Statistica Sinica Preprint No: SS-2024-0344	
Title	Statistical Inference for Functional Data over Multi Dimensional Domain
Manuscript ID	SS-2024-0344
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0344
Complete List of Authors	Qirui Hu and Lijian Yang
Corresponding Authors	Lijian Yang
E-mails	yanglijian@tsinghua.edu.cn
Notice: Accepted author version.	

Statistica Sinica

STATISTICAL INFERENCE FOR FUNCTIONAL DATA OVER MULTI-DIMENSIONAL DOMAIN

Qirui Hu, Lijian Yang

Tsinghua University

Abstract: This work develops inference tools for the mean function of functional data over a multi-dimensional domain. A two-step mean estimator based on tensor product spline estimates of individual trajectories is shown oracally efficient, i.e., it is asymptotically indistinguishable from the infeasible estimator using unobservable trajectories. Consistent estimates of covariance function as well as exact quantile of the limiting maximal deviation are obtained by innovative use of results on sharp comparison of Gaussian extreme distributions and quantiles, leading to asymptotic coverage and order $n^{-1/2}$ uniformly adaptive width of data-driven simultaneous confidence regions (SCRs). Also formulated are one-sided SCRs that can be used for testing against uniform upper and lower bound of the mean function. Extensive Monte Carlo experiments corroborate the theory, and a satellite ocean dataset collected by Copernicus Marine Environment Monitoring Service (CMEMS) illustrates how the proposed SCR is used.

Key words and phrases: Exact quantile; Functional data; Gaussian approximation; Simultaneous confidence region; Tensor product spline

1. Introduction

Functional data analysis (FDA) has been an important area of statistics research for over three decades. Functional data consist of observations of stochastic processes such as Electrocardiogram (ECG), Electroencephalogram (EEG) and human growth, see Ramsay and Dalzell (1991), Ferraty and Vieu (2006), Horváth and Kokoszka (2012) and Hsing and Eubank (2015), and the goal of FDA is to secure crucial information about the stochastic process.

There is already a rich collection of works on statistical inference of mean and covariance functions of functional data, such as Degras (2011), Cao et al. (2012), Gu et al. (2014), Zheng et al. (2014), Cao et al. (2016), Wang et al. (2020), Huang et al. (2022), Li and Yang (2023) and Zhong and Yang (2023). This body of work systematically develops various types of simultaneous confidence region (SCR), which is for an unknown function what a confidence interval is for an unknown parameter. Nonparametric SCRs are powerful tools for making global and uniform inference on unknown functions, a task often inadequately performed by pointwise confidence intervals. The lack of SCR is mainly due to the difficulty of obtaining limiting distribution of the uniform deviation in function estimation. See Wang and Yang (2009), Gu and Yang (2015) and Zheng et al. (2016) for SCR in nonparametric regression/generalized regression models, Wang et al. (2020), Yu et al. (2021) and Hu and Li (2024) for SCR over 2-dimensional irregular compact domains.

The goal of this paper is to extend the theory and methodology of SCR to functional data over multidimensional regular domain. Such extension include 2D satellite images, 3D human medical objects, or functional data of any dimension such as spatial temporal functional data. Take for example the global ocean temperature on 2020/9/15 at 0.00 am, available from Copernicus Marine Service. It can be viewed either as 2D surface data recorded on latitude and longitude grids of the earth, or 3D data if the sea depth is included as the third dimension (See Figure 1).

