TEST
https://doi.org/10.1007/s11749-018-0598-y

ORIGINAL PAPER

Robust simultaneous inference for the mean function of functional data

Italo R. Lima[1] · **Guanqun Cao**[1] · **Nedret Billor**[1]

Received: 16 December 2016 / Accepted: 10 July 2018

Abstract
A robust framework is proposed, based on polynomial spline estimation technique, for the estimation of the mean function of dense functional data, together with a simultaneous confidence band for the mean function. The robust simultaneous confidence band is also extended to the difference of mean functions of two populations. The performance of the proposed robust methods is evaluated with the simulation study and real data examples.

Keywords Confidence band · Functional data analysis · Least absolute deviation · Robust statistics

Mathematics Subject Classification 62G08 · 62G35

1 Introduction

Advancements in modern technology have enabled the collection of complex, high-dimensional data sets, such as curves, 2D or 3D images, and other objects living in a

Cao's research is supported in part by the Simons Foundation under Grant #354917 and National Science Foundation under Grant DMS 1736470. The authors thank the editor, the associate editor and reviewers for their constructive comments that have led to a dramatic improvement of the earlier version of this article.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11749-018-0598-y) contains supplementary material, which is available to authorized users.

✉ Guanqun Cao
gzc0009@auburn.edu

Italo R. Lima
italo@auburn.edu

Nedret Billor
billone@auburn.edu

[1] Department of Mathematics and Statistics, Auburn University, Auburn, AL 36849, USA

functional space, thus boosting the investigation of function data. This phenomenon affects all the fields involving applied statistics such as Geophysics (Ferraty et al. 2005), Environmetrics (Febrero et al. 2008), Ecology (Embling et al. 2012) and others. A general overview of functional data analysis (FDA) and an extensive list of references can be found in the seminal work of Ramsay and Silverman (2005) and Ferraty and Vieu (2006). Recent studies by Aneiros et al. (2017), Cuevas (2014) and Goia and Philippe (2016) contain latest advances in the FDA field. In addition, there are two excellent references by Hsing and Eubank (2015) and Horváth and Kokoszka (2012) which provide the core mathematical concepts related to the theoretical development of FDA and inferential methods such as statistical hypothesis tests for various functional data analytic settings.

Last two decades, many FDA techniques have been developed as extensions of multivariate data analysis techniques, such as regression, classification, clustering. However, the majority of these FDA methods require the homogeneity of data, i.e., free of outliers. It is well known that these methods are not robust in the presence of outlier curves. Although robustness has been studied extensively in multivariate data analysis framework, this has not been the case in the FDA. Only recently, there have been some studies of outlying-resistant methods for various FDA problems. Fraiman and Muniz (2001) propose trimmed mean to measure the centrality of a given curve within a group of curves based on data depth and Cuevas et al. (2007) study robust estimation and classification method for functional data via the notion of projection-based depth. Then, Gervini (2008) proposes an R-estimator for the location parameter and robust functional principal components based on the spherical principal components. López-Pintado and Romo (2009) propose depth-based approaches for the median and trimmed mean functions as well as a rank test for the problem of testing whether two groups of curves come from the same population. Recently, an R-estimator (robust estimator) of the median in Hilbert spaces is also proposed by Cardot et al. (2013).

Also, Bali et al. (2011) and Lee et al. (2013) propose R-estimators for the functional principal components. Kraus and Panaretos (2012) study the R estimation of the dispersion operator containing influential observations. More recently, Maronna and Yohai (2013) and Shin and Lee (2016) propose alternative R-estimators for the regression coefficient functions in functional linear regression models.

In this work, our objective is to develop outlying-resistant methods that can provide valid statistical inference even in the presence of a significant proportion of outlier curves. In particular, outlier-resistant simultaneous inference for the location function will be studied.

There are some studies discussing robust confidence intervals for the location parameter in the finite-dimensional setting, such as Du Mond and Lenth (1987) and Fraiman et al. (2001). In addition, Adrover et al. (2004) define globally R confidence intervals and p values for the location and simple linear regression models.

In the presence of outlier curves, the estimate of the functional mean, and therefore simultaneous confidence band (SCB), may be affected poorly which would yield misleading statistical conclusions. Further, the presence of outliers is amplified by the inherent complexity of functional spaces. This may lead to arising different types of outlier curves which add further complication to estimating functional mean and constructing SCB for the functional mean. It is of particular interest in data analysis

to build SCB for the mean function instead of point-wise confidence intervals and to develop global test statistics for the general hypothesis testing problem on the location functions. For example, Cao et al. (2012a, b) construct asymptotic SCB for the mean and derivative functions in FDA without considering any outliers. However, all the existing literature on constructing SCB for the mean function assume that observed functional data are free of outlier curves. To develop the R simultaneous inference for functional responses, we encounter many new challenges. First, the greater technical difficulty is to formulate SCB for a mean function of infinite-dimensional functional response and establish their theoretical properties. Second, unlike the scenarios considered in the classical FDA literature, in our settings, however, there is complex and unknown outlier structure.

In this paper, our main contribution is to construct R simultaneous confidence band (RSCB) based on R-estimators of the mean function and covariance function using the least absolute deviation and spline smoothing methods. We note that all currently available methods cannot be immediately used for constructing an R version of SCB for mean functions. We further extend the simultaneous inference to the two-sample case and evaluate the equality of mean functions from two groups when atypical curves exist. Our Monte Carlo results show that the proposed bands are superior to existing classical methods which do not account for atypical curves. As this paper is the first attempt to provide RSCB for functional data, there are some limitations. First, we only provide the asymptotic consistency of the proposed estimator. The theoretical justification of the proposed RSCB is a promising future work. Second, due to the richness of the types of functional outliers that may occur in functional data, it may not be possible to obtain one method that works for all. Therefore, the proposed method also needs careful inference under certain types of outliers, for example, the random bumps outlier (see the Supplement file).

The paper is organized as follows: We first introduce an FDA model in Sect. 2. Then we propose an R-estimator for the mean function when the functional data contain outliers. The RSCB based on the proposed R-estimators of the mean and covariance functions will be derived in Sect. 3. In addition, we extend this method to form the RSCB for the difference of mean functions of two populations. In Sects. 4 and 5, the performance of the proposed R methods and their robustness is demonstrated with simulation and real data examples. Finally, we conclude our paper with a discussion. Appendix contains technical proofs and additional simulation results.

2 Model

A functional data set can be defined as a collection of i.i.d. random samples, $\{\eta_i(x)\}_{i=1}^n$, where i is the subject index, from a smooth and square integrable random function $\eta(x) \in L^2$, with unknown mean function, $\mathbb{E}[\eta(x)] = m(x)$, and unknown covariance function $G(x, x') = \text{cov}[\eta(x), \eta(x')]$. For simplicity reasons, the domain of $\eta(\cdot)$ is assumed as $[0, 1]$. In this paper, we assume an equally spaced dense design, that is, each random curve $\eta_i(\cdot)$ is measured at the points $X_{ij} = j/N, 1 \le j \le N, 1 \le i \le n$, where N goes to infinity when sample size n goes to infinity. Then, the jth observation for the ith subject can be written as

$$Y_{ij} = \eta_i\,(j/N) + \sigma\varepsilon_{ij}, \tag{1}$$

where errors ε_{ij}'s are independent and assumed to satisfy $\mathbb{E}(\varepsilon_{ij}) = 0$ and $\mathbb{E}(\varepsilon_{ij}^2) = 1$, and the stochastic process $\mathbb{E}\int_{[0,1]}\eta^2(x)\mathrm{d}x < \infty$.

