Statistica Sinica Preprint No: SS-2017-0262	
Title	Hypothesis Testing for Multiple Mean and Correlation Curves with Functional Data
Manuscript ID	SS-2017-0262
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202017.0262

Complete List of Authors	Ao Yuan Hong-Bin Fang Haiou Li Colin O. Wu and Ming T. Tan
Corresponding Author	Ao Yuan
E-mail	yuanao@hotmail.com

Notice: Accepted version subject to English editing.

Hypothesis Testing for Multiple Mean and Correlation Curves with Functional Data

Ao Yuan[1], Hong-Bin Fang[1], Haiou Li[1], Colin O. Wu[2], Ming T. Tan[1,*]

[1]Department of Biostatistics, Bioinformatics and Biomathematics, Georgetown University Medical Center, Washington DC 20057 USA

[2]Office of Biostatistics Research, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892 USA

∗Correspondence: Ming T. Tan, mtt34@georgetown.edu

Abstract

Existing statistical methods of functional data analysis have mostly focused on using local smoothing estimators or some known basis approximations. In many applications with functional observations, the main objectives of statistical inferences are to test (a) the equivalence of a set of unknown mean curves, and (b) the correlation curves of the unknown stochastic processes, while the unknown curves of the functional data may not be approximated by a set of known basis functions. We propose a class of simple to use test statistics, for comparing mean and correlation curves of functional data without relying on the estimation with local smoothing and basis approximations, and studied their basic asymptotic properties. Through an application to the functional gene expression data, we show that our method gives practically meaningful results with minimal assumptions. Numerical justifications of our testing method is provided by a simulation study.

Some keywords: Comparing mean curves, correlation curve, functional data, gene expression profile, testing functional equivalence, uniform normed convergence.

1 Introduction

Functional data analysis with the objectives of evaluating multiple mean and correlation curves over time has been extensively used in biomedical studies. Nonparametric analysis with functional data include estimation and hypothesis testing of the unknown curves without relying on the potentially unrealistic parametric assumptions. A popular approach for the nonparametric inferences with functional data is to assume that the unknown curves belong to a "structured regression model" (Hastie, Tibshirani and Friedman, 2009), which can be approximated by some expansions of a class of known basis functions, so that the estimation and testing procedures can be constructed through the unknown coefficients of the basis expansions. Various basis approximation methods can be found in Shi, Weiss and Taylor (1996), Rice and Wu (2001), Huang, Wu and Zhou (2002), Müller and Yao (2008), and Li et al. (2010), among others. When the objective is only nonparametric estimation of the unknown curves, local smoothing methods, such as the local polynomials or smoothing splines are often used in conjunction with some structured regression models, e.g., Hoover et al. (1998), Wu and Chiang (2000) and Chiang, Rice and Wu (2001) and Fan and Zhang (2000). Cai and Yuan (2011) studied optimal convergence rate of general estimators and penalized smoothing estimator of mean function. Kim and Zhao (2013) introduced two self-normalization methods to overcome kernel smooth convergence issue for the sparse and dense longitudinal model. Local polynomials and smoothing splines are often used to estimate functional curves. With our best knowledge, the statistical properties in hypothesis testing with functional data have little been investigated.

In many biological studies, biological characteristics have been measured at various time points. the scientific question is how to compare the different curves of various biological characteristics and their correlations, while there is no *a priori* knowledge of

the suitable basis functions to approximate the curves. For example, a main objective in gene temporal expression study is to test whether different genes have the same mean expression profiles or some gene expression profiles are correlated.

Two sample hypothesis testing for functional data has been approached in many aspects. To test the differences in the mean functions, Zhang and Chen (2007) applied local polynomial kernel smoothing technique and constructed a L^2-norm-based global test statistic. Cao et al. (2012) constructed polynomial spline confidence band for mean curves and Degras (2011, 2017) constructed simultaneous confidence bands. Hypothesis test problem was address by testing if the bands for the difference contained zero function. Wang et al. (2017) considered unified empirical likelihood ratio tests based on functional concurrent linear model. Extension to k independent samples of curves were discussed in Cuevas et al. (2004) by introducing ANOVA-like test.

In order to test the equality of the distribution of two sets of curves, Hall and van Keilegom (2007) proposed a Cramer-von Mises (CVM)-type of test based on a second-order smoother. Estévez-Pérez & Vilar (2008) extended the method to k-sample cases and applied it to air quality data. Benko et al. (2009) studied this problem by first introducing the functional principal components decomposition along with bootstrap. Pomann et al. (2016) also decomposed the curves by functional principle component analysis (FPCA) and applied Anderson-Darling statistic. Some researches focused on detecting differences in the covariance structures of curves. Gaines et al. (2011) used a likelihood ratio-type approach. Fremdt et al. (2012) developed a chi-square asymptotic test statistic based on FPCA. Regularized M-test based on Hilbert-Schmidt norm was introduced by Kraus and Panaretos (2012). Horváth and Kokoszka (2012) describes an estimation and testing approach for mean and covariance curves based on the assumption that the subjects are observed over a very dense set of time points, hence, statistical inferences could be constructed as if the entire curves were observed. This assumption,

again, is unrealistic in most real studies. The above methods can surely be applied in practice with moderately dense functional data, like the method proposed here, both will produce inaccurate inferences if the observation grid is not sufficiently dense.

We study in this paper the problem of testing the equivalence of mean curves or correlation curves under the practical situation that the curves are observed at a set of time design points chosen based on the scientific objectives and the estimated curves can be constructed through interpolations at the observed time points. As pointed out by Cai and Yuan (2011, p.2332) that for estimation of mean function, smoothing (i.e. more complicated method) does not result in improved convergence rate, so we propose the simple method of empirical linear interpolation mean as the estimator of mean curve, for testing the difference in mean and for correlation functions between two curves. The asymptotic results of our testing procedures are more general in the stronger uniform sense than the comparable ones given in Horváth and Kokoszka (2012). The main advantage of the proposed method is a) existing methods first transform the observations into smoothed curves, take these curves as functional data, thus the results deviate more or less from the truth; while we directly use the observed raw data without first smoothing them, this is of particular meaning when the number of observation time points is not big, as for our problem. b) The proposed method uses the empirical mean functions to test the null hypothesis, makes it very simple to use. In comparison, most existing methods use splines, basis functions, reproducing kernels, kernel smoothers, etc. which are not as simple to use, and require more assumptions. The main limitation of this method is that it may not be able to used in some other problems outside the mean, such as functional regression.

We describe the data structure and our testing procedures in Section 2, and investigate the asymptotic properties of our proposed test statistics in Section 3. Finite sample properties of our test statistics are investigated through the temporal gene expression

study in Section 4 and a simulation study in Section 5. We conclude with a discussion in Section 6 and outline proofs of the main results in the Appendix in a separate online file.

2 Data Structure and Hypothesis Testing

2.1 Data Structure

Functional data can be viewed as observed stochastic processes on the real line or random functions in some functional spaces (e.g., Ramsay and Silverman, 2005). For a bivariate case of functional data, we consider stochastic processes $\{\big(X(t), Y(t)\big)^T : t \in \mathcal{T}\}$, where, given $t \in \mathcal{T}$, a bounded subset in $[0, \infty)$, $X(t)$ and $Y(t)$ are real valued random variables, which may be correlated. For each fixed $t \in \mathcal{T}$, let $\mu(t) = E[X(t)]$ and $\eta(t) = E[Y(t)]$ be the mean curves of $X(t)$ and $Y(t)$, respectively, and

$$R(t) = \frac{E\big\{[X(t) - \mu(t)]\,[Y(t) - \eta(t)]\big\}}{\sigma_{X(t)}\sigma_{Y(t)}} \tag{1}$$

be the correlation curve of $X(t)$ and $Y(t)$, where $\sigma_{X(t)}$ and $\sigma_{Y(t)}$ are the corresponding standard deviations of $X(t)$ and $Y(t)$. As t changes within $\mathcal{T}$, $\mu(t)$, $\eta(t)$, $\sigma_{X(t)}$, $\sigma_{Y(t)}$ and $R(t)$ are the curves for the means, standard deviations and correlation coefficient over $t \in \mathcal{T}$.