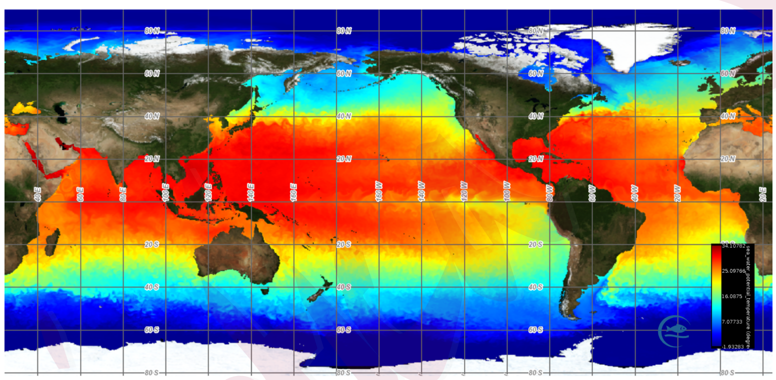

Figure 1: Heat map of global ocean temperature on 2020/9/15 at 0.00 am.

A functional random variable $\left\{\eta\left(\boldsymbol{x}\right), \boldsymbol{x} \in [0,1]^D\right\}$ is a square-integrable continuous stochastic process defined over the D-dimensional rectangle $[0,1]^D$, i.e., $\eta\left(\cdot\right) \in \mathcal{C}\left([0,1]^D\right)$ a.s. and $\mathbb{E}\left\|\eta\right\|^2 = \mathbb{E}\left(\sup_{\boldsymbol{x}\in[0,1]^D}\left|\eta\left(\boldsymbol{x}\right)\right|\right)^2 < +\infty$, with mean and covariance functions $m\left(\boldsymbol{x}\right) = \mathbb{E}\left\{\eta\left(\boldsymbol{x}\right)\right\} \in \mathcal{C}\left([0,1]^D\right), G\left(\boldsymbol{x},\boldsymbol{x}'\right) = \mathrm{Cov}\left\{\eta\left(\boldsymbol{x}\right), \eta\left(\boldsymbol{x}'\right)\right\} \in \mathcal{C}\left([0,1]^D \times [0,1]^D\right)$ respectively. Mercer lemma implies that $G\left(\boldsymbol{x},\boldsymbol{x}'\right) = \sum_{k=1}^{\infty}\lambda_k\psi_k\left(\boldsymbol{x}\right)\psi_k\left(\boldsymbol{x}'\right)$ for eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0, \sum_{k=1}^{\infty}\lambda_k < \infty$ and eigenfunctions $\left\{\psi_k\left(\boldsymbol{x}\right) \in \mathcal{C}\left([0,1]^D\right)\right\}_{k=1}^{\infty}$ such that $\int G\left(\boldsymbol{x},\boldsymbol{x}'\right)\psi_k\left(\boldsymbol{x}'\right)d\boldsymbol{x}' = \lambda_k\psi_k\left(\boldsymbol{x}\right)$, and $\left\{\psi_k\left(\boldsymbol{x}\right)\right\}_{k=1}^{\infty}$ form an orthonormal basis of $\mathcal{L}^2\left([0,1]^D\right)$. The standard Karhunen-Loève representation (See Adler and Taylor (2007)) is $\eta(\boldsymbol{x}) = m(\boldsymbol{x}) + \sum_{k=1}^{\infty} \xi_k \phi_k(\boldsymbol{x})$ in which the random coefficients $\{\xi_k\}_{k=1}^{\infty}$ are uncorrelated with mean 0 and variance 1, $\{\phi_k\}_{k=1}^{\infty}$ are rescaled eigenfunctions called functional principal components (FPCs) that satisfy $\phi_k = \sqrt{\lambda_k}\psi_k$ and $\int \{\eta(\boldsymbol{x}) - m(\boldsymbol{x})\} \phi_k(\boldsymbol{x})\, d\boldsymbol{x} = \lambda_k \xi_k$, $\forall k \in \mathbb{N}_+$.