The process $\eta(x)$ can be written, based on Karhunen–Loève L^2 representation, as $\eta(x) = m(x) + \sum_{k=1}^{\infty}\xi_k\phi_k(x)$, where ξ_k's are uncorrelated random coefficients with mean zero and variance one, and $\phi_k(x) = \sqrt{\lambda_k}\varphi_k(x)$. Here, $\{\lambda_k\}_{k=1}^{\infty}$ and $\{\varphi_k(x)\}_{k=1}^{\infty}$ are the eigenvalues and eigenfunctions, which form an orthonormal basis of L^2, of the covariance function $G(x, x')$. We assume that $\phi_k(\cdot)$'s are kept in a descending order of λ_k's, i.e., $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$. Assume that $\lambda_k = 0$ for $k > \kappa$, where κ is a positive integer or ∞. This implies that the eigenvalue decomposition of $G(x, x')$ is

$$G(x, x') = \sum_{k=1}^{\kappa}\lambda_k\varphi_k(x)\varphi_k(x') = \sum_{k=1}^{\kappa}\phi_k(x)\phi_k(x'). \tag{2}$$

With this representation, we rewrite model (1) as

$$Y_{ij} = m\,(j/N) + e_{ij}, \tag{3}$$

where $e_{ij} = \sum_{k=1}^{\kappa}\xi_{ik}\phi_k(j/N) + \sigma\varepsilon_{ij}$, $1 \leq i \leq n$, $1 \leq j \leq N$. Although the existence of $\{\lambda_k\}_{k=1}^{\kappa}$ and $\{\phi_k\}_{k=1}^{\kappa}$, and the random coefficients $\{\xi_{ik}\}_{k=1}^{\kappa}$ are guaranteed mathematically, they are unknown and unobservable.

3 Method

3.1 The classical estimation of the mean function

We first review an estimation procedure that approximates the mean function by polynomial splines. Let $t_{1-p} = \cdots = t_0 = 0 < t_1 < \cdots < t_{N_m} < 1 = t_{N_m+1} = \cdots = t_{N_m+p}$ be equally spaced points over $[0, 1]$, called interior knots, in which $t_J = Jh_m$, $0 \leq J \leq N_m$, and $h_m = 1/(N_m + 1)$ is the distance between neighboring knots. Denote by $\mathcal{H}^{(p-2)}$ the pth-order spline space, i.e., $p-2$ times continuously differentiable functions on $[0, 1]$ that are polynomials of degree $p-1$ on $[t_J, t_{J+1}]$, $J = 0, \ldots, N_m$.

Cao et al. (2012b) propose to approximate the mean function $m(\cdot)$ by a linear combination of spline basis: $\hat{m}(x) = \sum_{J=1-p}^{N_m}\hat{\beta}_J B_{J,p}(x)$, where $B_{J,p}$ be the Jth B-spline basis of order p defined in de Boor (2001), and the coefficients $\{\hat{\beta}_{1-p}, \ldots, \hat{\beta}_{N_m}\}^{\mathrm{T}}$ are the solutions of the following least squares problem

$$\left\{\hat{\beta}_{1-p}, \ldots, \hat{\beta}_{N_m}\right\}^{\mathrm{T}} = \underset{\{\beta_{1-p},\ldots,\beta_{N_m}\}\in R^{N_m+p}}{\arg\min}\sum_{i=1}^{n}\sum_{j=1}^{N}\left\{Y_{ij} - \sum_{J=1-p}^{N_m}\beta_J B_J\,(j/N)\right\}^2.$$

Let $\bar{\mathbf{Y}} = (\bar{Y}_{\cdot 1}, \ldots, \bar{Y}_{\cdot N})^{\mathrm{T}}$, where $\bar{Y}_{\cdot j} = 1/n \sum_{i=1} Y_{ij}$, $j = 1, \ldots, N$. One advantage of the least square estimator is the existence of a closed-form solution. Indeed, applying elementary algebra, one obtains

$$\hat{m}(x) = \mathbf{B}_p(x) \left(\mathbf{B}^{\mathrm{T}}\mathbf{B}\right)^{-1} \mathbf{B}^{\mathrm{T}}\bar{\mathbf{Y}}, \tag{4}$$

where $\mathbf{B}_p(x) = (B_{1-p,p}(x), \ldots, B_{N_m,p}(x))$ and $\mathbf{B} = (\mathbf{B}_p^{\mathrm{T}}(1/N), \ldots, \mathbf{B}_p^{\mathrm{T}}(N/N))^{\mathrm{T}}$ is the design matrix.

3.2 The robust estimation of the mean function

The ordinary least square (OLS) estimation, used in Eq. (4), is susceptible to the presence of outliers. To circumvent this non-robustness, we propose to replace the OLS by the *least absolute deviation* (LAD). The mean function is estimated by a linear combination of spline basis,

$$\tilde{m}(x) = \sum_{J=1-p}^{N_m} \tilde{\beta}_J B_{J,p}(x), \tag{5}$$

where

$$\left\{\tilde{\beta}_{1-p}, \ldots, \tilde{\beta}_{N_m}\right\}^{\mathrm{T}} = \underset{\{\beta_{1-p},\ldots,\beta_{N_m}\}\in R^{N_m+p}}{\arg\min} \sum_{i=1}^{n}\sum_{j=1}^{N} \left| Y_{ij} - \sum_{J=1-p}^{N_m} \beta_J B_{J,p}(j/N) \right|.$$

LAD gives equal emphasis to all observations, in contrast to OLS which, by squaring the residuals, gives more weight to large residuals. This is helpful in studies where outliers do not need to be given greater weight than other observations. In this work, we use cubic spline ($p = 4$) basis and the number of interior knots N_m is taken to be $[2n^{1/9}]$ and $[a]$ denotes the integer part of a, which is according to the following Assumption (A3) and the recommendation in Cao et al. (2012b).

For any $r \in (0, 1]$, we denote $C^{q,r}[0, 1]$ as the space of Hölder continuous functions on $[0, 1]$, $C^{q,r}[0, 1] = \left\{\phi : \sup_{t\neq s, t,s\in[0,1]} \frac{|\phi^{(q)}(t)-\phi^{(q)}(s)|}{|t-s|^r} < +\infty\right\}$. The following technical assumptions are needed by the following Theorem 1.

(A1) *The regression function* $m \in C^{p_0-1,1}[0, 1]$. *The spline order in estimating* m *satisfies* $p \geq p_0$.

(A2) *The standard deviation* $\sigma > 0$ *and for any* $k = 1, 2, \ldots \kappa, \phi_k(x) \in C^{p,1}[0, 1]$ *and* $\min_{x\in[0,1]} G(x, x) > 0$;

(A3) *The number of knots* N_m *satisfies* $N_m \sim n^{1/(2p+1)}$ *and the number of observation* N *satisfies* $N \sim n$.

(A4) *Let* f_{ij} *be the density function of* e_{ij}. *Uniformly over* i *and* j, $f_{ij}(0)$ *is bounded from infinity, and it is bounded away from zero and has a bounded first derivative in the neighborhood of zero. In model* (1), *for any* i *and* j, *the measurement errors* ε_{ij} *have a symmetric distribution and are independent of* η_i.

Assumptions (A1)–(A3) are standard in the spline smoothing and FDA literature; see Cao et al. (2012b), for instance. In particular, (A1) guarantees the orders of the bias terms of the spline smoothers for $m(x)$. Assumption (A2) ensures the covariance function is a uniformly bounded function. Assumption (A3) implies the number of points on each curve N diverges to infinity as n goes to infinity, which is a well-developed asymptotic scenario for dense functional data. The smoothness of our estimator is controlled by the number of knots, which increases to infinity as specified in (A3). This increasing knots asymptotic framework guarantees the richness of the basis. Assumption (A4) is a standard assumption in quantile regression; see Theorem 1 in Wang et al. (2009), for example. Note that if ε_{ij}'s have different distributions across individuals, then the statement on Assumption (A4) does not depend on i.

