed at that site during each heartbeat; a complete ECG utilizes twelve electrodes, but fewer are often used for simpler diagnostic procedures. In our case, the data comes from two electrodes ($p = 2$). Each data set in the ECG database contains the measurements recorded by one electrode during one heartbeat. The data sets contained in each database were analyzed by experts, and a label of normal or abnormal (supraventricular premature beat) was assigned to each data set. The ECG database contains 200 ($n = 200$) data sets, where 133 were identified as normal and 67 were identified as abnormal. Data from ECG were recorded in 152 time points ($m_i = 152$, $i = 1, \ldots, n$) (Figure 2 on page 654).

The ECG database was used to discriminate between normal and abnormal heartbeats [26]. For illustrative purposes, we show that this also makes sense from the point of view of variability. Moreover, it seems that groups with greater variability are more difficult to classify, as we show for the ECG database at the end of this section.

Here, we compute the FMCV for normal and abnormal heartbeats separately. The basis function representation of the data was obtained using the Fourier and B-spline bases and $B_1 = B_2 = 5, 7, 9, 11, 13, 15$, if this was possible. (Odd values of the number of basis functions are dictated by the implementation of the Fourier basis in the R package fda [29] which was used.) To estimate the vectors of coefficients $\boldsymbol{\alpha}_i$, $i = 1, \ldots, n$, the least squares estimation method was used. To estimate the FMCV, we used the same estimators as in the simulation experiments of Section 3, namely