Since, in real applications, the subjects are assumed to be independent, but some subjects may have only $X(\cdot)$ or $Y(\cdot)$ observed, we denote by $\mathcal{S}_X$ the set of subjects with observations of $X(\cdot)$ only, $\mathcal{S}_Y$ the set of subjects with observations of $Y(\cdot)$ only, and $\mathcal{S}_{XY}$ the set of subjects with observations of $(X(\cdot), Y(\cdot))^T$. Let n_x, n_y and n_{xy} be the numbers of subjects in $\mathcal{S}_X$, $\mathcal{S}_Y$ and $\mathcal{S}_{XY}$, respectively. The number of subjects with observations of $X(\cdot)$, i.e. in $\mathcal{S}_X \cup \mathcal{S}_{XY}$, is $n_1 = n_x + n_{xy}$. The number of subjects with observations of $Y(\cdot)$, i.e., in $\mathcal{S}_Y \cup \mathcal{S}_{XY}$, is $n_2 = n_y + n_{xy}$. The total number of subject

is $n = n_x + n_y + n_{xy} = n_1 + n_2 - n_{xy}$. For simplicity, we assume that the observations of $X(\cdot)$ and $Y(\cdot)$ are made at $k(n)$ distinct time points $t_1 < \cdots < t_{k(n)}$. The results of this paper can be directly generalized to cases of $X(\cdot)$ and $Y(\cdot)$ observed at different time points, which comes at the expense of more complex notation. The observations for $\{X(t) : t \in \mathcal{T}\}$, $\{Y(t) : t \in \mathcal{T}\}$ and $\{(X(t), Y(t))^T : t \in \mathcal{T}\}$ are given by

$$\begin{cases} \mathbb{X}_{n_1} = \{X_{i,j} = X_i(t_j) + \epsilon_i(t_j) : \ i = 1, \ldots, n_1; \ j = 1, \ldots, k(n)\}, \\ \mathbb{Y}_{n_2} = \{Y_{i,j} = Y_i(t_j) + \xi_i(t_j) : \ i = n_x + 1, \ldots, n_x + n_2; \ j = 1, \ldots, k(n)\}, \\ (\mathbb{X}_{n_{xy}}, \mathbb{Y}_{n_{xy}}) = (X_{i,j}, Y_{i,j})^T : i = n_x + 1, \ldots, n_x + n_{xy}; \ j = 1, \ldots, k(n)\}. \end{cases} \tag{2}$$

where the $\epsilon_i(\cdot)$ and $\xi_i(\cdot)$'s are independent measurement noises, they are independent of the $X_i(\cdot)$'s and $Y_i(\cdot)$'s, $\epsilon_i(\cdot)$ iid $\epsilon(\cdot)$ and $\xi_i(\cdot)$ iid $\xi(\cdot)$, with the typical assumption $E[\epsilon_i(t)] = E[\xi_i(t)] = 0$ for all $t \in \mathcal{T}$. Clearly, the subjects in $(\mathbb{X}_{n_{xy}}, \mathbb{Y}_{n_{xy}})$ are the subjects in both $\mathbb{X}_{n_1}$ and $\mathbb{Y}_{n_2}$.

Let $X_i(t)$ and $Y_i(t)$ be the unknown subject-specific curves of $X(t)$ and $Y(t)$, respectively, for the ith subject. To define the curves observations of $X_i(t)$ and $Y_i(t)$ at any $t_1 \le t \le t_{k(n)}$ based on $\{\mathbb{X}_{n_1}, \mathbb{Y}_{n_2}\}$ of (2), we consider the curve observations based on the linear interpolations

$$\begin{cases} \mathcal{X}_{n_1} = \{X_{i,k(n)}(t) : \ i = 1, \ldots, n_1\}, \\ \mathcal{Y}_{n_2} = \{Y_{i,k(n)}(t) : \ i = n_x + 1, \ldots, n_x + n_2\}, \\ (\mathcal{X}_{n_{xy}}, \mathcal{Y}_{n_{xy}}) = \{(X_{i,k(n)}(t), Y_{i,k(n)}(t))^T : \ i = n_x + 1, \ldots, n_x + n_{xy}\}, \end{cases} \tag{3}$$

where $X_{i,k(n)}(t)$ is the linear interpolation based on $X_{i,j}$ defined by

$$\begin{cases} X_{i,k(n)}(t) = X_{i,1}, \quad \text{if } t \le t_1 \text{ and } t \in \mathcal{T}; \\ X_{i,k(n)}(t) = X_{i,k(n)}, \quad \text{if } t \ge t_{k(n)} \text{ and } t \in \mathcal{T}; \\ X_{i,k(n)}(t) = \frac{(t_{j+1} - t)X_{i,j} + (t - t_j)X_{i,j+1}}{t_{j+1} - t_j}, \quad \text{if } t_j \le t \le t_{j+1}, \end{cases} \tag{4}$$

and $Y_{i,k(n)}(t)$ is the linear interpolation based on $Y_{i,j}$ similarly defined as in (4). Then, $X_{i,k(n)}(t)$ and $Y_{i,k(n)}(t)$ are the observed subject-specific curves $X_i(t)$ and $Y_i(t)$. Although

the curve observations can also be constructed using other smoothing methods, such as splines, the linear interpolation is computationally simple and does not depend on the choice of smoothing parameters.

2.2 Curve Estimates

The mean curves $\mu(t)$ and $\eta(t)$ of $X(t)$ and $Y(t)$, respectively, can be naturally estimated by the sample means

$$\hat{\mu}_{n_1}(t) = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{i,k(n)}(t) \quad \text{and} \quad \hat{\eta}_{n_2}(t) = \frac{1}{n_2}\sum_{i=n_x+1}^{n_x+n_2} Y_{i,k(n)}(t). \tag{6}$$

To evaluate the covariance curve of $X(t)$ and $Y(t)$, we need to first consider the moments and their estimators. Let $\mu_{(k,r)}(t) = E\big[X^k(t)Y^r(t)\big]$ and $\mu_{(r_1,r_2,r_3,r_4)}(s,t) = E\big[X^{r_1}(s)X^{r_2}(t)Y^{r_3}(s)Y^{r_4}(t)\big]$ be the moments of $\big(X(t),Y(t)\big)^T$ with non-negative integers r, k and r_l, $l = 1,\ldots,4$. Estimators of $\mu_{(k,r)}(t)$ and $\mu_{(r_1,r_2,r_3,r_4)}(s,t)$ constructed using $\big(X_{i,k(n)}(t), Y_{i,k(n)}(t)\big)^T$ are given by

$$\mu_{n,(k,r)}(t) = \frac{1}{n_{xy}}\sum_{i=n_x+1}^{n_x+n_{xy}} X^r_{i,k(n)}(t)Y^r_{i,k(n)}(t)$$

and

$$\mu_{n,(r_1,r_2,r_3,r_4)}(t) = \frac{1}{n_{xy}}\sum_{i=n_x+1}^{n_x+n_{xy}} X^{r_1}_{i,k(n)}(s)X^{r_2}_{i,k(n)}(t)Y^{r_3}_{i,k(n)}(s)Y^{r_4}_{i,k(n)}(t).$$

By (1), the estimator of $R(t)$ based on $\big(X_{i,k(n)}(t), Y_{i,k(n)}(t)\big)^T$ is

$$R_n(t) = \frac{\mu_{n,(1,1)}(t) - \mu_{n,(1,0)}(t)\mu_{n,(0,1)}(t)}{\sqrt{\big[\mu_{n,(2,0)}(t) - \mu^2_{n,(1,0)}(t)\big]\big[\mu_{n,(0,2)}(t) - \mu^2_{n,(0,1)}(t)\big]}}. \tag{7}$$

Let $\boldsymbol{\mu}(t) = \big(\mu_{(1,0)}(t), \mu_{(0,1)}(t), \mu_{(2,0)}(t), \mu_{(0,2)}(t), \mu_{(1,1)}(t)\big)'$ be the vector of moments. The estimator of $\boldsymbol{\mu}(t)$ is then

$$\boldsymbol{\mu}_n(t) = \big(\mu_{n,(1,0)}(t), \mu_{n,(0,1)}(t), \mu_{n,(2,0)}(t), \mu_{n,(0,2)}(t), \mu_{n,(1,1)}(t)\big)'. \tag{8}$$

The asymptotic distribution of $R_n(t)$ is then evaluated through the asymptotic properties of $\boldsymbol{\mu}_n(t)$.

2.3 Hypotheses and Test Statistics

2.3.1 Testing the Equivalence of Mean Curves

As illustrated in Duan et al. (2003) and Fang et al. (2012), a main scientific question is to test whether the two stochastic processes $\{X(t) : t \in \mathcal{T}\}$ and $\{Y(t) : t \in \mathcal{T}\}$ have the same mean profiles, i.e., $\mu(t) = \eta(t)$, for all $t \in \mathcal{T}$. A well-known approach for testing the equivalence of two mean curve is the Kolmogorov-Smirnov type test statistics $n \sup_{t\in\mathcal{T}} \left|\hat{\mu}_{n_1}(t) - \hat{\eta}_{n_2} n(t)\right|$, which is not robust because the supreme may be influenced by outliers.