In the following theorem, we show that the proposed R-estimator, $\tilde{m}(x)$, is a consistent estimator of the true mean function, $m(x)$. The proof of this theorem is provided in Appendix.

Theorem 1 *Under Assumptions* (A1)–(A4)*, one has*

$$\frac{1}{N}\sum_{j=1}^{N}\{\tilde{m}(j/N) - m(j/N)\}^2 = O_P(n^{-2p/(1+2p)}).$$

3.3 The RSCB for the mean function

In the following, we mimic the construction of SCB procedure in Cao et al. (2012b) to obtain the RSCB for the mean function. Namely, first we obtain an estimator of the $(1-\alpha)100\%$ quantile, $Q_{1-\alpha}$, of the absolute maxima distribution for a standardized Gaussian process $\zeta(x)$, with $\mathbb{E}[\zeta(x)] = 0$, and $\mathbb{E}[\zeta^2(x)] = 1$, and covariance function $\mathbb{E}\left[\zeta(x)\zeta(x')\right] = G(x,x')\left\{G(x,x)G(x',x')\right\}^{-1/2}$. For any $\alpha \in (0,1)$, we denote $Q_{1-\alpha}$ the $100(1-\alpha)$th percentile of the absolute maxima distribution of $\zeta(x)$, i.e., $\mathbb{P}(\sup_{x\in[0,1]}|\zeta(x)| < Q_{1-\alpha}) = 1-\alpha$, $0 < \alpha < 1$. This can be done using a Monte Carlo simulation, that is, we first simulate $\zeta_l(x) = \tilde{G}(x,x)^{-1/2}\sum_{k=1}^{\kappa} Z_{k,l}\sqrt{\tilde{\lambda}_k}\tilde{\varphi}_k(x)$, where $\tilde{G}(x,x)$ is an R-estimator of covariance function G and $Z_{k,l}$ follows i.i.d. standard normal distribution, $1 \le k \le \kappa$, $l = 1,\ldots,L$, and L is a large positive integer, say 1000. We provide the definition of $\tilde{G}(x,x)$ in the next section. Next, we choose the number κ of eigenfunctions by using the following standard "fraction of variation explained", i.e., select the number of eigenvalues that can explain, say, 95% of the variation in the data. Then an estimator for $Q_{1-\alpha}$, $\hat{Q}_{1-\alpha}$, is obtained as the empirical $(1-\alpha)$ quantile of the set $\{\sup_{x\in[0,1]}|\zeta_l(x)|, l = 1,\ldots,L\}$.

Similar to the SCB investigated in Cao et al. (2012b), RSCB for the mean function may be calculated as $\tilde{m}(x) \pm n^{-1/2}\tilde{G}(x,x)^{1/2}\hat{Q}_{1-\alpha}$. However, simulations and real examples show that the RSCB does not have consistent coverage rates for relatively small sample sizes. To solve this problem, we propose the $(1-\alpha)100\%$ RSCB for the mean function as

$$\tilde{m}(x) \pm c(n)\cdot n^{-1/2}\tilde{G}(x,x)^{1/2}\hat{Q}_{1-\alpha},$$

where $c(n)$ is a finite-sample correction factor and depends on sample size n (Pison et al. 2002; Wilcox 2005). In order to determine the value of $c(n)$, we conduct a simulation study and discuss it in Sect. 4.2.

Theorem 1 is only a first step toward a theoretical justification of the proposed methodology. The asymptotic theory of the confidence bands based on $\tilde{m}$ is indeed a challenging work. First, the explicit formula for $\tilde{m}$ is not available. Thus, it is not possible to obtain an estimate for the covariance function of $\tilde{m}$ and the error decomposition. Hence, the strong approximation used in Cao et al. (2012b) cannot be applied directly here. Second, the asymptotic distribution of the LAD estimator depends on the value of the error density at 0. Due to its infinite-dimensional structure, an analogous assumption for the process $\eta(x)$ is not feasible.

3.4 The robust estimator for the covariance function

In order to construct an R confidence band for the mean function, we also need to obtain an R-estimator for the covariance function, $G(x, x')$. The covariance function defined in (2) can be recovered if the eigenfunction/eigenvalue decomposition is known. A popular estimator for the eigenfunctions and eigenvalues is derived from the eigenvalue decomposition of the empirical or the smoothed covariance matrix. Such method has been widely used in FDA (Yao et al. 2005; Cao et al. 2012b). However, this estimator is sensitive to the presence of outliers.

An alternative R-estimator, the spherical principal components, is proposed in Locantore et al. (1999). The first step in this method is to normalize each sample curve to mitigate the effect of outliers. That is, the covariance function is replaced by the *normalized covariance function*, i.e.,

$$\rho(x, x') = \mathbb{E}\left\{ \frac{[\eta(x) - m(x)][\eta(x') - m(x')]}{\|\eta(\cdot) - m(\cdot)\|_2^2} \right\},$$

where $\|\cdot\|_2$ is the usual L^2-norm. We can then find the eigenfunction/eigenvalue decomposition of this new operator, leading to eigenfunctions φ_k^* and eigenvalues λ_k^*. When $\kappa < \infty$ and $\{\xi_k\}_{k=1}^{\kappa}$ has a symmetric and interchangeable marginal distribution, Gervini (2008) has shown that $\varphi_k^* = \varphi_k$, that is, the eigenfunctions of the covariance operator are the same as the eigenfunctions of normalized one, and they are in the same order. However, the corresponding two types of eigenvalues are not necessarily be the same. These R-estimators are generally more robust to outliers than the commonly used sample mean and principal components. Moreover, Boente et al. (2014) extend the above results to the case of $\kappa = \infty$.

Since it is impossible to have the entire stochastic process and the true mean function in the practice, we use the discretized sample version of the normalized covariance function,

$$\rho_n(j/N, j'/N) = \frac{1}{n} \sum_{i=1}^{n} \frac{(Y_{ij} - \tilde{m}(j/N))(Y_{ij'} - \tilde{m}(j'/N))}{\|Y_{i\cdot} - \tilde{m}(\cdot)\|_2^2}, \quad 1 \le j,\ j' \le N, \quad (6)$$

where $\|Y_{i\cdot} - \tilde{m}(\cdot)\|_2^2 = 1/N \sum_{j=1}^{N} \left(Y_{ij} - \tilde{m}(j/N)\right)^2$ and $\tilde{m}(\cdot)$ is the proposed R-estimator of the mean function, defined in (5). Similarly to the population version, we can find the eigenvector decomposition $\{\varphi_k^*(j/N)\}_{k=1}^{\kappa}$ for the sample version of the normalized covariance function and the eigenfunctions, φ_k^* are called *spherical principal components* (Locantore et al. 1999) as ρ_n is the sample covariance function of the centered curves projected on the unit sphere.

The presence of measurement errors in our model adds a layer of contamination to the calculation of (6), but due to the aforementioned robustness of the spherical principal components, if the measurement errors are not excessive, φ_k^* will be close to the actual φ_k. In order to reduce the effects of measurement errors further, we smooth each φ_k^* using the B-splines which yield the smoothed eigenfunctions $\tilde{\varphi}_k(\cdot)$.