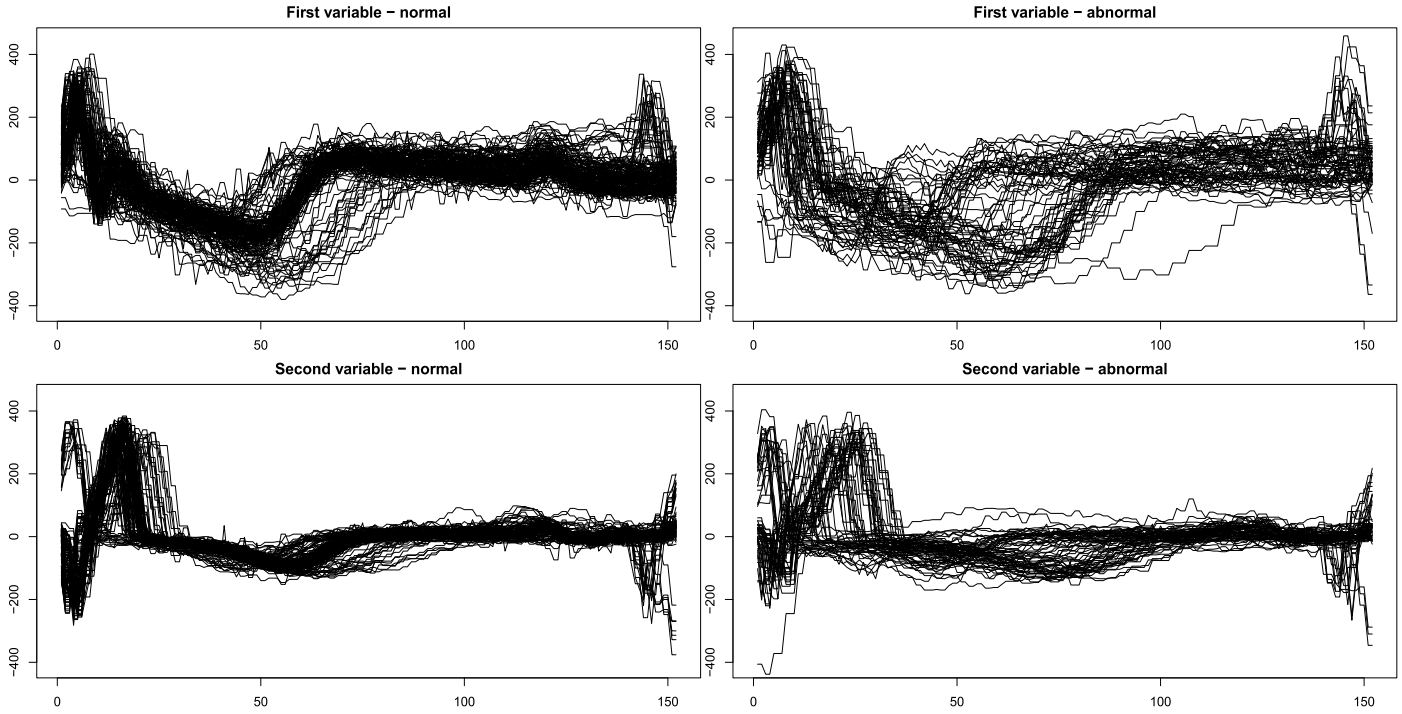

Figure 2. ECG data set.

the classical, MCD and S estimators, as well as the orthogonalized Gnanadesikan-Kettenring (OGK) estimator [11, 22]. Additionally, the standard errors (SE) were obtained by the bootstrap method, based on 1000 bootstrap samples. The estimated FMCVs and SEs are presented in Figure 3 on page 655.

Since to apply the MCD and S estimators one may need a larger number of observations, they may not be applicable when the number of basis functions is greater. This is seen when computing the SEs in this example: for the abnormal group, the MCD and S estimators do not exist for bootstrap samples. In contrast, the OGK estimator is a robust estimator of location and dispersion for high-dimensional data sets, and it can be used in this scenario. Moreover, for the normal group, the SEs of the MCD and S estimators increase fairly rapidly as the number of basis functions increases. These facts illustrate a limitation of some of the robust methods, namely that their application may require a greater amount of data. Fortunately, the SEs of the classical and OGK estimates are quite stable for all values of B_i, $i = 1, 2$. The results are also quite stable for both of the bases used, although sometimes greater differences can be observed.

In the normal group, we observe that the FMCVs for the classical and S estimators are greater than those for the MCD and OGK estimators. This confirms that the classical and S estimators may overestimate the FMCV, as was indicated in the simulations (Section 3). This may be caused by the presence of outlying observations (especially in the normal group) as indicated by the outlier detection method of [3]. Nevertheless, the FMCVs for the normal group are significantly smaller than those for the abnormal group. This implies that the ECGs for heartbeats representing cardiac pathology exhibit much greater variability than those for normal heartbeats. Thus, the FMCV confirms the correctness of the division of the heartbeats into normal and abnormal groups.

Finally, we justify our claim that the groups with greater variability may be more difficult to classify. For this purpose, we applied two different classification rules to the ECG data set, namely the linear discriminant analysis in the space of the multivariate functional discriminant coordinates [14] (the FDC classifier) and the classifier based on the functional logistic regression [13] (the FLR classifier). For simplicity, we used the Fourier basis only and the same B_1 and B_2 as for the FMCV. The leave-one-out (LOO) classification errors computed separately for all observations, the normal and abnormal groups are presented in Figure 4 on page 655. In fact, the LOO classification errors of both classifiers for the abnormal group are much greater than for the normal group, which justifies our claim at least for this particular data set.

5. CONCLUDING REMARKS

In functional data analysis, the data are considered as curves or functions. These appear naturally in many scientific fields where repeated measurements are taken in time or space. To extend the range of methods of functional data