A functional data set consists of n i.i.d. realizations $\left\{\eta_i(\boldsymbol{x}), \boldsymbol{x} \in [0,1]^D\right\}_{i=1}^{n}$ of $\eta(\cdot)$. The i-th trajectory $\eta_i(\boldsymbol{x}) = m(\boldsymbol{x}) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k(\boldsymbol{x})$, where $\{\xi_{ik}\}_{k=1}^{\infty}$, $i = 1, \ldots, n$, are i.i.d. copies of random coefficients $\{\xi_k\}_{k=1}^{\infty}$ called functional principal components scores (FPC scores). Actual functional data are observed sample points from trajectories $\{\eta_i(\cdot)\}_{i=1}^{n}$ with noises. Denote the equidistanced sample points $\boldsymbol{x}_{j_1 \ldots j_D} = (j_1/N_1, \ldots, j_D/N_D)$, $1 \le j_d \le N_d, 1 \le d \le D$, $R_i(\cdot) = \sum_{k=1}^{\infty} \xi_{ik}\phi_k(\cdot)$ and

$$Y_{i,j_1 \ldots j_D} = m(\boldsymbol{x}_{j_1 \ldots j_D}) + R_i(\boldsymbol{x}_{j_1 \ldots j_D}) + \sigma_i(\boldsymbol{x}_{j_1 \ldots j_D})\, \varepsilon_{i,j_1 \ldots j_D}, \tag{1.1}$$

where measurement errors $\varepsilon_{i,j_1 \ldots j_D}, 1 \le i \le n$, $1 \le j_d \le N_d, 1 \le d \le D$ are i.i.d. with $\mathbb{E}\varepsilon_{i,j_1 \ldots j_D} \equiv 0, \mathbb{E}\varepsilon_{i,j_1 \ldots j_D}^2 \equiv 1$, augmented by standard deviation functions $\sigma_i(\cdot)$. The data generating equation (1.1) extends the one-dimensional functional data setup of Cao et al. (2012), Cao et al. (2016).

A tensor product spline estimator of mean function $m(\cdot)$ is shown oracally efficient, i.e., it is asymptotically equivalent to the sample mean of trajectories fully observed without errors. SCRs are then constructed for $m(\cdot)$ by maxima deviation distribution of the proposed spline estimator. Proposition 1 and Theorem 5 amend one oversight in SCR theory for $m(\cdot)$, i.e., existence and consistent estimation of exact quantile for the maxima deviation process, by applying latest results from Yang (2025a) and Yang (2025b). Sharp comparison of Gaussian extreme distributions from Chernozhukov et al. (2015) establishes consistency in Theorem 5 of a data-driven quantile for the exact quantile of maximal error of multidimensional $m(\cdot)$. Another theoretical advance is asymptotic coverage and $n^{-1/2}$ uniformly adaptive width of the data-driven SCRs by consistent estimation of multidimensional covariance function in Theorem 4, both nonexisting in previous works Huang et al. (2022), Li and Yang (2023) and Zhong and Yang (2023).

Lower/upper SCRs in Subsection 3.3 generate tests against uniform upper/lower bound of $m(\cdot)$ with desired level and power according to Theorem 6. As an example, testing against uniform upper bound is carried out in Section 5 for the ocean temperature data by lower SCRs, yielding distinct outcomes over low- and high- altitude domains. This example illustrates that scientific questions from many disciplines such as environmental science and neuroscience, can be appropriately answered by testing uniform lower/upper bound of a functional mean with one-sided SCRs.

This work also successfully handles other technical challenges. First, smart vectorization makes partial sum Gaussian approximation for the multidimensional array

$\{\varepsilon_{i,j_1 \ldots j_D}\}_{1 \le i \le n, 1 \le j_d \le N_d, 1 \le d \le D}$ the same way for one-dimensional sequence of measurement errors in existing works. Second, unnatural assumption about finitely many distinct distributions of FPC scores in existing works Huang et al. (2022), Li and Yang (2023) and Zhong and Yang (2023) is dropped, by applying the new and explicit form of strong Gaussian approximation from Götze and Zaitsev (2010) instead of the classic implicit form in Csőrgő and Révész (1981).