To estimate the covariance function in (2), we need to recover the eigenvalues λ_k. Notice that the kth eigenvalue is the variance of the projection of the centralized stochastic process on the kth eigenfunction, that is $\lambda_k = \text{Var}\left(\langle \eta(\cdot) - m(\cdot), \varphi_k(\cdot)\rangle\right)$. Adapting this to our model, we can use the proposed R-estimator $\tilde{m}$ and the R-estimator of the eigenfunctions $\tilde{\varphi}_k$ to obtain $\tilde{\lambda}_k$ as the square of the R-estimator of the scale of $\langle Y_{i\cdot} - \tilde{m}(\cdot), \tilde{\varphi}_k(\cdot)\rangle$. Having estimated the eigenvalues and eigenfunctions, we can recover the R-estimator of the covariance function as $\tilde{G}(x, x') = \sum_{k=1}^{\kappa} \tilde{\lambda}_k \tilde{\varphi}_k(x)\tilde{\varphi}_k(x')$. The R estimation of the covariance function can be summarized in the following steps.

Step 1. *Estimate eigenfunction* $\tilde{\varphi}_k(\cdot)$.
Apply the spline smoothing $\tilde{\varphi}_k(x) = \mathbf{B}_p(x)\left(\mathbf{B}^{\intercal}\mathbf{B}\right)^{-1}\mathbf{B}^{\intercal}\boldsymbol{\varphi}_k^*$, where $\boldsymbol{\varphi}_k^* = (\varphi_k^*(1/N), \ldots, \varphi_k^*(N/N))^{\intercal}$, $k = 1, 2, \ldots, N$, are the eigenvectors of the sample normalized covariance matrix ρ_n in (6).

Step 2. *Estimate eigenvalue* $\tilde{\lambda}_k$.
Using an R scale estimator of the projection of the centralized data onto $\tilde{\varphi}_k$. The chosen R-estimator is the Huber's M-estimator (Huber and Ronchetti 2009), i.e., the solution of the equation

$$H(\sigma) = \frac{1}{n}\sum_{i=1}^{n} \rho_\gamma\left(\frac{\frac{1}{N}\sum_{j=1}^{N}(Y_{ij} - \tilde{m}(j/N))\tilde{\varphi}_k(j/N)}{\sigma}\right) = 0.5,$$

where $\rho_\gamma(u) = \rho(u/\gamma), \gamma > 0$, and ρ is the Tukey's bisquare function. This choice guarantees that the scale estimators of i.i.d. observations has $1/2$ breakdown point. We use $\gamma = 1.5$ in the following numerical studies, which is the suggested value in Huber and Ronchetti (2009). This leads to the estimators $\tilde{\lambda}_k^{1/2}$. Squaring these values results in the estimator $\tilde{\lambda}_k$ of the eigenvalues.

Step 3. *Recover the R covariance estimator* $\tilde{G}(x, x')$.
Define $\tilde{G}(x, x') = \sum_{k=1}^{N} \tilde{\lambda}_k \tilde{\varphi}_k(x)\tilde{\varphi}_k(x')$, $x, x' \in [0, 1]$.

3.5 The RSCB for the difference of two mean functions

The framework proposed here to obtain a RSCB for the mean function can be extended to obtain a RSCB for the difference of the mean functions of two populations. Denote

for $d = 1, 2$ the samples coming from each population, satisfying the model defined in (3)

$$Y_{dij} = m_d\left(j/N\right) + \sum_{k=1}^{\kappa_d} \xi_{dik}\phi_{dk}\left(j/N\right) + \sigma_d\left(j/N\right)\varepsilon_{dij}, \quad 1 \le i \le n_d,\ 1 \le j \le N,$$

with covariance function $G_d(x, x') = \sum_{k=1}^{\kappa_d} \phi_{dk}(x)\phi_{dk}(x')$, respectively. Define the ratio of two-sample sizes as $\hat{r} = n_1/n_2$ and assume that $\lim_{n_1\to\infty} \hat{r} = r > 0$. For each group, we can obtain the R-estimator for the mean function as described in Sect. 3.2.

Following the procedure in Sect. 3.4, we can obtain the R-estimators for the covariance function of each group, $\tilde{G}_d(\cdot,\cdot)$, $d = 1, 2$, then proceed as Sect. 3.3, by first defining $\zeta_{12}(x)$, $x \in [0, 1]$ the Gaussian process with zero mean, $\mathbb{E}[\zeta_{12}(x)] = 0$, unit variance $\mathbb{E}[\zeta_{12}^2(x)] = 1$, and covariance function

$$\mathbb{E}\left[\zeta_{12}(x)\zeta_{12}(x')\right] = \frac{\tilde{G}_1(x, x') + r\tilde{G}_2(x, x')}{\left\{\tilde{G}_1(x, x) + r\tilde{G}_2(x, x)\right\}^{1/2}\left\{\tilde{G}_1(x', x') + r\tilde{G}_2(x', x')\right\}^{1/2}},$$

where $x, x' \in [0, 1]$. Analogue to the one-sample case, the quantile $\hat{Q}_{12,1-\alpha}$ can be estimated using simulations. The RSCB for $m_1(x) - m_2(x)$ is given as

$$(\tilde{m}_1(x) - \tilde{m}_2(x)) \pm c(n_1, n_2) \cdot n_1^{1/2}\left[\tilde{G}_1(x, x) + \hat{r}\tilde{G}_2(x, x)\right]^{1/2}\hat{Q}_{12,1-\alpha},$$

where the correction factor is selected as $c(n_1, n_2) = 1 + \left(1 + \min(n_1, n_2)^{0.20}\right)^{-1}$, similarly to the discussion in Sect. 4.2. The confidence band for the difference of the mean functions can be used to perform a hypothesis test of the form $H_0 : m_1(x) \equiv m_2(x)$, $\forall x \in [0, 1]$ versus $H_A : m_1(x) \neq m_2(x), \exists x \in [0, 1]$. The test can be performed by calculating the appropriate $(1-\alpha) \times 100\%$ confidence band. Although the p-value cannot be calculated directly, it can be estimated by finding the largest α when H_0 is rejected.

4 Simulation

In this section, we perform a simulation study to select the correction factor and compare the performance of the proposed RSCB with the (non-robust) method proposed by Cao et al. (2012b) for the mean function and the difference of mean functions for two populations. We use empirical confidence band coverage rate as a performance criterion. Since outlier curves often have different types of outlying behaviors in the functional data setting, we consider several types of outliers in the assessment of the performance of the RSCB. In Appendix, we also provide the comparison of SCB

and RSCB under different mean/covariance functions setting and additional outlier types.

4.1 Simulation setting

We first generate data from the simulation model in Cao et al. (2012b), i.e.,

$$Y_{ij} = m\,(j/N) + \sum_{k=1}^{2} \xi_{ik}\phi_k\,(j/N) + \sigma\epsilon_{ij}, \quad 1 \le j \le N,\ 1 \le i \le n.$$

In this model, ξ_{ik} for $k = 1, 2$ and ϵ_{ij} for $1 \le j \le N,\ 1 \le i \le n$ are generated from $N(0, 1)$. The number of subjects is n and the number of observations per curve is taken as $N = n$ for functional samples. The mean function, eigenvector functions and the noise level are taken as $m(x) = 10+\sin\{2\pi(x-1/2)\}$, $\phi_1(x) = -2\cos\{(\pi(x-1/2)\}$, $\phi_2(x) = \sin\{\pi(x - 1/2)\}$, and $\sigma = 0.5$, respectively. Note that this model implies that $\lambda_1 = 2$, $\lambda_2 = 0.5$, but this information is not used a priori.

Under this functional model, we introduce outlier curves (Y^o_{ij}) to the generated functional sample by contaminating a subset, I_O, of the original functional sample. The contamination proportion varies from 0 to 0.5, at 0.05 increment. In order to determine the influence of different types of outliers on constructing SCB for the mean function, we consider three different types of outliers, mainly with localized influence, which mimic the types of outlying curves encountered in the real data sets in Sect. 5.