We propose in this section two test statistics based on "two-sided" or "one-sided" alternative hypotheses. Our "two-sided" null and alternative hypotheses are

$$H_0 : \mu(t) = \eta(t) \text{ for all } t \in \mathcal{T}, \text{ vs. } H_1 : \mu(t) \neq \eta(t) \text{ on some } \mathcal{A} \subset \mathcal{T}. \tag{9}$$

while, our "one-sided" null and alternative hypotheses are

$$H_0 : \mu(t) = \eta(t) \text{ for all } t \in \mathcal{T}, \text{ vs. } H_1 : \mu(t) > \eta(t) \text{ for all } t \in \mathcal{T}, \tag{10}$$

where $\mathcal{A}$ has positive Lebesgue measure. For testing H_0 vs. H_1 in (9), we use the test statistic,

$$L_n = \frac{1}{|\mathcal{T}|} \frac{n_1 n_2}{n} \int_{\mathcal{T}} \left[\hat{\mu}_{n_1}(t) - \hat{\eta}_{n_2}(t)\right]^2 dt, \tag{11}$$

where $|\mathcal{T}|$ is the length of $\mathcal{T}$. For testing H_0 vs. H_1 in (10), we use the test statistic

$$D_n = \frac{1}{|\mathcal{T}|} \sqrt{\frac{n_1 n_2}{n}} \int_{\mathcal{T}} \left[\hat{\mu}_{n_1}(t) - \hat{\eta}_{n_2}(t)\right] dt. \tag{12}$$

Let $d_n(t_j) = \hat{\mu}_{n_1}(t_j) - \hat{\eta}_{n_2}(t_j)$. Since $\hat{\mu}_{n_1}(t) - \hat{\eta}_{n_2}(t)$ is obtained through a linear interpolation on $T_n = \{t_1, ..., t_{k(n)}\}$,

$$d_n(t) = \frac{(t - t_j) d_n(t_j) + (t_{j+1} - t) d_n(t_{j+1})}{t_{j+1} - t_j}, \qquad t \in [t_j, t_{j+1}].$$

If we define $\hat{\mu}_{n_1}(t) = \hat{\eta}_{n_2}(t) \equiv 0$ for $t \leq t_1$ or $t \geq t_{k(n)}$, then L_n and D_n is simplified as

$$L_n = \frac{1}{3|\mathcal{T}|}\frac{n_1 n_2}{n}\sum_{j=1}^{k(n)-1}\Big(d_n^2(t_{j+1}) + d_n^2(t_j) + d_n(t_j)d_n(t_{j+1})\Big)(t_{j+1} - t_j) \tag{13}$$

and

$$D_n = \frac{1}{|\mathcal{T}|}\sqrt{\frac{n_1 n_2}{n}}\sum_{j=1}^{k(n)-1}\frac{d_n(t_j) + d_n(t_{j+1})}{2}(t_{j+1} - t_j). \tag{14}$$

The approximate critical values for L_n and D_n are derived through their asymptotic distributions in Section 3.

2.3.2 Testing the Correlation function

The second objective is to test whether $X(t)$ and $Y(t)$ are correlated for $t \in \mathcal{T}$ based on the correlation curve $R(t)$ defined in (1). A natural formulation is to test the null hypothesis that $X(t)$ and $Y(t)$ are uncorrelated for all the time points in $\mathcal{T}$ versus the "two-sided" alternative

$$H_0 : R(t) = 0 \ \text{ for all } t \in \mathcal{T}, \text{ vs. } H_1 : R(t) \neq 0 \text{ on some } \mathcal{A} \subset \mathcal{T}, \tag{15}$$

or the uncorrelated $X(t)$ and $Y(t)$ versus the one-sided alternative

$$H_0 : R(t) = 0 \ \text{ for all } t \in \mathcal{T}, \text{ vs. } H_1 : R(t) > 0 \text{ (or } R(t) < 0) \text{ for all } t \in \mathcal{T}. \tag{16}$$

The "one-sided" alternative of (16) suggests that the processes $X(t)$ and $Y(t)$ are always positively (or negatively) correlated for all t in $\mathcal{T}$. In this case we only consider $n = n_1 = n_2 = n_{xy}$, as correlation inference needs paired data sets. Although other cases can also be handled with missing data methods, they are not our main goal here. Based on the empirical correlation curve $R_n(t)$ of (7), in the definition of $\mu_{n,(k,r)}(t)$ and $\mu_{n,(r_1,r_2,r_3,r_4)}(t)$ with the summation sign $\sum_{i=n_x+1}^{n_x+n_{xy}}$ replaced by $\sum_{i=1}^{n}$, we consider the test statistics

$$S_n = \frac{n}{|\mathcal{T}|}\int_{\mathcal{T}} R_n^2(t)dt \ \text{ and } \ W_n = \frac{\sqrt{n}}{|\mathcal{T}|}\int_{\mathcal{T}} R_n(t)dt \tag{17}$$

for the hypotheses in (15) and (16), respectively, where, as in (11), $|\mathcal{T}|$ is the length of $\mathcal{T}$. The approximate critical values for S_n and W_n are derived in Section 3.

3 Asymptotic Distributions of Test Statistics

We develop in this section the asymptotic distributions of the test statistics of Section 2. These asymptotic distributions lead to the critical values of the test statistics. Proofs of the theorems of this section are given in the appendices.

3.1 Asymptotic Distributions of Mean Test Statistics

We first consider the hypotheses (9) and (10) and their testing statistics L_n and D_n as defined in (13) and (14). Let

$$d(t) = \mu(t) - \eta(t) \quad \text{and} \quad d_n(t) = \hat{\mu}_{n_1}(t) - \hat{\eta}_{n_2}(t)$$

be the real and estimated difference of mean curves. Let $R_{xx}(s,t) = Cov\big[X(s), X(t)\big]$, $R_{yy}(s,t) = Cov\big[Y(s), Y(t)\big]$, $R_{xy}(s,t) = Cov\big[X(s), Y(t)\big]$, $R_{yx}(s,t) = Cov\big[Y(s), X(t)\big]$, $R_{\epsilon}(s,t) = Cov\big[\epsilon(s), \epsilon(t)\big]$ and $R_{\xi}(s,t) = Cov\big[\xi(s), \xi(t)\big]$ be the covariance curves of the corresponding stochastic processes. Denote

$$R_{11}(s,t) = R_{xx}(s,t) + R_{\epsilon}(s,t), \qquad R_{12}(s,t) = R_{xy}(s,t),$$

$$R_{21}(s,t) = R_{yx}(s,t), \qquad R_{22}(s,t) = R_{yy}(s,t) + R_{\xi}(s,t).$$

For the asymptotic properties described below, we denote by $\ell^{\infty}(\mathcal{T})$ the space of functions on $\mathcal{T}$ equipped with the supreme norm, and $\overset{D}{\Rightarrow}$ the uniform weak convergence in $\ell^{\infty}(\mathcal{T})$. The following theorem shows that, under some mild conditions, $d_n(t) - d(t)$ can be uniformly weakly approximated by a Gaussian process in the sense that the process $[(n_1 n_2)/n]^{1/2}[d_n(\cdot) - d(\cdot)]$ converges weakly to a Gaussian process. This result of weak convergence as a whole process on $\mathcal{T}$, instead of point-wise convergence on some $t \in \mathcal{T}$, allows us to study the weak limit of the test statistics constructed by the functional of $\widehat{\mu}_n(\cdot)$ and $\widehat{\eta}_n(\cdot)$, such as L_n and D_n. Let $\delta_{k(n)} = \max\{t_{j+1} - t_j : j = 0, 1, ..., k(n)\}$.

Before the statement of Theorem 1 below, we first give an outline of the proof how it works. Note that

$$\begin{aligned}
&\sqrt{\tfrac{n_1 n_2}{n}}\Big\{\big[\widehat{\mu}_{n_1}(t)-\mu(t)\big]-\big[\widehat{\eta}_{n_2}(t)-\eta(t)\big]\Big\}\\
&\quad=\sqrt{\tfrac{n_1 n_2}{n}}\Big\{\big[\widehat{\mu}_{n_1}(t)-\mu_{k(n)}(t)\big]-\big[\widehat{\eta}_{n_2}(t)-\eta_{k(n)}(t)\big]\Big\}\\
&\qquad+\sqrt{\tfrac{n_1 n_2}{n}}\Big\{\big[\mu_{k(n)}(t)-\mu(t)\big]-\big[\eta_{k(n)}(t)-\eta(t)\big]\Big\}.
\end{aligned}$$

We will show that under the given conditions, the second term on the right hand side above is $o(1)$ uniformly in t, deal with the first term using the empirical process theory, check the conditions for uniform Donsker class, and then identify the weak limit process.