The rest of the paper is organized as follows. Section 2 introduces the tensor spline estimator of the mean function. Section 3 presents main asymptotic results of the proposed estimator. Implementation details and simulation results are reported in Section 4. Our proposed SCR method is illustrated by the ocean temperature data in Section 5. All technical proofs, tables and figures of simulation results are collected in the Supplemental Material.

2. Estimator of the mean function

This section describes a two-step estimator for mean function $m(\cdot)$.

If all the trajectories $\{\eta_i(\cdot)\}_{i=1}^n$ were observed, one natural estimator of $m(\cdot)$ would be the sample average of n trajectories

$$\overline{m}(\cdot) = \overline{\eta}(\cdot) = n^{-1} \sum_{i=1}^{n} \eta_i(\cdot). \tag{2.2}$$

This would-be estimator $\overline{m}(\cdot)$ is, however, infeasible as it makes use of latent trajectories $\eta_i(\cdot)$. The following two-step estimator $\widehat{m}(\cdot)$ (or $\widehat{m}_p(\cdot)$) mimics $\overline{m}(\cdot)$ by using estimated $\widehat{\eta}_i(\cdot)$ in place of $\eta_i(\cdot)$

$$\widehat{m}(\cdot) = \widehat{m}_p(\cdot) = n^{-1}\sum_{i=1}^{n}\widehat{\eta}_i(\cdot), \tag{2.3}$$

where the estimated multi-dimensional trajectories $\widehat{\eta}_i(\cdot)$ are described in details below. For univariate case, see Cao et al. (2012), Wang et al. (2020), Li and Yang (2023), Huang et al. (2022).

For each $d = 1, \ldots, D$, denote by $\{t_{J_d}\}_{J_d=1}^{N_{s_d}}$ a sequence of equally-spaced points, called interior knots for variable direction d, that divide $[0, 1]$ into $(N_{s_d} + 1)$ equal subintervals, so $t_{J_d} = J_d h_{s_d}, 0 \leq J_d \leq N_{s_d} + 1$ where $h_{s_d} = 1/(N_{s_d} + 1)$ denotes the distance between neighbouring knots. These knots in D dimensions divide unit cube $[0, 1]^D$ into $\prod_{d=1}^{D}(N_{s_d} + 1)$ sub-rectangles of sizes $h_{s_1} \times \cdots \times h_{s_D}$. Denoted by $\mathscr{H}_d^{p-2} = \mathscr{H}_d^{p-2}[0, 1]$ the space of p-th order spline space in direction d, i.e. $(p - 2)$ times continuously differentiable functions on $[0, 1]$ that are polynomials of degree $(p - 1)$ on all subintervals $[t_{J_d}, t_{J_d+1}], 0 \leq J_d \leq N_{s_d}$. Explicitly

$$\mathscr{H}_d^{p-2} = \left\{\sum_{J_d=1-p}^{N_{s_d}} \lambda_{J_d,p} B_{J_d,p}(\cdot), \lambda_{J_d,p} \in \mathbb{R}\right\},$$

where $B_{J_d,p}$ is J_d-th spline basis of order p as defined in De Boor (2001)

The tensor product spline space $\mathscr{H}^{p-2,D}\left[0,1\right]^{D}$ is the tensor product of $\mathscr{H}_{d}^{p-2}\left[0,1\right], 1 \leq d \leq D$, i.e.

$$\begin{aligned}\mathscr{H}^{p-2,D}[0,1]^{D} &= \mathscr{H}_{1}^{p-2}\otimes\cdots\otimes\mathscr{H}_{D}^{p-2}\\ &= \left\{\sum_{J_1=1-p}^{N_{s_1}}\cdots\sum_{J_D=1-p}^{N_{s_D}}\lambda_{J_1\ldots J_D,p}B_{J_1\ldots J_D,p}^{[D]}\left(\boldsymbol{x}\right), \lambda_{J_1\ldots J_D,p}\in\mathbb{R}, \boldsymbol{x}\in[0,1]^{D}\right\},\end{aligned}$$

where for $\boldsymbol{x}=(x_1,\ldots,x_D)^{\mathsf{T}}\in[0,1]^{D}$, the $(J_1,\ldots,J_D)$-th tensor product B spline basis $B_{J_1\ldots J_D,p}^{[D]}\left(\boldsymbol{x}\right)=\prod_{d=1}^{D}B_{J_d,p}\left(x_d\right)$.