1. *Peak outliers* To simulate an outlier with a punctual influence, each curve was contaminated at a single measurement point, j^*/N, by adding a fixed value s (outliers strength), that is,

$$Y^o_{ij^*} = Y_{ij^*} + s, i \in I_O, \quad j^* = \lfloor 0.05N \rfloor.$$

   This produces a peak outlier curve with a *peak* at the point j^*/N.
2. *Bump outliers* This type is an extension of the peak outliers and the contamination occurs in an interval, $[b_0, b_1]$, rather than at a single point, that is,

$$Y^o_{ij^*} = Y_{ij^*} + s, \quad i \in I_O, \quad j^*/N \in [b_0, b_1].$$

   This type of outlier is present in the ozone data set considered in Sect. 5. In the simulation, the interval is chosen as $[b_0, b_1] = [0.5, 0.53]$.
3. *Step outliers* A further extension of the bump outliers is created by contaminating the curve in the interval $[c_i, 1]$, where c_i is randomly chosen from $[0.5, 1]$ for each outlying curve, that is,

$$Y^o_{ij^*} = Y_{ij^*} + s, \quad i \in I_O, \quad j^*/N \ge c_i.$$

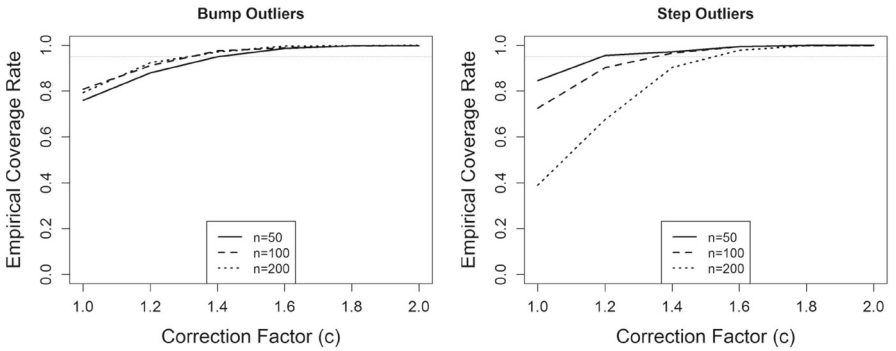

Fig. 1 Empirical coverage rates of 95% RSCB when the correction factor c varies between 1 and 2

4.2 Finding correction factor

The motivation of applying correction factor is to improve the consistency of RSCB, as we find the estimator $\tilde{G}$ tends to underestimate the true value for small sample sizes. We follow the idea introduced in Pison et al. (2002) to find an appropriate correction factor based on simulation. In this simulation, we set the contamination ratio of the outliers as 20% and the outliers strength as $s = 5$. Three sample sizes $n = 50, 100$ and 200 are considered and the correction factor c is varied for values between 1 and 2. For each n, the validity of the confidence band is tested at 100 points $\{1/100, \ldots, 99/100, 1\}$, against the value of the true mean function defined as in the simulation setting. The proportion of the valid confidence bands is then evaluated. The simulation is repeated 500 times. Figure 1 presents the empirical coverage rates of 95% RSCB, i.e., $\tilde{m}(x) \pm c \cdot n^{-1/2} \tilde{G}(x, x)^{1/2} \hat{Q}_{1-\alpha}$, for the cases of bump and step outliers. Using the empirical coverage rates for various c, we select $c(n) = 1 + (1 + n^{0.20})^{-1}$ as the correction factor to improve the empirical coverage rates. Notice that $c(n)$ is greater than 1, which helps to improve consistency for small sample cases, and has a diminishing influence as the sample size increases, i.e., $\lim_{n\to\infty} c(n) = 1$. Although an exhaustive study of all possible sample sizes and outliers is impossible, additional simulation results in Appendix have also shown that this selection of correction factor performed well under different simulation settings and outliers.

4.3 An illustrative example

We first show the performance of the proposed method for one-sample SCB construction on an illustrative toy example. Therefore, we generate a functional sample of $n = 100$ from the model defined in Sect. 4.1 and contaminate the data by using all three types of outliers with the contamination proportion 20% and, for this illustrative example, a strength of $s = 20$. We first construct the 95% confidence band using the proposed RSCB (black) and non-robust (red) SCB for the mean function for each outlier type. We also construct the non-robust SCB and RSCB for the mean function when the sample does not have outlier curves to assess the consistency of the proposed

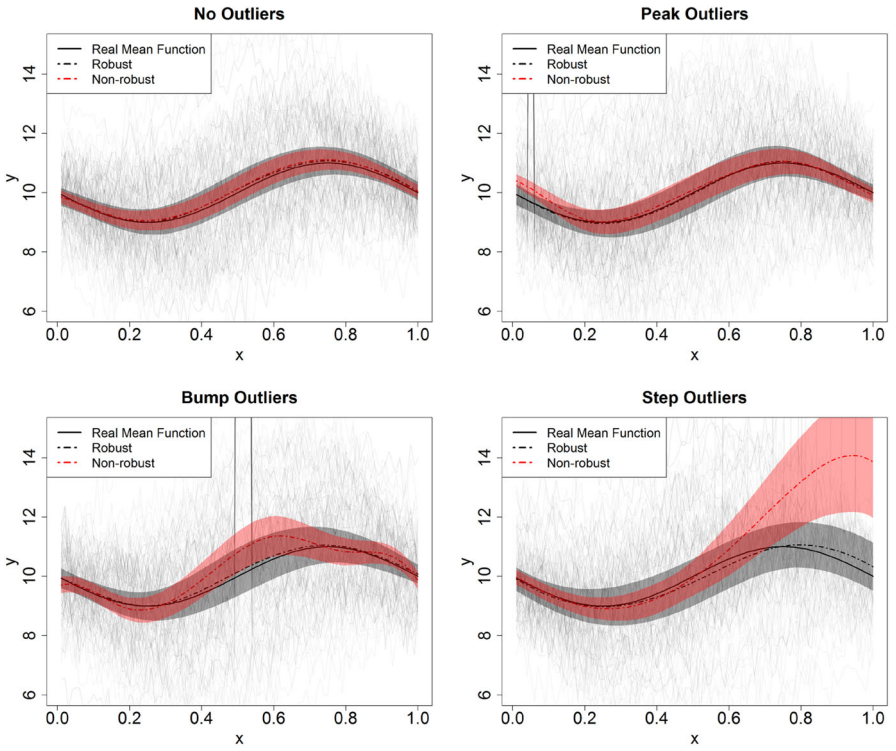

Fig. 2 Comparison between RSCB (black) and non-robust SCB (red) for a simulated dataset (color figure online)

RSCB. Figure 2 depicts the effects of outlier curves on the non-robust SCB and RSCB methods.

The first graph (top left in Fig. 2) for no outlier case shows that the proposed RSCB behaves the same as the non-robust SCB when there is no outlier curve in the data; therefore, it is consistent. For peak and bump outliers (top right and bottom left in Fig. 2), which are considered as very localized outliers, the width of the non-robust SCB is widened around the outlier location and the estimate of the true mean function is strongly affected, deviating from the true mean function, which results in a confidence band that does not cover the mean function, whereas the RSCB is robust to outlier curves even though it is widened slightly around the outlier location, but by a small factor, and the estimated mean is barely affected by the presence of these outliers.

For the step outlier (bottom right in Fig. 2), the mean function estimation and the non-robust SCB are affected dramatically since the estimate of the mean function is deteriorated, whereas the RSCB is not affected even though the band is widened a little around the outlier location. This illustrative example provides evidence that when there are outlier curves in a functional dataset, estimate of the mean function and non-robust SCB are both affected badly, while the proposed RSCB based on the R mean estimator performs well for different types of outlier curves. Further, the proposed RSCB is consistent since it performs well when there are no outlier curves.