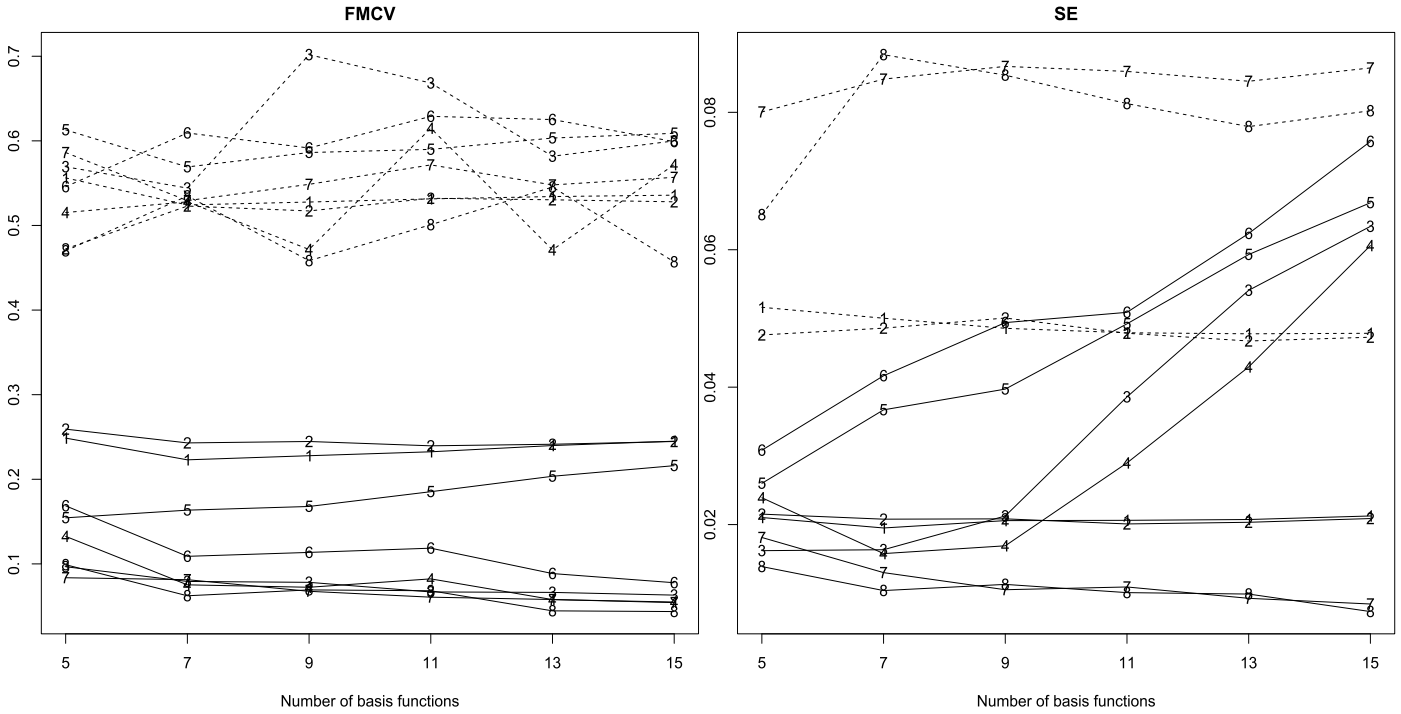

Figure 3. Estimated FMCVs and SEs for the normal (solid line) and abnormal (dashed line) groups, where 1 – classical estimator and Fourier basis, 2 – classical estimator and B-spline basis, 3 – MCD estimator and Fourier basis, 4 – MCD estimator and B-spline basis, 5 – S-estimator and Fourier basis, 6 – S-estimator and B-spline basis, 7 – OGK estimator and Fourier basis, 8 – OGK estimator and B-spline basis.