Let $\mathcal{P}$ and $\mathcal{Q}$ be the collections of all probability measures (P,Q) of the random processes $(X(\cdot)+\epsilon(\cdot), Y(\cdot)+\eta(\cdot))$. Denote $G=\sup_{t\in\mathcal{T}}\left([X(t)+\epsilon(t)]^2+[Y(t)+\eta(t)]^2\right)^{1/2}$.

Theorem 1. *Assume that $\mathcal{T}$ is bounded, $\lim_{n\to\infty}\frac{n_j}{n}=\gamma_j$ for $j=1,2$, $\lim_{n\to\infty}\frac{n_{xy}}{n}=\gamma_{12}$, $\sqrt{n}\delta_{k(n)}\to 0$, that $E_{(P,Q)}\big(G^2\big)<\infty$, and that for $\delta_n\to 0$, $\sup_{|t-s|\le\delta_n}E_{(P,Q)}\Big(\big[X(t)+\epsilon(t)-X(s)-\epsilon(s)\big]^2+\big[Y(t)+\eta(t)-Y(s)-\eta(s)\big]^2\Big)\to 0$. Then, as $n_1n_2/n\to\infty$,*

$$\sqrt{\frac{n_1 n_2}{n}}\Big\{\big[\hat{\mu}_{n_1}(\cdot)-\mu(\cdot)\big]-\big[\hat{\eta}_{n_2}(\cdot)-\eta(\cdot)\big]\Big\}\overset{D}{\Rightarrow}W(\cdot), \tag{18}$$

where $W(\cdot)$ is the mean zero Gaussian process on $\mathcal{T}$ with covariance function

$$R(s,t)=E[W(s)\,W(t)]=\gamma_2R_{11}(s,t)-\gamma_{12}[R_{12}(s,t)+R_{21}(s,t)]+\gamma_1R_{22}(s,t).$$

The following remarks illustrate some special cases of Theorem 1.

Remark 1. In practice, the condition $\sup_{(P,Q)\in(\mathcal{P},\mathcal{Q})}E_{(P,Q)}(G^2)<\infty$ in Theorem 1 is not stringent, since $\mathcal{T}$ is a bounded set.

Remark 2. Since $\gamma_{12}\le\min\{\gamma_1,\gamma_2\}$, if $n_1>0$ and $\gamma_1=0$, then $\gamma_{12}=0$, $\gamma_2=1$ and $R(s,t)=R_{11}(s,t)$. Similarly, if $n_2>0$ and $\gamma_2=0$, then $R(s,t)=R_{22}(s,t)$. For these two case, i.e., $n_1=o(n)$ or $n_2=o(n)$, we can not test the equivalence of the mean curves $\mu(t)$ and $\eta(t)$. If $n_{12}=0$ or $\gamma_{12}=0$, the samples for $X(t)$ and $Y(t)$ are independent or asymptotically independent, so that $R(s,t)=\gamma_2R_{11}(s,t)+\gamma_1R_{22}(s,t)$. $\square$

Remark 3. If $n_1 = n_2 = n_{xy}$, we only observe the paired sample of $\big(X(t), Y(t)\big)^T$. In this case, $R(s,t) = \big[R_{11}(s,t) - R_{12}(s,t) - R_{21}(s,t) + R_{22}(s,t)\big]/2$ and Theorem 1 should be modified with $\sqrt{n_1 n_2/n}$ replaced by $\sqrt{n} = \sqrt{n_2} = \sqrt{n_1}$. □

3.2 Rejection Regions for Mean Test Statistics

By definition $R(s,t)$ is symmetric and square integrable on $\mathcal{T} \times \mathcal{T}$, and there are eigenvalues λ's and associated eigenfunction $h(t)$ such that

$$\int_{s \in \mathcal{T}} R(s,t) h(s) ds = \lambda h(t), \quad t \in \mathcal{T}. \tag{20}$$

Following the results of Theorem 1, we summarize in the following theorem the asymptotic distributions of the test statistics L_n and D_n given in (11) and (12).

Theorem 2. *Assume that the conditions of Theorem 1 are satisfied and $R(t,s)$ defined in (18) is continuous. Then, under the null hypothesis H_0 of (9) and (10),*

(i) $L_n \to |\mathcal{T}|^{-1} \sum_{k=1}^{\infty} \lambda_k Z_k^2$ in distribution as $n \to \infty$, and

(ii) $D_n \to N\big(0, \sigma^2\big)$ in distribution as $n \to \infty$,

where $\sigma^2 = |\mathcal{T}|^{-2} \int_{\mathcal{T}\times\mathcal{T}} R(s,t)\, ds\, dt$, λ_j's are the eigenvalues of $R(s,t)$, and Z_j's are independent identically distributed $N(0,1)$ random variables. □

Since, by Theorem 2, the asymptotic variances of L_n and D_n depend on the unknown eigenvalues λ_k and $\sigma^2 = |\mathcal{T}|^{-2} \int_{\mathcal{T}\times\mathcal{T}} R(s,t)\, ds\, dt$, a consistent estimator of $R(s,t)$ is needed to compute the approximate critical value for the test statistics. In practice, $R_{11}(s,t)$ can be estimated by

$$\begin{aligned} \widehat{R}_{11}(s,t) \;=\; & \tfrac{1}{n_1} \textstyle\sum_{i=1}^{n_1} \big[X_{i,k(n)}(t_s^*)\, X_{i,k(n)}(t_t^*)\big] \\ & - \tfrac{1}{n_1^2} \Big[\textstyle\sum_{i=1}^{n_1} X_{i,k(n)}(t_s^*)\Big] \Big[\textstyle\sum_{i=1}^{n_1} X_{i,k(n)}(t_t^*)\Big], \end{aligned} \tag{21}$$

where, for $a = s$ and t, $t_a^* = \arg\min\{|t_j - a| : j = 1, \ldots, k(n)\}$. Similarly, $R_{12}(s,t)$, $R_{21}(s,t)$ and $R_{22}(s,t)$ can be estimated by their corresponding estimates $\widehat{R}_{12}(s,t)$, $\widehat{R}_{21}(s,t)$ and $\widehat{R}_{22}(s,t)$, and, consequently, $R(s,t)$ can be estimated by

$$\widehat{R}(s,t) = \frac{n_1}{n}\widehat{R}_{11}(s,t) - \frac{n_{xy}}{n}[\widehat{R}_{12}(s,t) + \widehat{R}_{21}(s,t)] + \frac{n_2}{n}\widehat{R}_{22}(s,t).$$

Given that the function $R(t,s)$ does not have a close form expression, the eigenvalues λ_k are unknown and have to be estimated from the data. Let $\widehat{\mathbf{R}}_{k(n)}$ be the $k(n) \times k(n)$ matrix $\mathbf{R}_{k(n)} = \big(\widehat{R}(t_i, t_j)\big)$ of the estimators of $R(s,t)$ at the observed times $T_n = \{t_1, \ldots, t_{k(n)}\}$. As typical in practice, we compute the eigenvalues $\widehat{\lambda}_1, \ldots, \widehat{\lambda}_{k(n)}$ of $\widehat{\mathbf{R}}_{k(n)}$, and approximate the limit distribution of L_n by the distribution of

$$\widehat{L}_n = \frac{1}{|\mathcal{T}|}\sum_{j=1}^{k(n)} \widehat{\lambda}_j Z_j^2. \tag{22}$$

For a given significant level α, the rejection region for the null hypothesis in (9) is

$$L_n > Q_n(1-\alpha), \tag{23}$$

where the approximate critical value $Q_n(1-\alpha)$ is the $(1-\alpha)$th upper quantile of the distribution of $\widehat{L}_n$.

Remark 4. Results on eigenvalue estimation can be found elsewhere. For example, let $\lambda_1, ..., \lambda_p$ be the p largest eigenvalues, $\hat{\lambda}_j$'s be the estimates, by Theorem 2.7 in Horv*á*th and Kokoszka (2012, P.31), $E(\hat{\lambda}_j - \lambda_j)^2 = O(n^{-1})$, for all $1 \le j \le p$ for any fixed p. In application, we only need the first p largest eigenvalues, for some fixed p, such as $p = 10$. Recall in basis expansions, often only the first $k(\le 10)$ basis are used for well approximation, here the situation is similar.

For the rejection region of the test statistic D_n, we substitute $\int_{\mathcal{T}\times\mathcal{T}} R(s,t)\,ds\,dt$ with

a summation of $\widehat{R}(s,t)$ and estimate σ^2 by

$$\widehat{\sigma}_n^2 = \frac{1}{|\mathcal{T}|^2} \sum_{j_1=1}^{k(n)} \sum_{j_2=1}^{k(n)} \Big[\widehat{R}(t_{j_1}, t_{j_2}) \, (t_{j_1+1} - t_{j_1}) \, (t_{j_2+1} - t_{j_2}) \Big]. \tag{24}$$

Given that D_n is approximately normal when n is sufficiently large, the rejection region for the null hypothesis in (10) is approximated by

$$D_n > \widehat{\sigma}_n \, \Phi^{-1}(1-\alpha) \tag{25}$$

for a given significant level of α, where $\Phi(\cdot)$ is the cumulative distribution function of the $N(0,1)$ distribution.