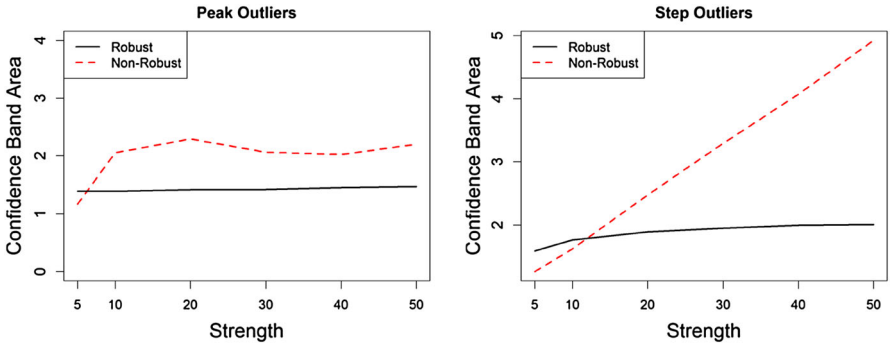

Fig. 3 Comparison of the area of the RSCB (black) and non-robust SCB (red), for varying strengths of outliers (color figure online)

Another important metric is the relationship between the strength of the outlier curves and the total area of the confidence band. To illustrate this metric, we generate functional samples of size $n = 50$ from the model defined in Sect. 4.1 and contaminate the data by using peak and step outliers with contamination proportion 20%, and the varying strengths, $s = 5, 10, 20, 30, 40, 50$. The results are presented in Fig. 3. Note that the area of non-robust SCB is strongly influenced by the strength of the outliers, while the area of the RSCB is maintained at an almost constant level. Similar results are obtained for the bump outliers, which are not shown here.

4.4 Simulation for the SCB for the mean function and the difference of two mean functions

Case I: SCB for the mean function, $m(x)$

To evaluate the performance of the proposed RSCB method for the mean function, we calculate the empirical coverage rate. We generate functional samples from the model in Sect. 4.1 for sample sizes $n = 30, 50$, and 100 with strength is $s = 5$. Each simulation is repeated 500 times.

The empirical coverage rates for contamination proportions varying from 0.05 to 0.50 are presented in Fig. A.1 in Supplementary file as well as in Table 1. The results for non-contaminated datasets are presented in Table 2. For the first two types of outliers (top two graphs in Fig. A.1 and Table 1), the proposed RSCB method maintains the nominal level 95% and has breakdown point close to 50%, while the non-robust SCB breaks down immediately in the presence of outliers and shows a rapid decrease in the empirical coverage rate. Hence, we conclude that proposed RSCB performs superiorly to the non-robust SCB for these two cases.

For the step outlier case, the RSCB method still performs well (bottom graph in Fig. A.1 and Table 1). Although the RSCB method does not maintain 50% breakdown point as in the case of peak and bump outliers, but still has reasonably good breakdown point (20% to 30%), whereas the non-robust SCB breaks down immediately in the presence of even small contamination proportions.

Table 1 Comparison of robust (R) and non-robust (NR) empirical coverage rates of 95% SCB for four types of outlier curves

Outlier type	n	Method	Contamination proportion									
			0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
Peak	30	R	0.91	0.92	0.92	0.93	0.91	0.90	0.91	0.94	0.93	0.94
		NR	0.71	0.44	0.26	0.08	0.05	0.04	0.03	0.01	0.02	0.01
	50	R	0.89	0.92	0.92	0.92	0.91	0.88	0.89	0.92	0.93	0.89
		NR	0.77	0.20	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	100	R	0.93	0.93	0.95	0.94	0.94	0.92	0.93	0.91	0.88	0.87
		NR	0.90	0.81	0.70	0.52	0.32	0.17	0.07	0.02	0.00	0.00
	200	R	0.94	0.97	0.94	0.94	0.96	0.96	0.95	0.95	0.91	0.92
		NR	0.93	0.90	0.85	0.77	0.62	0.52	0.34	0.22	0.12	0.02
Bump	30	R	0.93	0.91	0.92	0.88	0.87	0.92	0.95	0.93	0.93	0.94
		NR	0.81	0.70	0.56	0.27	0.13	0.11	0.14	0.13	0.12	0.06
	50	R	0.92	0.92	0.88	0.87	0.88	0.87	0.90	0.88	0.91	0.91
		NR	0.86	0.68	0.50	0.14	0.03	0.00	0.00	0.00	0.00	0.00
	100	R	0.93	0.94	0.91	0.89	0.86	0.87	0.90	0.88	0.89	0.85
		NR	0.80	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	200	R	0.94	0.93	0.90	0.81	0.71	0.67	0.72	0.70	0.59	0.63
		NR	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Step	30	R	0.93	0.92	0.95	0.96	0.98	0.99	0.99	0.99	1.00	0.82
		NR	0.81	0.83	0.67	0.17	0.02	0.00	0.00	0.00	0.00	0.00
	50	R	0.91	0.96	0.96	0.98	0.97	0.98	0.99	1.00	0.98	0.84
		NR	0.91	0.59	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	100	R	0.93	0.94	0.94	0.92	0.93	0.91	0.88	0.86	0.79	0.35
		NR	0.73	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	200	R	0.94	0.90	0.81	0.66	0.54	0.31	0.19	0.07	0.02	0.00
		NR	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 2 Comparison of robust (R) and non-robust (NR) empirical coverage rates of 95% SCB for datasets with no outliers

n	30	50	100	200
R	0.93	0.93	0.96	0.96
NR	0.81	0.92	0.93	0.94

Case II: SCB for the difference of two mean functions, $m_1(x)-m_2(x)$

We also conduct simulation to evaluate the performance of the RSCB method for the difference between two mean functions, by testing the hypotheses described in Sect. 3.5,

$$H_0 : m_1(x) \equiv m_2(x), \quad \forall x \in [0, 1] \text{ versus } H_A : m_1(x) \neq m_2(x), \quad \exists\, x \in [0, 1]. \tag{7}$$

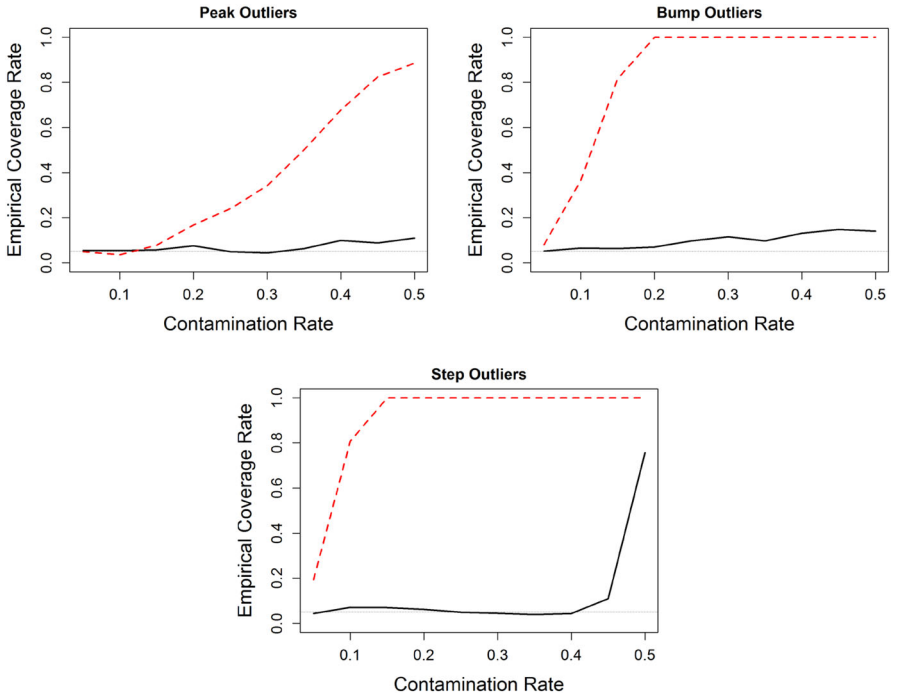

Fig. 4 Empirical type I error of hypothesis test for the difference of means for two populations, with nominal value $\alpha = 0.05$. Non-robust method (red) versus robust method (black) (color figure online)

We employ the same model in Sect. 4.1 for the one sample case. In this simulation setup, $n_1 = 100$, and $n_2 = 130$ correspond to the sample sizes for the first and the second population, respectively, $N = 100$ are the number of measurement points for both samples, and outlier curves are introduced to the first population.