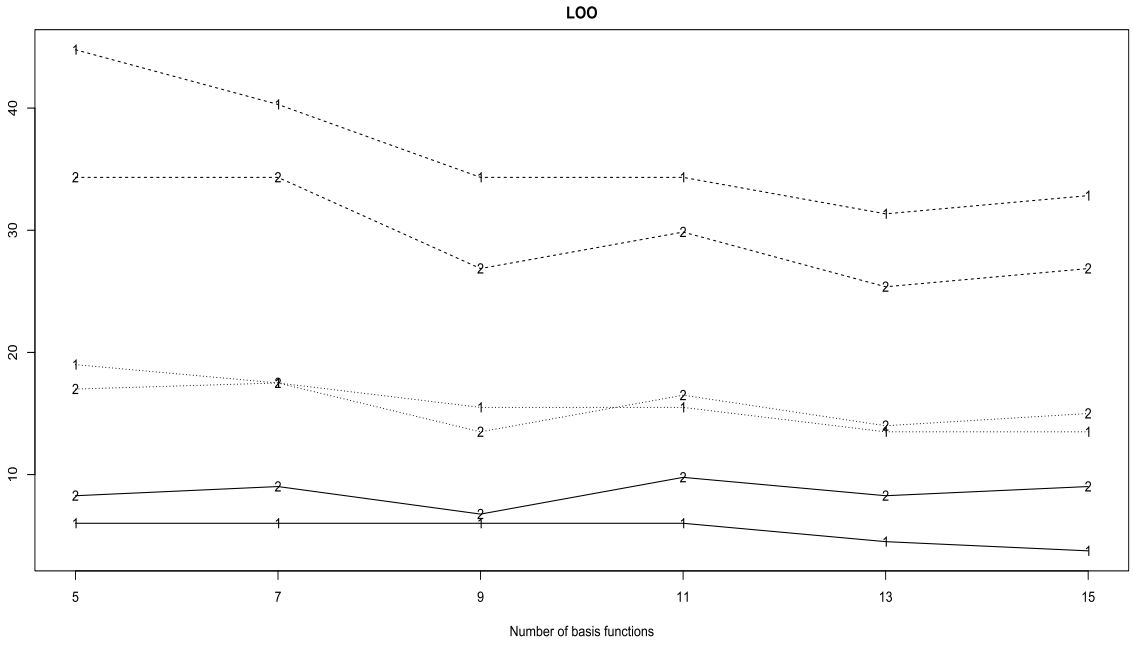

Figure 4. LOO classification errors (as percentages) for the normal (solid line), abnormal (dashed line) groups and all observations (dotted line), where 1 – FDC classifier, 2 – FLR classifier.

analysis, we have defined the functional multivariate coefficient of variation for both univariate and multi-dimensional functional data. Its theoretical properties show that the FMCV is well-defined and admits a reasonable interpretation. We have also proved a simple form of the FMCV using the basis expansion of the data, which is easy to implement. By an application to ECG data, treated as functional data, we illustrated the use of the FMCV to compare the relative variation of different groups and indicated that the groups with greater variability may be more difficult to classify. Possible applications of the FMCV also include comparison of the performance or reproducibility of different techniques or equipment, when they are described by certain functional variables.

The practical performance of the classical and robust estimators of the FMCV has been tested by simulation studies. For small values of the FMCV, its estimators perform very satisfactorily, but for greater values, they may overestimate the FMCV. Robust estimators were constructed by plugging in the robust estimators of the location and dispersion parameters to the basis form of the FMCV; that is, the MCD and S estimators. Under normal data without outliers, the classical sample mean and covariance matrix outperform the robust estimators, but the reverse is the case with non-normal data or in the presence of outliers. The S-estimator appears to perform better than the MCD estimator in the absence of contamination or under small contamination, but in other cases the position is reversed. In the real data application, we also noted that these robust estimators (and many others) may require a fairly large number of observations in order to be applied. Therefore, there may be a need to consider other estimators, such as OGK, which are also designed for high-dimensional data.

SUPPLEMENTARY MATERIAL

The Supplementary Materials contain the all simulation results of Section 3 (http://intlpress.com/site/pub/files/_supp/sii/2019/0012/0004/SII-2019-0012-0004-s003.pdf) as well as the R code reproducing the simulation results of Section 3 and the results of the real data example from Section 4 (http://intlpress.com/site/pub/files/_supp/sii/2019/0012/0004/SII-2019-0012-0004-s004.zip).