The power for D_n to detect the difference $\Delta = |\mathcal{T}|^{-1} \int_T \big[\mu(t) - \eta(t)\big] \, dt$ under H_1 of (10) is

$$\beta(\Delta) = P\big[D_n > \sigma \, \Phi^{-1}(1-\alpha) \big| H_1\big] \approx 1 - \Phi\big[\Phi^{-1}(1-\alpha) - \sqrt{n} \, \Delta/\sigma_n\big],$$

and the estimated power is

$$\widehat{P}\big[D_n > \sigma \Phi^{-1}(1-\alpha) | H_1\big] = 1 - \Phi\big[\Phi^{-1}(1-\alpha) - \sqrt{n}\Delta/\widehat{\sigma}_n\big].$$

3.3 Asymptotic Distributions of Moment Curves

To obtain the asymptotic distributions of the correlation test statistics S_n and W_n of (17), we first develop the asymptotic properties of the moment estimates $\boldsymbol{\mu}_n(t)$ of (8). For notational simplicity, our derivation is only for the case of $n_{xy} = n$. Results of this section also hold for the case of $n_{xy} < n$ by substituting n with n_{xy}. The next theorem shows that the stochastic process $\sqrt{n}\big[\boldsymbol{\mu}_n(\cdot) - \boldsymbol{\mu}(\cdot)\big]$ converges weakly to a R^5-valued mean zero Gaussian process on $\mathcal{T}$. This result is used to derive the asymptotic distributions of the test statistics S_n and W_n in (17).

Theorem 3. *Assume conditions in Theorem 1. Then, as $n \to \infty$,*

$$\sqrt{n}\big[\boldsymbol{\mu}_n(\cdot) - \boldsymbol{\mu}(\cdot)\big] \overset{D}{\Rightarrow} \mathbf{W}(\cdot),$$

where $\mathbf{W}(\cdot)$ *is the* R^5*-valued mean zero Gaussian process on* $\mathcal{T}$ *with matrix covariance function* $\Omega(s,t) = Cov\big[\mathbf{Z}(s), \mathbf{Z}(t)\big]$ *and* $\mathbf{Z}(t) = \big(X(t)+\epsilon(t), Y(t)+\xi(t), [X(t)+\epsilon(t)]^2, [Y(t)+\xi(t)]^2, [X(t)+\epsilon(t)][Y(t)+\xi(t)]\big)^T$.

Since the test statistics S_n and W_n are functions of the moment process $\boldsymbol{\mu}(t)$, Theorem 3 suggests that the asymptotic distributions of S_n and W_n can be derived from the distribution of the Gaussian process $\mathbf{W}(t)$ and the covariance structure $\Omega(s,t)$ of $\mathbf{Z}(t)$.

3.4 Rejection Regions for Correlation Test Statistics

Based on the results of Theorem 3, we first derive here the asymptotic distributions of the test statistics S_n and W_n and then present their approximate rejection regions for the tests of (15) and (16).

For any R^5-valued vector $\mathbf{z} = (z_1, \ldots, z_5)$, we define

$$H(\mathbf{z}) = \frac{z_5 - z_1 z_2}{\sqrt{(z_3 - z_1^2)(z_4 - z_2^2)}}, \tag{26}$$

and its derivative $\dot{H}(\mathbf{z}) = \big(\partial H(\mathbf{z})/\partial z_1, ..., \partial H(\mathbf{z})/\partial z_5\big)$, where

$$\frac{\partial H(\mathbf{z})}{\partial z_1} = \frac{z_1 z_5 - z_2 z_3}{z_3 - z_1^2}\frac{\partial H(\mathbf{z})}{\partial z_5}, \qquad \frac{\partial H(\mathbf{z})}{\partial z_2} = \frac{z_2 z_5 - z_1 z_4}{z_4 - z_2^2}\frac{\partial H(\mathbf{z})}{\partial z_5},$$
$$\frac{\partial H(\mathbf{z})}{\partial z_3} = \frac{z_1 z_2 - z_5}{2(z_3 - z_1^2)}\frac{\partial H(\mathbf{z})}{\partial z_5}, \qquad \frac{\partial H(\mathbf{z})}{\partial z_4} = \frac{z_1 z_2 - z_5}{2(z_4 - z_2^2)}\frac{\partial H(\mathbf{z})}{\partial z_5},$$
$$\frac{\partial H(\mathbf{z})}{\partial z_5} = \frac{1}{\sqrt{(z_4 - z_2^2)(z_3 - z_1^2)}}.$$

Based on the results of Theorem 3, the next Theorem shows the asymptotic distributions of the statistics S_n and W_n.

Theorem 4. *Assume that the conditions of Theorem 3 are satisfied. Under the null hypothesis* H_0 *of (15) and (16), we have that, as* $n \to \infty$,

$$S_n \to |\mathcal{T}|^{-1} \sum_{j=1}^{\infty} \lambda_j Z_j^2 \quad \textit{in distribution} \tag{27}$$

and

$$W_n \to N\big(0, \sigma_2^2\big) \quad \textit{in distribution,} \tag{28}$$

where $\sigma_2^2 = |\mathcal{T}|^{-2} \int_{\mathcal{T}\times\mathcal{T}} Q(s,t)\, ds\, dt$, $Q(s,t) = \dot{H}\big[\boldsymbol{\mu}(s)\big]\Omega(s,t)\dot{H}^T\big[\boldsymbol{\mu}(t)\big]$ *with* $\Omega(s,t)$ *defined in Theorem 3, the* λ_j*'s are the eigenvalues of* $Q(s,t)$*, and the* Z_j*'s are i.i.d.* $N(0,1)$ *random variables.* □

From Theorem 4, for given nominal level α, we get the asymptotic rejection regions for H_0 of (15) and (16), based on S_n and W_n respectively as

$$S_n > G^{-1}(1-\alpha), \quad \text{and} \quad W_n > \widehat{\sigma}_{2,n}\, \Phi^{-1}(1-\alpha),$$

where $G^{-1}(1-\alpha)$ is the $(1-\alpha)$-th upper quantile of the distribution in the right hand side of (27), and the empirical estimator $\widehat{\sigma}_{2,n}$ is a consistent estimate of σ_2.

The power for S_n to detect the difference $\Delta = |\mathcal{T}|^{-1} \int_T R^2(t)\, dt$ under H_1 of (15) is

$$\beta(\Delta) = P\big[S_n > \chi^2(1-\alpha) - n\Delta \big| H_1\big] \approx 1 - G\big[\chi^2(1-\alpha) - n\Delta\big],$$

where $G(\cdot)$ is the distribution function of the right hand side of (27).

The power for W_n to detect the difference $\Delta = |\mathcal{T}|^{-1} \int_T R(t)\, dt$ under H_1 of (16) is

$$\beta(\Delta) = P\big[W_n > \sigma_2 \Phi^{-1}(1-\alpha) | H_1\big] \approx 1 - \Phi\big[\Phi^{-1}(1-\alpha) - \sqrt{n}\Delta/\widehat{\sigma}_{2,n}\big].$$

4 Application to Temporal Gene Expression Study

We apply the proposed testing procedures of Sections 2 and 3 to the temporal gene expression (TGE) data analyzed by Fang et al. (2012). The dataset contains repeated observations from a high-throughput gene expressions of 18 genes in *P. aeruginosa* expressed in 24 different biological conditions with various antibiotics (such as, *AMp, kam, Cm, Tc etc.*) at different concentrations. Then, 24 biological conditions can be as 24

independent samples. Gene expression outcome was measured by the log-scaled counts per second (CPS) every 30 minutes for 24 hours, which resulted in a total of 48 equally spaced observation time points. Details of the design and biological objectives of the experiment have been described in Duan et al. (2003) and Fang et al. (2012). Two main statistical objectives of the experiment are to test whether two selected genes of interest (a) have the same mean expression curves, and (b) are correlated over time.

For the purpose of illustration, we apply our testing methods of Sections 2 and 3 to the gene expression data of three genes, PA3897 (narL), PA2997 (nqrC), and PA0649 (trpG), which are listed in Table 1 of Fang et al. (2012). Among them, PA2997 and PA0649 are important genes related to energy metabolism, and PA3897 is related to the two-component response regulator NarL. Figure 1 shows the observed gene expression trajectories of these three genes among the 24 "biological conditions".