The results of the simulation are presented in Fig. 4 for the three types of outliers, using RSCB (red) and using non-robust SCB (black) as proposed in Cao et al. (2012b). For all cases, the type I error is not maintained for the non-robust method, while for the RSCB method, the type I error is kept at, or close to the nominal value. For the step outliers, the type I error is close to the nominal value for small contamination proportions, deteriorating for large contamination proportions, albeit at a much slower rate than the non-robust SCB. This is evidence that our proposed RSCB method has a superior performance compared to the classical one.

5 Applications

We illustrate our approach on two datasets: octane dataset for the one-sample case and ground-level ozone concentration dataset for the two-sample case.

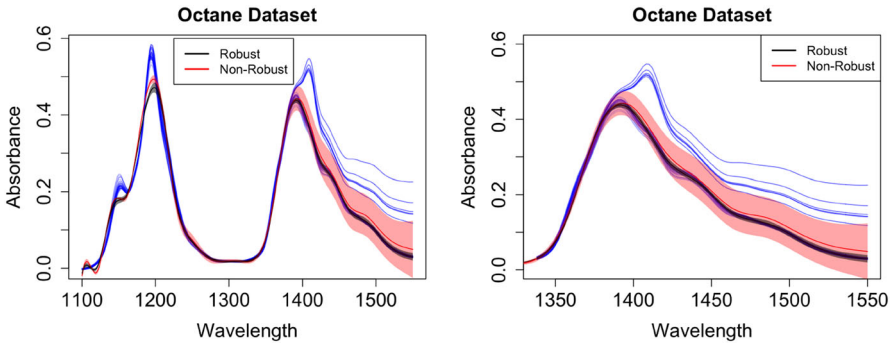

Fig. 5 95% SCB comparison for the octane dataset. Non-robust (red) versus robust (black) methods. Left: full spectrum. Right: magnified on the second half of the spectrum (color figure online)

5.1 Octane dataset

This dataset consists of 39 near-infrared (NIR) spectra of gasoline sample, obtained from Esbensen et al. (1994). It is known that six of the samples contain added ethanol, which corresponds to an upward translation on the upper wavelength, 1390 onward, interval of the spectrum. This is considered as the step outliers described in Sect. 4.

The R estimation of the mean and the 95% RSCB are calculated for this dataset, as well as the mean estimator and confidence band using the method in Cao et al. (2012b). The results are presented in Fig. 5, showing the full spectrum measure (left panel) and magnified on the second half of the spectrum to display the differences more apparently between the non-robust and robust SCBs (right panel).

We observe that the R mean estimator remains close to the non-outlying curves, while the non-robust estimate of the mean function is heavily influenced by the outliers, resulting in an upward shift. The non-robust SCB is also heavily influenced by the outliers, translating in a vast wide band on the second half of the spectrum. However, the proposed RSCB maintains a consistent width across the whole spectrum.

5.2 Ground-level ozone concentration dataset

This dataset consists of hourly average measurements of ground-level ozone (O_3) concentrations from a monitoring station in Richmond, BC, Canada, from the years of 2004 to 2012. The presence of ozone at ground level is highly undesirable and considered a serious air pollutant. Since the concentration of ground-level ozone typically peaks at summer months, only the month of August is analyzed, resulting in 31 samples, with 24 measurement points for each sample.

The same dataset was studied in Boente and Salibian-Barrera (2015), and using S-estimators for the principal components, the presence of outliers was detected in the year of 2005. For illustrative purposes, we take the ozone levels for the year of 2005 as one sample and the ozone levels for the year of 2007, which we know there are no outlier curves (Boente and Salibian-Barrera 2015), as the other sample. The plot of the ground level O_3 concentration for years 2005 and 2007 is presented in Fig. 6, top

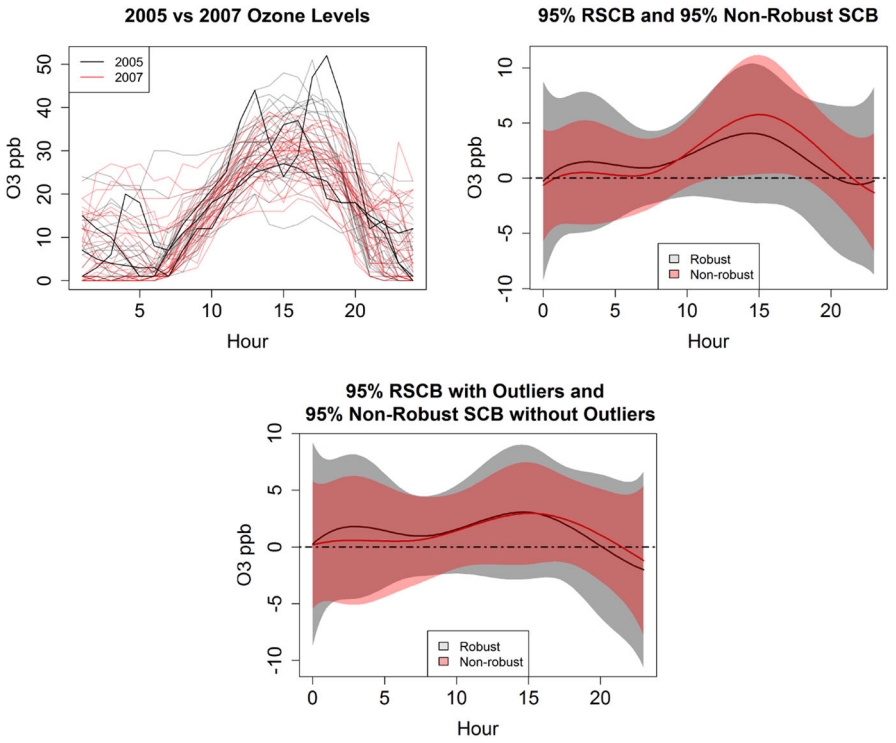

Fig. 6 Top left: O_3 levels in years of 2005 (gray and black) and 2007 (red) in Richmond. Black lines are the outliers. Top right: 95% non-robust SCB (red) and RSCB (gray) for the difference between the mean functions of the two years. Bottom: 95% non-robust SCB (red) and RSCB (gray) for the difference between the mean functions of the two years, keeping outliers for RSCB, excluding outliers for non-robust SCB (color figure online)

left panel, with the year of 2005 in gray/black and the year of 2007 in red. The outliers detected by Boente and Salibian-Barrera (2015) are highlighted.