The tensor product spline estimator of trajectory $\eta_i\left(\cdot\right)$ is

$$\widehat{\eta}_i\left(\cdot\right)=\underset{g(\cdot)\in\mathscr{H}^{p-2,D}}{\arg\min}\sum_{1\leq j_d\leq N_d,1\leq d\leq D}\left\{Y_{i,j_1\ldots j_D}-g\left(\boldsymbol{x}_{j_1\ldots j_D}\right)\right\}^{2}, \qquad (2.4)$$

which is used in (2.3) to compute the two-step estimator $\widehat{m}\left(\boldsymbol{x}\right)$.

3. Asymptotic results

3.1 Assumptions

Throughout the paper, for vector $\boldsymbol{x}=(x_1,\ldots,x_D)^{\mathsf{T}}\in\mathbb{R}^{D}$, denote the Euclidean norm $\|\boldsymbol{x}\|=\sqrt{\sum_{d=1}^{D}x_d^2}$. To represent multivariate mixed derivatives, let the set $\mathbb{N}^{D}$

of D-indices be

$$\mathbb{N}^D = \{\boldsymbol{\alpha} : \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_D)^{\mathsf{T}}, \alpha_d \in \mathbb{N}, 1 \leq d \leq D\}.$$

For any $\boldsymbol{\alpha} \in \mathbb{N}^D$, denote $|\boldsymbol{\alpha}| = \sum_{d=1}^{D} \alpha_d$ and the $\boldsymbol{\alpha}$-th partial derivative of a D-variate function $\phi(\cdot)$ as

$$\partial^{\boldsymbol{\alpha}} \phi(\boldsymbol{x}) = \partial_{x_1}^{\alpha_1} \ldots \partial_{x_D}^{\alpha_D} \phi(\boldsymbol{x}).$$

The space $\mathcal{C}[0,1]^D$ of continuous functions on $[0,1]^D$ is Banach with respect to the sup norm $\|\phi\|_\infty = \sup_{\boldsymbol{x} \in [0,1]^D} |\phi(\boldsymbol{x})|$. Denote for any $\nu \in (0,1]$, and any $q \in \mathbb{N}$, denote the space of

$$\mathcal{C}^{q,\nu}[0,1]^D = \left\{\phi, \partial^{\boldsymbol{\alpha}}\phi \in \mathcal{C}^{0,\nu}[0,1]^D, \forall \boldsymbol{\alpha} \in \mathbb{N}^D, |\boldsymbol{\alpha}| = q\right\},$$

with seminorm

$$\|\phi\|_{q,v} = \max_{|\boldsymbol{\alpha}|=q} \sup_{\boldsymbol{x} \neq \boldsymbol{x}', \boldsymbol{x}, \boldsymbol{x}' \in [0,1]^D} \frac{|\partial^{\boldsymbol{\alpha}}\phi(\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}}\phi(\boldsymbol{x}')|}{\|\boldsymbol{x} - \boldsymbol{x}'\|^\nu}.$$

Denote also

$$\mathcal{C}^{q}[0,1]^D = \left\{\phi, \partial^{\boldsymbol{\alpha}}\phi \in \mathcal{C}[0,1]^D, \forall \boldsymbol{\alpha} \in \mathbb{N}^D, |\boldsymbol{\alpha}| = q\right\},$$

equipped with seminorm $\max_{|\boldsymbol{\alpha}|=q} \|\partial^{\boldsymbol{\alpha}}\phi\|_\infty$.