APPENDIX A. PROOF OF THEOREM 2.1

Proof of Theorem 2.1. Let $Y_i(t) = X_i(t) - \mu_i(t)$, $t \in [a,b]$, $i = 1, \ldots, p$. Then $\mathrm{E}Y_i = 0$ and $Y_i(t)$, $t \in [a,b]$, $i = 1, \ldots, p$, are square integrable, since:

$$\begin{aligned}
\mathrm{E}\|Y_i\|^2 &= \mathrm{E}\int_a^b Y_i^2(t)\,dt \\
&= \mathrm{E}\int_a^b (X_i(t) - \mu_i(t))^2\,dt \\
&= \mathrm{E}\int_a^b X_i^2(t)\,dt \\
&\quad -2\mathrm{E}\int_a^b X_i(t)\mu_i(t)\,dt \\
&\quad +\mathrm{E}\int_a^b \mu_i^2(t)\,dt \\
&= \mathrm{E}\|X_i\|^2 - 2\mathrm{E}\langle X_i, \mu_i\rangle + \|\mu_i\|^2 \\
&= \mathrm{E}\|X_i\|^2 - \|\mu_i\|^2 < \infty
\end{aligned}$$

(see [18] p. 23, for evidence of the last equality). Now, we conclude similarly to [18] (p. 23–24) that

$$\begin{aligned}
\mathrm{Var}(\langle \mu_{*i}, X_i\rangle) &= \mathrm{Var}(\langle \mu_{*i}, X_i\rangle - \langle \mu_{*i}, \mu_i\rangle) \\
&= \mathrm{Var}(\langle \mu_{*i}, X_i - \mu_i\rangle) \\
&= \mathrm{Var}(\langle \mu_{*i}, Y_i\rangle) \\
&= \mathrm{E}(\langle \mu_{*i}, Y_i\rangle^2) \\
&= \mathrm{E}\left(\int_a^b \mu_{*i}(t)Y_i(t)\,dt\right)^2 \\
&= \int_a^b\int_a^b \mathrm{E}\left(Y_i(t)Y_i(s)\right)\mu_{*i}(t)\mu_{*i}(s)\,dtds \\
&= \int_a^b C_{Y_i}(\mu_{*i})(t)\mu_{*i}(t)\,dt \\
&= \langle C_{Y_i}(\mu_{*i}), \mu_{*i}\rangle,
\end{aligned}$$

where C_{Y_i} is the covariance operator of $Y_i(t)$, $t \in [a,b]$. Therefore, $\mathrm{Var}(\langle \mu_{*i}, X_i\rangle)$ exists, which implies the existence of $\mathrm{Var}(\langle \boldsymbol{\mu}_*, \mathbf{X}\rangle)$ as

$$\begin{aligned}
\mathrm{Var}(\langle \boldsymbol{\mu}_*, \mathbf{X}\rangle) &= \mathrm{Var}\left(\sum_{i=1}^p \langle \mu_{*i}, X_i\rangle\right) \\
&= \sum_{i=1}^p \mathrm{Var}\left(\langle \mu_{*i}, X_i\rangle\right) \\
&\quad + 2\sum_{1\le i<j\le p} \mathrm{Cov}\left(\langle \mu_{*i}, X_i\rangle\langle \mu_{*j}, X_j\rangle\right) \\
&\le \sum_{i=1}^p \mathrm{Var}\left(\langle \mu_{*i}, X_i\rangle\right) \\
&\quad + 2\sum_{1\le i<j\le p} \sqrt{\mathrm{Var}\left(\langle \mu_{*i}, X_i\rangle\right)\mathrm{Var}\left(\langle \mu_{*j}, X_j\rangle\right)}.
\end{aligned}$$

The second statement follows from the first and from the following observation:

$$\begin{aligned}
\mathrm{E}(\langle \boldsymbol{\mu}_*, \mathbf{X}\rangle) &= \mathrm{E}\left(\sum_{i=1}^p \langle \mu_{*i}, X_i\rangle\right) \\
&= \sum_{i=1}^p \mathrm{E}\left(\langle \mu_{*i}, X_i\rangle\right) \\
&= \sum_{i=1}^p \langle \mu_{*i}, \mu_i\rangle
\end{aligned}$$

$$= \frac{1}{\|\boldsymbol{\mu}\|}\sum_{i=1}^{p}\langle \mu_i, \mu_i\rangle$$
$$= \|\boldsymbol{\mu}\|,$$

which completes the proof. □

ACKNOWLEDGEMENTS

The authors would like to thank the Reviewers for their very constructive suggestions which led to an improvement in this paper.