We set up our hypotheses for testing whether there is a difference between the ozone mean functions of the years 2005 and 2007 in Richmond, Canada. The outliers in the dataset are similar to the bump outliers described in Sect. 4. The ground level O_3 concentration remained the same, except for the aforementioned outliers in 2005 (Fig. 6, top left panel). The R method does not reject the null hypothesis at a significant level $\alpha = 0.05$, while the non-robust method proposed in Cao et al. (2012b) rejects the null hypothesis, with an empirical p value calculated as 0.015. The 95% SCB of the difference between the mean functions of the ground level O_3 concentration in years of 2005 and 2007 is presented in the top right panel of Fig. 6. We also calculate the 95% SCB for the difference between the mean functions with the outliers kept for the RSCB and excluding the outliers for the non-robust SCB. This is presented in the bottom panel of Fig. 6. This plot provides a comparison of the SCB between the robust and non-robust methods, highlighting that the former successfully works, despite the presence of outliers.

6 Discussion

In this paper, we proposed a novel method for robustly estimating the mean function of the functional data and a method to obtain an RSCB for the mean function with the help of a robust estimator for the covariance function. To the best of our knowledge, this work is the first manuscript investigating the construction of RSCB accounting for the presence of outlying curves. Although we primarily focus on the computational issue of RSCB, we also proved the consistency of the proposed LAD estimator for the mean function when the measurement errors have a symmetric distribution. Moreover, we obtained a finite-sample correction factor that results in consistent coverage rates for the RSCB for small sample sizes. Although an exhaustive study of all possible sample sizes and outliers is impossible, we conducted a simulation study to determine the correction factor for three additional types of outliers. We showed that this selection of correction factor performed well under different simulation settings and outliers.

Further, we consider that our contribution to software development for computing and visualizing SCB and RSCB with functional data is significant. The R-codes written for the non-robust and robust SCB in this manuscript which can be downloaded at http://auburn.edu/~gzc0009/software.html will be instrumental in applied research, and these SCB methods will provide a strong visualization tool for interdisciplinary audiences.

As suggested by one referee, we may consider model (1) under elliptical processes setting. Let η be elliptical with location m and scatter operator G, $\eta(x) = m(x) + \sum_{k=1}^{\infty} \xi_k \phi_k(x)$, where for any k the random vector $(\xi_1, \ldots, \xi_k)^{\mathrm{T}}$ has an spherical distribution. Furthermore, the scale function $G = \sum_{k=1}^{\infty} \phi_k \otimes \phi_k$ with $\phi_k(x) = \sqrt{\lambda_k}\varphi_k(x)$ and $\lambda_1 \geq \lambda_2 \geq \cdots$ and $\{\varphi_k(x)\}_{k=1}^{\infty}$ are the eigenvalues and eigenfunctions. By considering elliptical processes setting, it allows to identify location m without requiring any moment conditions. The interested readers are referred to Boente et al. (2014). We also notice that $\tilde{m}(x)$ does not have the same asymptotic variance as the estimators considered in Cao et al. (2012b) and $\tilde{G}$ is not a consistent estimator, so more careful and intensive work is necessary to investigate the asymptotic covariance of LAD estimator $\tilde{m}(x)$. We leave these issues for future work.

References

Adrover J, Salibian-Barrera M, Zamar R (2004) Globally robust inference for the location and simple linear regression models. J Stat Plan Inference 119(2):353–375

Aneiros G, Bongiorno EG, Cao R, Vieu P (2017) Functional statistics and related fields. Springer series in statistics. Springer, New York

Bali JL, Boente G, Tyler DE, Wang JL (2011) Robust functional principal components: a projection-pursuit approach. Ann Stat 39(6):2852–2882

Boente G, Salibian-Barrera M (2015) S-estimators for functional principal component analysis. J Am Stat Assoc 110(511):1100–1111

Boente G, Barrera MS, Tyler DE (2014) A characterization of elliptical distributions and some optimality properties of principal components for functional data. J Multivar Anal 131:254–264

Cao G, Wang J, Wang L, Todem D (2012a) Spline confidence bands for functional derivatives. J Stat Plan Inference 142(6):1557–1570

Cao G, Yang L, Todem D (2012b) Simultaneous inference for the mean function based on dense functional data. J Nonparametric Stat 24(2):359–377

Cardot H, Cénac P, Zitt PA et al (2013) Efficient and fast estimation of the geometric median in hilbert spaces with an averaged stochastic gradient algorithm. Bernoulli 19(1):18–43
Cuevas A (2014) A partial overview of the theory of statistics with functional data. J Stat Plann Inference 147:1–23
Cuevas A, Febrero M, Fraiman R (2007) Robust estimation and classification for functional data via projection-based depth notions. Comput Stat 22(3):481–496
de Boor C (2001) A practical guide to splines. Springer, New York
Du Mond CE, Lenth RV (1987) A robust confidence interval for location. Technometrics 29(2):211–219
Embling CB, Illian J, Armstrong E, van der Kooij J, Sharples J, Camphuysen KC, Scott BE (2012) Investigating fine-scale spatio-temporal predator-prey patterns in dynamic marine ecosystems: a functional data analysis approach. J Appl Ecol 49(2):481–492
Esbensen K, Schönkopf S, Midtgaard T, Guyot D (1994) Multivariate analysis in practice: a training package. Camo, Trondheim
Febrero M, Galeano P, González-Manteiga W (2008) Outlier detection in functional data by depth measures, with application to identify abnormal nox levels. Environmetrics 19(4):331–345
Ferraty F, Vieu P (2006) Nonparametric functional data analysis: theory and practice. Springer series in statistics. Springer, New York
Ferraty F, Rabhi A, Vieu P (2005) Conditional quantiles for dependent functional data with application to the climatic "el niño" phenomenon. Sankhyā Indian J Stat 67(2):378–398
Fraiman R, Muniz G (2001) Trimmed means for functional data. Test 10(2):419–440
Fraiman R, Yohai VJ, Zamar RH (2001) Optimal robust M-estimates of location. Ann Stat 29(1):194–223
Gervini D (2008) Robust functional estimation using the median and spherical principal components. Biometrika 95(3):587–600
Goia A, Philippe V (2016) An introduction to recent advances in high/infinite dimensional statistics. J Multivar Anal 146:1–6
Horváth L, Kokoszka P (2012) Inference for functional data with applications. Springer series in statistics. Springer, New York
Hsing T, Eubank R (2015) Theoretical foundations of functional data analysis, with an introduction to linear operators. Wiley series in probability and statistics. Wiley, Chichester
Huber P, Ronchetti E (2009) Robust statistics. Wiley series in probability and statistics. Wiley, Hoboken
Kraus D, Panaretos VM (2012) Dispersion operators and resistant second-order functional data analysis. Biometrika 99(4):813–832
Lee S, Shin H, Billor N (2013) M-type smoothing spline estimators for principal functions. Comput Stat Data Anal 66:89–100
Locantore N, Marron J, Simpson D, Tripoli N, Zhang J, Cohen K, Boente G, Fraiman R, Brumback B, Croux C et al (1999) Robust principal component analysis for functional data. Test 8(1):1–73
López-Pintado S, Romo J (2009) On the concept of depth for functional data. J Am Stat Assoc 104(486):718–734
Maronna RA, Yohai VJ (2013) Robust functional linear regression based on splines. Comput Stat Data Anal 65:46–55
Pison G, Van Aelst S, Willems G (2002) Small sample corrections for lts and mcd. Metrika 55(1):111–123
Ramsay JO, Silverman B (2005) Functional data analysis, 2nd edn. Springer, New York
Shin H, Lee S (2016) An RKHS approach to robust functional linear regression. Stat Sin 26:255–272
Wang HJ, Zhu Z, Zhou J (2009) Quantile regression in partially linear varying coefficient models. Ann Stat 37(6b):3841–3866
Wilcox RR (2005) Introduction to robust estimation and hypothesis testing. Academic Press, New York
Yao F, Müller HG, Wang JL (2005) Functional data analysis for sparse longitudinal data. J Am Stat Assoc 100(470):577–590