Constraints on constants related to the model are the following.

$$\nu, \mu \in (0,1], \quad q \in \mathbb{N}, \quad p^* = q + \mu, \tag{3.5}$$

$$\theta \in \left(0, \min\left\{\frac{2p^*}{1+p^*}, 2\nu\right\}\right), \tag{3.6}$$

$$\beta_2 \in \left(0, \min\left\{D/2, \nu - \frac{\theta}{2}, 1 - \frac{\theta}{2} - \frac{\theta}{2p^*}\right\}/D\right), \tag{3.7}$$

$$\omega_0 > \max\left\{4, \frac{4\theta}{\nu(2 - D\beta_2 - \theta)}\right\}, \tag{3.8}$$

and for smoothing parameter γ,

$$\max\left\{\frac{\theta}{2p^*} + \frac{2\theta}{p^*\omega_0}, 1 - \nu\right\} < \gamma < 1 - D\beta_2 - \frac{\theta}{2}. \tag{3.9}$$

The following assumptions are needed for theoretical results.

(A1) The mean function $m(\cdot) \in \mathcal{C}^{q,\mu}[0,1]^D$ for $q \in \mathbb{N}, \mu \in (0,1]$, $p^* = q + \mu$ in (3.5)

(A2) The standard deviation functions $\sigma_i(\cdot) \in \mathcal{C}^{0,\nu}[0,1]^D$ satisfy $\max_{1\le i\le n} \|\sigma_i\|_\infty \le M_\sigma$, $\max_{1\le i\le n} \|\sigma_i\|_{0,\nu} \le M_\sigma$ for ν in (3.5) and $M_\sigma > 0$.

(A3) Denote $N_{\max} = \max_{1\le d\le D} N_d, N_{\min} = \min_{1\le d\le D} N_d$, as $n \to \infty$ $N_{\max}/N_{\min} = \mathcal{O}(1)$, and $N = N_{\min}$ satisfies $N \gg n^{1/\theta}$ for θ in (3.6). Let $T_d = N_1 \times \cdots \times N_d, d = 1, \ldots, D$, then for all $d = 1, \ldots, D$, $T_d \propto N^d$.

(A4) There are constants $0 < c_G < C_G < \infty$ such that $c_G \leq G(\boldsymbol{x}, \boldsymbol{x}) \leq C_G, \forall \boldsymbol{x} \in [0,1]^D$ and for $k \in \mathbb{N}_+, \phi_k(\cdot) \in \mathcal{C}^{q,\mu}[0,1]^D$ with $\sum_{k=1}^{\infty} \{\|\phi_k\|_\infty + \|\phi_k\|_{0,\mu} + \|\phi_k\|_{q,\mu}\} < \infty$.

(A5) On the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, the FPC scores $\{\xi_{ik}\}_{i,k\in\mathbb{N}_+}$ are independent over $k \in \mathbb{N}_+$ and i.i.d. over $i \in \mathbb{N}_+$, the measurement errors $\{\varepsilon_{i,j_1\ldots j_D}\}_{i,j_1\ldots j_D\in\mathbb{N}_+}$ are i.i.d., and $\{\xi_{ik}\}_{i,k\in\mathbb{N}_+}$ are independent of $\{\varepsilon_{i,j_1\ldots j_D}\}_{i,j_1\ldots j_D\in\mathbb{N}_+}$. For $C_1, C_2 \in (0,+\infty), \gamma_1, \gamma_2 \in (1,+\infty), \beta_1 \in (0,1/2)$, β_2 in (3.7) $\kappa_n = \mathcal{O}(n^\omega)$, for some $\omega > 0$, there are i.i.d. $N(0,1)$ variables $\{Z_{ik,\xi}\}_{i=1,k=1}^{n,\kappa_n}, \{Z_{i,j_1,\ldots,j_D,\varepsilon}\}_{i=1,j_1,\ldots,j_D=1}^{n,N_1,\ldots,N_D}$ on a richer probability space $\left(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}}\right)$ in which all $\xi_{ik}, \varepsilon_{i,j_1\ldots j_D}$'s are embedded, such that