Received 25 September 2018

REFERENCES

[1] Aerts, S., Haesbroeck, G. and Ruwet, C. (2015). Multivariate coefficients of variation: comparison and influence functions. *J. Multivar. Anal.* **142** 183–198. MR3412747

[2] Albert, A. and Zhang, L. (2010). A novel definition of the multivariate coefficient of variation. *Biom. J.* **52** 667–675. MR2757012

[3] Arribas-Gil, A. and Romo, J. (2014). Shape outlier detection and visualization for functional data: the outliergram. *Biostatistics* **15** 603–619.

[4] Boente, G., Barrera, M. S. and Tyler, D. E. (2014). A characterization of elliptical distributions and some optimality properties of principal components for functional data. *J. Multivar. Anal.* **131** 254–264. MR3252648

[5] Coffey, N., Hinde, J. and Holian, E. (2014). Clustering longitudinal profiles using P-splines and mixed effects models applied to time-course gene expression data. *Comput. Stat. Data Anal.* **71** 14–29. MR3131951

[6] Davies, P. L. (1987). Asymptotic behavior of S-estimators of multivariate location parameters and dispersion matrices. *Ann. Statist.* **15** 1269–1292. MR0902258

[7] Delaigle, A. and Hall, P. (2013). Classification using censored functional data. *J. Amer. Statist. Assoc.* **108** 1269–1283. MR3174707

[8] Epifanio, I. and Ventura-Campos, N. (2014). Hippocampal shape analysis in Alzheimer's disease using functional data analysis. *Stat. Med.* **33** 867–880. MR3249093

[9] Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis: Theory and practice*. Springer, New York. MR2229687

[10] Giacofci, M., Lambert-Lacroix, S., Marot, G. and Picard, F. (2013). Wavelet-based clustering for mixed-effects functional models in high dimension. *Biometrics* **69** 31–40. MR3058049

[11] Gnanadesikan, R. and Kettenring, J. R. (1972). Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics* **28** 81–124.

[12] Goldsmith, J. and Schwartz, J. E. (2017). Variable selection in the functional linear concurrent model. *Stat. Med.* **36** 2237–2250. MR3660128

[13] Górecki, T., Krzyśko, M. and Wołyński, W. (2015). Classification problem based on regression models for multidimensional functional data. *Statistics in Transition New Series* **16** 97–110.

[14] Górecki, T., Krzyśko, M., Waszak, Ł. and Wołyński, W. (2018). Selected statistical methods of data analysis for multivariate functional data. *Statist. Papers* **59** 153–182. MR3765940

[15] Górecki, T. and Smaga, Ł. (2015). A comparison of tests for the one-way ANOVA problem for functional data. *Comput. Statist.* **30** 987–1010. MR3433439

[16] Górecki, T. and Smaga, Ł. (2017a). Multivariate analysis of variance for functional data. *J. Appl. Stat.* **44** 2172–2189. MR3670297

[17] Górecki, T. and Smaga, Ł. (2017b). mfds: Multivariate functional data sets. R package version 0.1.0, https://github.com/Halmaris/mfds Accessed 20 September 2018.

[18] Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*. Springer-Verlag, New York. MR2920735

[19] Kayano, M. and Konishi, S. (2009). Functional principal component analysis via regularized Gaussian basis expansions and its application to unbalanced data. *J. Stat. Plann. Inference* **139** 2388–2398. MR2508000

[20] Keser, I. K. and Kocakoç, I. D. (2015). Smoothed functional canonical correlation analysis of humidity and temperature data. *J. Appl. Stat.* **42** 2126–2140. MR3373723

[21] Kim, J. S., Maity, A. and Staicu, A.-M. (2018). Additive nonlinear functional concurrent model. *Stat. Interface* **11** 669–685. MR3858523

[22] Maronna, R. and Zamar, R. (2002). Robust estimation of location and dispersion for high-dimensional datasets. *Technometrics* **44** 307–317. MR1939680

[23] Martin-Barragan, B., Lillo, R. and Romo, J. (2014). Interpretable support vector machines for functional data. *European J. Oper. Res.* **232** 146–155.