[Figure 1 is inserted here]

We computed the test statistics L_n and D_n for testing the equality of two log-scaled CPS mean curves in (9) and (10), and the test statistics S_n and W_n for testing the log-scaled CPS correlation curve of two genes in (15) and (16). Table 1 shows the values of the test statistics for the genes "PA2997 vs. PA3897" and "PA0649 vs. PA2997", the corresponding hypotheses, and the approximated p-values of the test statistics computed using the asymptotic distributions of Sections 3.2 and 3.4. These results show that the mean expression functions of gene PA2997 and gene PA3897 are significantly different with $p < 0.001$ for the test statistic L_n, and there is no significant difference between the mean expression functions of gene PA2997 and gene PA0694 with $p = 0.284$ and $p = 0.319$ for the test statistics L_n and D_n, respectively. The mean expressions of gene PA2997 are higher across the 24-hour time period than that of gene PA3897 with $p < 0.001$ for the test statistic D_n.

[Table 1 is inserted here]

Figure 2 shows the estimated correlation coefficient functions of gene pairs (PA2997, PA3897) and (PA0649, PA2997) across the 24-hour time period. The results of Table 1 for hypotheses (15) and (16) show that the genes PA3897 and PA2997 are not significantly correlated across the 24-hour time range with $p = 0.280$ and 0.5048 for the test statistics W_n and S_n, respectively, and the two genes PA2997 and PA0694, which are related to energy metabolism, have significantly positive correlation across the 24-hour time range with $p < 0.001$ for the one-sided test based on W_n and $p = 0.045$ for the two-sided test based on S_n.

[Figure 2 is inserted here]

5 Simulation

To investigate the finite sample properties of the proposed methods, and compare with the commonly used local linear smoothing (Loess) and spline methods, we present here several simulation studies, which are designed to mimic the practical situations with moderate sample sizes. We consider separately the tests for the equality of two mean curves and the tests for the correlation function between two stochastic processes. For each case, the simulation is based on 5000 replications, and the mean values of estimates over the replications are reported. The detailed descripitons of the simulation are given in a separate online file, here we only describe the results.

5.1 Testing the Equality of Mean Curves

5.1.1 Simulation for Test with Two-Sided Alternatives

For a series of alternatives determined by the C values, we present in Table 2 the corresponding differences $\Delta_1 = |\mathcal{T}|^{-1} \int_{\mathcal{T}} \left[\mu(t) - \eta(t)\right]^2 dt$, the values of the test statistic L_n, the empirical Type I error and the empirical powers under different values of Δ_1. These results suggest that the Type I error of the test is close to the nominal level of 0.05 and the power of the test is greater than 80% if $C \geq 0.50$ or $\Delta_1 \geq 0.035$.

[Table 2 is inserted here]

5.1.2 Simulation for Test with One-Sided Alternatives

Table 3 shows the values of the test statistic D_n, the empirical Type I error and the empirical powers under a series of C^* values and their corresponding values of $\Delta_2 = |\mathcal{T}|^{-1} \int_{\mathcal{T}} \left(\mu(t) - \eta^*(t)\right) dt$. These results show that the Type I error for D_n is close to the nominal level of 0.05 and the power is more than 80% if $C^* \geq 0.30$ or $\Delta_2 \geq 0.096$.

[Table 3 is inserted here]

5.2 Testing Correlation Functions

5.2.1 Simulation for Test with Two-Sided Alternatives

Table 4 shows the values of the test statistic S_n, the empirical Type I error and the empirical powers under a series of ρ values and their corresponding values of $\Delta_3 = |\mathcal{T}|^{-1} \int_{\mathcal{T}} R^2(t)\, dt$. These results show that the Type I error for S_n is close to the nominal level of 0.05 and the power is above 80% if $\rho \geq 0.115$ or $\Delta_3 \geq 0.0066$.

[Table 4 is inserted here]

5.2.2 Simulation for Test with One-Sided Alternatives

For testing the hypotheses in (16) using the test statistic W_n, our samples are generated the same way as in Section 5.2.1 except that the $\rho(t)$ of (29) is replaced by $\rho^*(t) = \rho\big|\sin(8 - t/10)\big|$, so that $\rho^*(t)$ does not change signs. Again, ρ determines the difference of the correlation curve $R(t)$ from zero.

[Table 5 is inserted here]

Table 5 shows the values of the test statistic W_n, the empirical Type I error and the empirical powers under a series of ρ values ant their corresponding values of $\Delta_4 = |\mathcal{T}|^{-1}\int_{\mathcal{T}} R(t)\,dt$. These results show that, again, the Type I error for W_n is close to the nominal level of 0.05 and the power is above 80% if $\rho \geq 0.08$ or $\Delta_4 \geq 0.0575$.

6 Discussion

We developed a class of simple nonparametric procedures based on linear interpolations of the observed functional data for testing multiple mean and correlation curves. As a direct response to the practical needs in biological studies, such as the temporal gene expression study of Section 4, our testing procedures serves as an alternative to existing methods such basis expansion, splines, smoothing, etc, and is simple to use. Our testing procedures also have the advantage of being computationally simple and having straightforward biological interpretations. The asymptotic properties of the test statistics suggest that our testing procedures can be theoretically justified under minimal assumptions, which is crucial in many biological studies. However, the actual use still need approximate of eigenvalues, similar to some other methods such as the FPCA.

Also, inference inaccuracy is expected due to the sparcity of the data. Our simulation results demonstrate that the approximate rejection regions obtained from the asymptotic distributions of the test statistics generally have satisfactory performance under practical settings. Further extensions of our method may include testing procedures involving structured nonparametric models with time-varying covariates or functional linear models.

Acknowledgement. The research of Yuan, Fang and Tan is partially supported by the National Cancer Institute (NCI) grant R01CA164717 and P30CA 051008.

REFERENCES

Antoniadis, A., Sapatinas, T. (2007). Estimation and inference in functional mixed-effects models. *Computational Statistics and Data Analysis*, **51**, 47934813.

Benko, M., Härdle, W., Kneip, A. et al. (2009). Common functional principal components. *The Annals of Statistics*, **37**(1), 1-34.

Cao, G., Yang, L., Todem, D. (2012). Simultaneous inference for the mean function based on dense functional data. *Journal of nonparametric statistics*, **24**(2), 359-377.

Cao, Y., Han, Y., Meng, D., Li, D., Jin, Q., Lin, Y., Cai, Y. (2016). Structural, evolutionary, and functional analysis of the class III peroxidase gene family in Chinese pear (Pyrus bretschneideri). *Frontier Plant Science* , **9**, https://doi.org/10.3389/fpls.2016.01874.

Chiang, C.T., Rice, J.A., Wu, C.O. (2001). Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables, *Journal of the American Statistical Association*, **96**, 605-619.

Cuevas, A., Febrero, M., Fraiman, R. (2004). An ANOVA test for functional data, *Computational Statistics and Data Analysis*, **47**, 111-122.

Degras, D. (2011). Simultaneous confidence bands for nonparametric regression with functional data. *Statistica Sinica*, **21**, 1735-1765.

Degras, D. (2017). Simultaneous confidence bands for the mean of functional data. *WIREs Computational Statistics*, e1397. doi:10.1002/wics.1397

Duan, K., Keerthi, S.S., Poo, A.N. (2003). Evaluation of simple performance measures for tuning SVM hyperparameters, *Neurocomputing*, **51**, 41-59.

Estévez-Pérez, G., & Vilar, J. A. (2008). Functional anova starting from discrete data: an application to air quality data. *Environmental and Ecological Statistics*, **20**, 495-515.

Fan, J., Zhang, J. T. (2000). Two-step estimation of functional linear models with applications to longitudinal data. *Journal of the Royal Statistical Society*, B, **62**, 303-322.

Fang, H-B, Deng, D., Tian, G-L, Shen, L., Duan, K., Song, J. (2012). Analysis for temporal gene expressions under multiple biological conditions, *Statistics in Biosciences*, **4**, 282-299.

Faraway, J.J. (1997). Regression analysis for a functional response. *Technometrics*, **39**, 254261.

Ferraty, F., Mas, A., Vieu, P. (2007). Nonparametric regression on functional data: Inference and practical aspects. *Australian & New Zealand Journal of Statistics*, **49** (3), 267286.

Fremdt, S., Steinebach, J.G., Horváth, L., Kokoszka, P. (2012). Testing the equality of covariance operators in functional samples. *Scandinavian Journal of Statistics*, **40**, 38-152.

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

Hastie, T., Tibshirani, R., Friedman, J. (2009). *Elements of Statistical Learning*, Springer.

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

Hoover, D.R., Rice, J.A., Wu, C.O., Yang, L.P. (1998). Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data, *Biometrika*, **85**, 809-822.