$$\mathbb{P}\left\{\max_{1\leq k\leq\kappa_n}\max_{1\leq t\leq n}\left|\sum_{i=1}^{t}\xi_{ik} - \sum_{i=1}^{t}Z_{ik,\xi}\right| > n^{\beta_1}\right\} < C_1 n^{-\gamma_1}$$

$$\mathbb{P}\left\{\max_{1\leq i\leq n, 1\leq t\leq T_D}\left|\sum_{j=1}^{t}\varepsilon_{i,j_1(j)\ldots j_D(j)} - \sum_{j=1}^{t}Z_{i,j_1(j)\ldots j_D(j),\varepsilon}\right| > N^{D\beta_2}\right\} < C_2 N^{-\gamma_2},$$

where $j_1(\cdot)\ldots j_D(\cdot)$ are functions $\mathbb{N}_+ \to \mathbb{N}_+$ defined by $j_D(t) = [T_{D-1}^{-1}(t-1)] + 1$ iteratively, and

$$j_{D-d}(t) = \begin{cases} N_{D-d} - [t-1-(j_{D-d+1}(t)-1)\,T_{D-d}]\,T_{D-d-1}^{-1}, & \text{if } j_{D-d+1}(t) \text{ is even} \\ [t-1-(j_{D-d+1}(t)-1)\,T_{D-d}]\,T_{D-d-1}^{-1}+1, & \text{otherwise.} \end{cases}$$

where $[a]$ denotes the integer part of a, and T_{D-d}'s are given in Assumption (A3). Further, $\sup_{k\in\mathbb{N}_+}\mathbb{E}\left|\xi_{ik}\right|^{\omega_0}<\infty$ for ω_0 in (3.8).

(A6) The number N_{s_d} of interior knots in direction $d, 1\le d\le D$ satisfy $\max_{1\le d\le D} N_{s_d}/\min_{1\le d\le D} N_{s_d}=\mathcal{O}(1)$, and one denotes $N_s=\min_{1\le d\le D} N_{s_d}$, $h_s=(N_s+1)^{-1}$. For γ in (3.9) and some $\tau>0$, $N_s N^{-\gamma}+N_s^{-1}N^{\gamma}=\mathcal{O}(\log^{\tau} N)$ as $N\to\infty$. While for some $\vartheta>0, \sum_{k=K_n}^{\infty}\|\phi_k\|_\infty=\mathcal{O}\left(n^{-\vartheta}\right)$ as $N\to\infty$, in which $K_n=\prod_{d=1}^{D}(N_{s_d}+p)$ is the dimension of the tensor product spline space $\mathscr{H}^{p-2,D}[0,1]^D$.

Assumptions (A1) and (A2) ensure smoothness of mean function and standard deviation function of measurement errors. Assumption (A3) dominates the sample size n by numbers N_d of observations per subject, by the same philosophy as in Wang et al. (2020), Huang et al. (2022), Li and Yang (2023), Zhong and Yang (2023). The bounded smoothness of eigenfunctions and their decay rate are guaranteed by Assumption (A4). Assumption (A6) regulates the number of knots for spline smoothing, in terms of smoothness of mean and standard deviation functions, and order of sample size and dimensionality. The high level Assumption (A5) on strong approximation is guaranteed by the following elementary Assumption (A5') according to Lemma S.10.