[24] Martínez-Camblor, P. and Corral, N. (2011). Repeated measures analysis for functional data. *Comput. Stat. Data Anal.* **55** 3244–3256. MR2825407

[25] Morris, J. S. (2015). Functional regression. *Annual Review of Statistics and Its Application* **2** 321–359.

[26] Olszewski, R. T. (2001). Generalized feature extraction for structural pattern recognition in time-series data. Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, PA, http://www.cs.cmu.edu/~bobski Accessed 20 September 2018.

[27] Ramsay, J. O. and Silverman, B. W. (2002). *Applied functional data analysis. Methods and case studies*. Springer, New York. MR1910407

[28] Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*, 2nd ed. Springer, New York. MR2168993

[29] Ramsay, J. O., Wickham, H., Graves, S. and Hooker, G. (2018). fda: Functional data analysis. R package version 2.4.8, http://CRAN.R-project.org/package=fda Accessed 20 September 2018.

[30] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ Accessed 20 September 2018.

[31] Reyment, R. A. (1960). Studies on Nigerian Upper Cretaceous and Lower Tertiary Ostracoda: part 1. Senonian and Maastrichtian Ostracoda. *Stockholm Contributions in Geology* **7** 1–238.

[32] Rincón, M. and Ruiz-Medina, M. D. (2012). Local wavelet-vaguelette-based functional classification of gene expression data. *Biom. J.* **54** 75–93. MR2868981

[33] Rousseeuw, P. J. (1985). Multivariate estimation with high breakdown point. In W. Grossmann, G. Pflug, I. Vincze, & W. Wertz (Eds.), *Mathematical Statistics and Applications* (pp. 283–297). Dordrecht: Reidel Publishing. MR0851060

[34] Shmueli, G. (2010). To explain or to predict? *Statist. Sci.* **25** 289–310. MR2791669

[35] Smaga, Ł. (2017). Repeated measures analysis for functional data using Box-type approximation – with applications. *REVSTAT* (in press). MR2716455

[36] Sørensen, H., Goldsmith, J. and Sangalli, L. M. (2013). An introduction with medical applications to functional data analysis. *Stat. Med.* **32** 5222–5240. MR3141370

[37] Todorov, V. and Filzmoser, P. (2009). An object-oriented framework for robust multivariate analysis. *Journal of Statistical Software* **32** 1–47.

[38] Todorov, V. and Pires, A. M. (2007). Comparative performance of several robust linear discriminant analysis methods. *REVSTAT* **5** 63–83. MR2365933

[39] Wang, J. L., Chiou, J. M. and Müller, H. G. (2016). Functional data analysis. *Annual Review of Statistics and Its Application* **3** 257–292.

[40] Van Valen, L. (1974). Multivariate structural statistics in natural history. *J. Theoret. Biol.* **45** 235–247.

[41] Voinov, V. G. and Nikulin, M. S. (1996). *Unbiased estimators and their applications, Vol. 2, Multivariate Case*. Dordrecht: Kluwer. MR1422256
[42] Zhang, J. T. (2013). *Analysis of variance for functional data*. Chapman & Hall, London.

Mirosław Krzyśko
Interfaculty Institute of Mathematics and Statistics
The President Stanisław Wojciechowski State University of Applied Sciences in Kalisz
Nowy Świat 4
62–800 Kalisz
Poland
E-mail address: mkrzysko@amu.edu.pl

Łukasz Smaga
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Uniwersytetu Poznańskiego 4
61–614 Poznań
Poland
E-mail address: ls@amu.edu.pl