Huang, J.Z., Wu, C.O., Zhou, L. (2002). Varying Coefficient Models and Basis Function Approximations for the Analysis of Repeated Measurements. *Biometrika*, **89**, 111-128.

Kim, S. and Zhao, Z. (2013). Unified inference for sparse and dense longitudinal models. *Biometrika*, **100**, 203-212.

Kraus, D. and Panaretos, V. M. (2012). Disperson operators and resistant second-order functional data analysis. *Biometrika*, **99**, 813832.

Li, Y., Wang, N. and Carroll, R. J. (2010). Generalized functional linear models with semiparametric singleindex interactions. *Journal of the American Statistical Association*, **105**, 621-633.

Müller, H-G., and Yao, F. (2008). Functional additive models. *Journal of the American Statistical Association*, **103**, 1534-1544.

Pomann, G.M., Staicu, A.M., Gosh, M. (2016). Two-sample distribution-free test for functional data with application to a diffusion tensor imaging study of multiple sclerosis. *Journal of the Royal Statistical Society*, Series C, **65**, 395-414.

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

Rice, J.A., Wu, C.O. (2001). Nonparametric mixed effects models for unequally sampled noisy curves, *Biometrics*, **57**, 253-259.

Shi, M., Weiss, R.E., Taylor, J.M.C. (1996). An analysis of paediatric AIDS CD4 counts using flexible random curves, *Applied Statistics*, **45**, 151-163.

Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics*, John Wiley & Sons.

van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes*, Springer.

Wang, H.L., Zhong, P-S., Cui , Y.H., Li, Y. (2017). Unified empirical likelihood ratio tests for functional concurrent linear models and the phase transition from sparse to dense functional data. *Journal of the Royal Statistical Society*, B., to appear.

Wu, C.O., Chiang, C.T. (2000). Kernel smoothing on varying coefficient models with longitudinal dependent variable. *Statist Sinica*, **10**, 433-456.

Zhang, J-T. and Chen, J. (2007). Statistical inference for functional data. *The Annals of Statistics*, **35**, 1052-1079.

Zhang, J-T. (2011). Statistical inferences for linear models with functional responses. *Statistica Sinica*, **21**, 1431-1451.

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

Zhou, L., Huang, J. Z. and Carroll, R. J. (2008). Joint modelling of paired sparse functional data using principal components. *Biometrika*, **95**, 601-619.

Table 1: Hypotheses, test statistics and the corresponding p-values for two selected pairs of genes of the Temporal Gene Expression Study.

Gene pair	Hypotheses	Proposed			Loess		Spline	
		Statistic	Value	P-value	value	P-value	value	P-value
PA2997 vs PA3897	$\mu = \eta$ vs $\mu \geq \eta$	D_n	4.0621	< 0.0001	4.0684	< 0.0001	4.0618	< 0.0001
	$\mu = \eta$ vs $\mu \neq \eta$	L_n	19.6430	< 0.0001	19.6964	< 0.0001	19.6082	< 0.0001
	$R = 0$ vs $R > 0$	W_n	0.2852	0.3016	0.3324	0.2721	0.1224	0.4120
	$R = 0$ vs $R \neq 0$	S_n	0.9652	0.1907	1.0568	0.1559	1.0011	0.1686
PA0649 vs PA2997	$\mu = \eta$ vs $\mu \geq \eta$	D_n	-0.2261	0.7209	-0.2246	0.7196	-0.2261	0.7209
	$\mu = \eta$ vs $\mu \neq \eta$	L_n	0.5424	0.0852	0.5423	0.0896	0.5417	0.0796
	$R = 0$ vs $R > 0$	W_n	2.3851	< 0.0001	2.3978	< 0.0001	1.0017	0.0040
	$R = 0$ vs $R \neq 0$	S_n	6.5203	0.0011	6.5846	0.0007	6.5297	0.0008

Table 2: Test statistics L_n and its empirical Type I error and powers under different values of C and effective differences Δ_1 for the two-sided test of the mean curves. All results are based on $N = 5000$ replicates.

Time point	Constant	Effective difference	Proposed		Loess		Spline	
k	C	Δ_1	L_n	Power	Statistic	Power	Statistic	Power
10	0	0	0.0928	0.0394	0.0910	0.0398	0.0870	0.0382
	0.24	0.0076	0.2348	0.1410	0.2332	0.1298	0.1903	0.0878
	0.28	0.0103	0.2880	0.2186	0.2847	0.2118	0.2327	0.1304
	0.32	0.0135	0.3479	0.3658	0.3441	0.3536	0.2838	0.2076
	0.36	0.0170	0.4148	0.5738	0.4115	0.5580	0.3440	0.3514
	0.40	0.0210	0.4899	0.7832	0.4868	0.7772	0.4138	0.5636
30	0	0	0.0964	0.0590	0.0860	0.0472	0.0809	0.0460
	0.24	0.0076	0.3181	0.3784	0.3061	0.3404	0.2286	0.1616
	0.28	0.0103	0.4013	0.6740	0.3856	0.6254	0.2910	0.2956
	0.32	0.0135	0.4938	0.9086	0.4774	0.8802	0.3662	0.5622
	0.36	0.0170	0.6026	0.9892	0.5814	0.9836	0.4544	0.8342
	0.40	0.0210	0.7171	0.9994	0.6976	0.9996	0.5554	0.9722
50	0	0	0.0948	0.0530	0.0812	0.0482	0.0794	0.0470
	0.24	0.0076	0.3272	0.4192	0.3042	0.3408	0.2436	0.1920
	0.28	0.0103	0.4111	0.7324	0.3847	0.6380	0.3118	0.3658
	0.32	0.0135	0.5079	0.9432	0.4776	0.9026	0.3932	0.6668
	0.36	0.0170	0.6176	0.9942	0.5829	0.9858	0.4876	0.9152
	0.40	0.0210	0.7403	1.0000	0.7006	0.9998	0.5951	0.9896
100	0	0	0.0952	0.0594	0.0781	0.0508	0.0768	0.0508
	0.24	0.0076	0.3313	0.4522	0.3055	0.3688	0.2586	0.2354
	0.28	0.0103	0.4159	0.7824	0.3878	0.6862	0.3326	0.4604
	0.32	0.0135	0.5142	0.9668	0.4827	0.9304	0.4198	0.7920
	0.36	0.0170	0.6256	0.9994	0.5903	0.9964	0.5202	0.9688
	0.40	0.0210	0.7502	1.0000	0.7106	1.0000	0.6337	0.9992

Table 3: Test statistics D_n and its empirical Type I errors and powers under different values of C^* and effective differences Δ_2 for the one-sided test of the mean curves. All results are based on $N = 5000$ replicates.

Time point	Constant	Effective difference	Proposed		Loess		Spline	
k	C	Δ_2	D_n	Power	Statistic	Power	Statistic	Power
10	0	0	0.0002	0.0970	0.0003	0.1046	0.0002	0.0964
	0.10	0.0315	0.1059	0.3917	0.1183	0.4423	0.1067	0.3959
	0.15	0.0472	0.1587	0.6075	0.1772	0.6649	0.1598	0.6107
	0.20	0.0629	0.2116	0.7972	0.2362	0.8472	0.2129	0.8028
	0.25	0.0786	0.2644	0.9252	0.2951	0.9512	0.2659	0.9262
30	0	0	0.0010	0.0764	0.0009	0.0862	0.0009	0.0766
	0.10	0.0315	0.0994	0.3399	0.1050	0.3637	0.1006	0.3435
	0.15	0.0472	0.1486	0.5534	0.1571	0.5840	0.1503	0.5606
	0.20	0.0629	0.1978	0.7611	0.2092	0.7845	0.2000	0.7669
	0.25	0.0786	0.2470	0.9083	0.2612	0.9193	0.2496	0.9125
50	0	0	-0.0038	0.0624	-0.0039	0.0708	-0.0038	0.0632
	0.10	0.0315	0.0949	0.3206	0.0992	0.3396	0.0957	0.3232
	0.15	0.0472	0.1442	0.5277	0.1508	0.5541	0.1454	0.5311
	0.20	0.0629	0.1935	0.7397	0.2023	0.7582	0.1950	0.7435
	0.25	0.0786	0.2429	0.9001	0.2539	0.9083	0.2445	0.9035
100	0	0	-0.0001	0.0644	-0.0002	0.0730	-0.0001	0.0648
	0.10	0.0315	0.0983	0.3156	0.1006	0.3292	0.0987	0.3170
	0.15	0.0472	0.1476	0.5274	0.1510	0.5414	0.1481	0.5302
	0.20	0.0629	0.1968	0.7565	0.2014	0.7623	0.1975	0.7567
	0.25	0.0786	0.2460	0.9099	0.2518	0.9087	0.2468	0.9105

Table 4: Test statistics, empirical Type I errors and empirical powers under different values of ρ and effective differences Δ_3 for testing zero correlation function with two-sided alternatives. All results are based on $N = 5000$ replicates.

Time point	Constant	Effective Difference	Proposed		Loess		Spline	
k	ρ	Δ_3	S_n	Power	Statistic	Power	Statistic	Power
10	0	0	1.1160	0.0724	1.1160	0.0724	1.1156	0.0734
	0.12	0.0069	1.4678	0.2138	1.4678	0.2138	1.2849	0.1378
	0.15	0.0108	1.6694	0.3166	1.6694	0.3166	1.3832	0.1780
	0.18	0.0156	1.9166	0.4574	1.9166	0.4574	1.5040	0.2314
	0.20	0.0192	2.1069	0.5566	2.1069	0.5566	1.5979	0.2766
30	0	0	1.0277	0.0460	1.0360	0.0540	1.0340	0.0432
	0.12	0.0069	1.3653	0.3178	1.3640	0.2814	1.1613	0.1000
	0.15	0.0108	1.5553	0.5280	1.5491	0.4484	1.2349	0.1436
	0.18	0.0156	1.7877	0.7426	1.7755	0.6442	1.3242	0.2022
	0.20	0.0192	1.9662	0.8630	1.9494	0.7530	1.3922	0.2520
50	0	0	1.0207	0.0456	1.0208	0.0570	1.0211	0.0394
	0.12	0.0069	1.3529	0.4124	1.3436	0.2720	1.1558	0.1010
	0.15	0.0108	1.5401	0.6792	1.5254	0.4148	1.2307	0.1476
	0.18	0.0156	1.7690	0.8952	1.7478	0.5886	1.3229	0.2064
	0.20	0.0192	1.9449	0.9620	1.9187	0.7018	1.3945	0.2566
100	0	0	1.0096	0.0442	1.0122	0.0592	1.0106	0.0630
	0.12	0.0069	1.3390	0.6514	1.3334	0.2690	1.2165	0.1570
	0.15	0.0108	1.5243	0.9162	1.5136	0.4088	1.3331	0.2158
	0.18	0.0156	1.7509	0.9922	1.7338	0.5862	1.4747	0.2942
	0.20	0.0192	1.9249	0.9994	1.9029	0.7034	1.5841	0.3592

Table 5: Test statistics, empirical Type I errors and empirical powers under different values of ρ and effective differences Δ_4 for testing zero correlation function with one-sided alternatives. All results are based on $N = 5000$ replicates.

Time point	Constant	Effective difference	Proposed		Loess		Spline	
k	ρ	Δ_4	W_n	Power	Statistic	Power	Statistic	Power
10	0	0	0.0096	0.0500	0.0096	0.0500	0.0119	0.0580
	0.05	0.0306	0.2225	0.1688	0.2225	0.1688	0.1710	0.1124
	0.07	0.0428	0.3077	0.2364	0.3077	0.2364	0.2352	0.1436
	0.08	0.0489	0.3503	0.2780	0.3503	0.2780	0.2671	0.1618
	0.10	0.0611	0.4355	0.3748	0.4355	0.3748	0.3317	0.2062
30	0	0	0.0020	0.0556	-0.0004	0.0536	-0.0045	0.0552
	0.05	0.0306	0.2162	0.3402	0.2136	0.2408	0.1320	0.1326
	0.07	0.0428	0.3018	0.5130	0.2991	0.3574	0.1865	0.1736
	0.08	0.0489	0.3446	0.6060	0.3419	0.4264	0.2141	0.2006
	0.10	0.0611	0.4264	0.9234	0.4275	0.5660	0.2684	0.2576
50	0	0	-0.0014	0.0568	0.0059	0.0636	0.0047	0.0590
	0.05	0.0306	0.2125	0.4736	0.2192	0.2208	0.1449	0.1392
	0.07	0.0428	0.2980	0.7000	0.3046	0.3266	0.2012	0.1842
	0.08	0.0489	0.3408	0.7910	0.3473	0.3848	0.2293	0.2122
	0.10	0.0611	0.4264	0.9234	0.4326	0.5140	0.2854	0.2736
100	0	0	-0.0001	0.0530	0.0052	0.0648	0.0039	0.0626
	0.05	0.0306	0.2138	0.7188	0.2184	0.2258	0.1790	0.1500
	0.07	0.0428	0.2994	0.9230	0.3037	0.3338	0.2491	0.1960
	0.08	0.0489	0.3421	0.9678	0.3464	0.3944	0.2842	0.2210
	0.10	0.0611	0.4277	0.9962	0.4317	0.5282	0.3545	0.2754

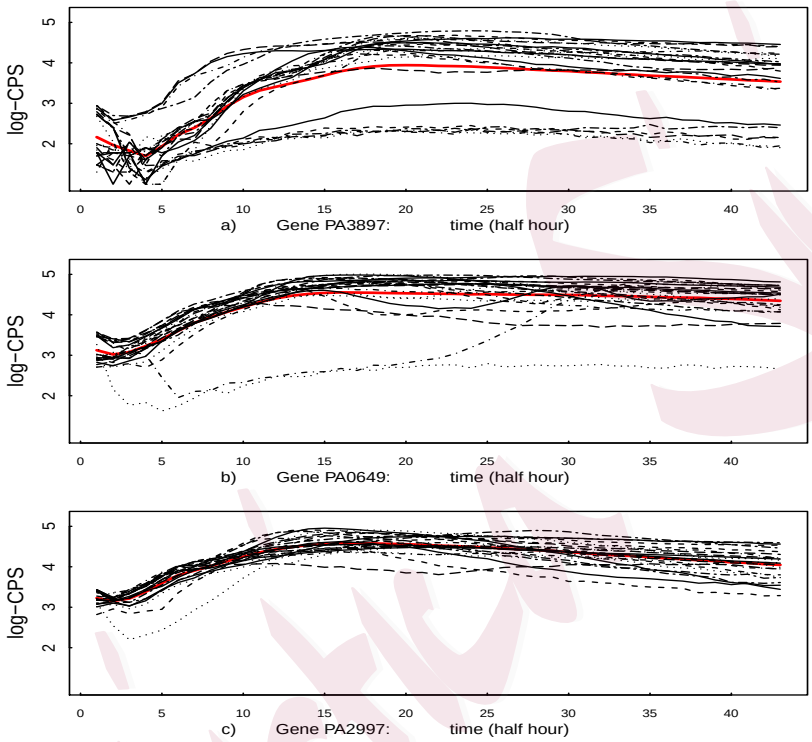

Figure 1: The observed expressions of three genes, PA3897 (narL), PA2997 (nqrC) and PA0649 (trpG), in 24 "biological conditions" over the 21 hours experimental time period. The red lines are the corresponding estimated mean functions of expression profiles.

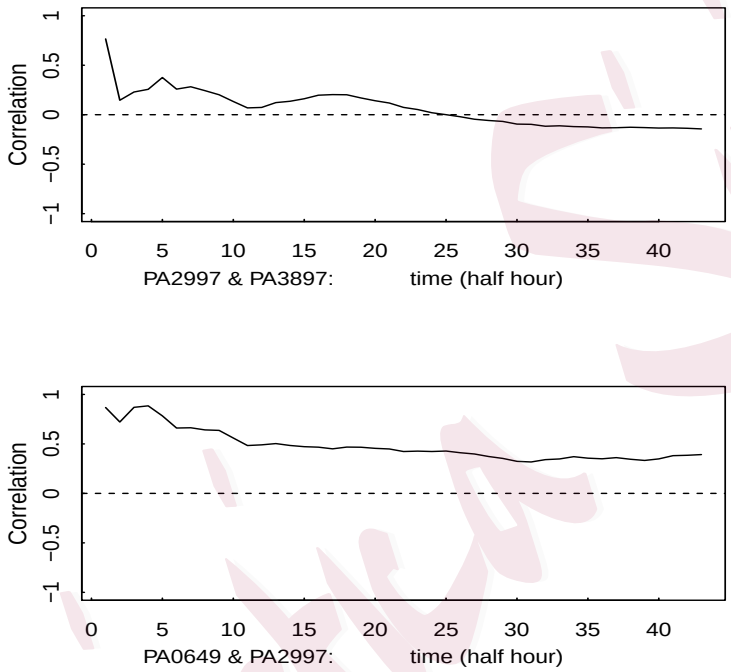

Figure 2: The estimated correlation coefficient curves of two pairs of genes over the 21 hours experimental time period. Top panel: PA2997 vs. PA3897. Bottom panel: PA0649 vs. PA2997.