(A5') On the probability space $(\Omega,\mathcal{A},\mathbb{P})$, the FPC scores $\{\xi_{ik}\}_{i,k\in\mathbb{N}_+}$ are independent over $k \in \mathbb{N}_+$ and i.i.d. over $i \in \mathbb{N}_+$, the measurement errors $\{\varepsilon_{i,j_1\ldots j_D}\}_{i,j_1\ldots j_D \in \mathbb{N}_+}$ are i.i.d., and $\{\xi_{ik}\}_{i,k\in\mathbb{N}_+}$ are independent of $\{\varepsilon_{i,j_1\ldots j_D}\}_{i,j_1\ldots j_D\in\mathbb{N}_+}$. For constants $\varpi_1 > 4+2\omega$ and $\varpi_2 > (D+1+\theta)/\varpi_3$, $\sup_{k\in\mathbb{N}_+}\mathbb{E}\left|\xi_{1k}\right|^{\varpi_1} + \mathbb{E}\left|\varepsilon_{1,1\ldots1}\right|^{\varpi_2} < \infty$ in which $\omega > 0$ and $\varpi_3 = \min\{D/2, \nu-\theta/2, 1-\theta/2-\theta/(2p^*)\}$.

Remark 1 Assumptions (A1)-(A6) are easily met in many practical situations. For the case $D=2$ one may use as default $q=3, \nu=1, p^*=4, \theta=1/2, \gamma=15/64, \omega_0 = 1/4, d_N = \log\log N$.

Remark 2 The function set $j_1(\cdot), \ldots, j_D(\cdot)$ in Assumption (A5) can be replaced by any function set satisfying $\max_{s,t,\in\{1,\ldots,T_D\},|s-t|\le 1}\left\|\boldsymbol{x}_{j_1(s),\ldots,j_D(s)} - \boldsymbol{x}_{j_1(t),\ldots,j_D(t)}\right\| = \mathcal{O}\left(N^{-1}\right)$.

Remark 3 In contrast to Cao et al. (2012), Cao et al. (2016), Wang et al. (2020), Huang et al. (2022), Li and Yang (2023) and Zhong and Yang (2023) that used strong approximation in Csőrgő and Révész (1981), the explicit and more flexible strong approximation tools of Götze and Zaitsev (2010) used in this paper allows for infinitely many distinct distributions of FPC scores ξ_{ik}.

3.2 Error decomposition

For simplicity, denote by

$$\mathbf{B}_p(\boldsymbol{x}) = \Big(B^{[D]}_{1-p\ldots1-p,p}(\boldsymbol{x}), \ldots, B^{[D]}_{N_{s_1}\ldots1-p,p}(\boldsymbol{x}), \ldots, B^{[D]}_{1-pN_{s_2}\ldots1-p,p}(\boldsymbol{x})$$

$$\ldots B^{[D]}_{N_{s_1}N_{s_2}\ldots 1-p,p}\left(\boldsymbol{x}\right),\ldots,B^{[D]}_{N_{s_1}\ldots N_{s_D},p}\left(\boldsymbol{x}\ \right)\Big)^{\top}, \qquad \forall \boldsymbol{x}\in[0,1]^{D}.$$

$$\begin{aligned}\mathbf{X} &= \left(\mathbf{B}_p\left(\boldsymbol{x}_{1\ldots 1}\right),\ldots,\mathbf{B}_p\left(\boldsymbol{x}_{N_1\ldots 1}\right),\mathbf{B}_p\left(\boldsymbol{x}_{12\ldots 1}\right),\ldots,\mathbf{B}_p\left(\boldsymbol{x}_{N_1 2\ldots 1}\right),\ldots,\right. \\ &\quad \left.\mathbf{B}_p\left(\boldsymbol{x}_{N_1N_2\ldots 1}\right),\ldots,\mathbf{B}_p\left(\boldsymbol{x}_{N_1\ldots N_D}\right)\right)^{\top} = \left(\mathbf{B}_p\left(\boldsymbol{x}_{j_1\ldots j_D}\right)\right)^{\top}_{1\le j_d\le N_d,1\le d\le D}\end{aligned} \tag{3.10}$$

Denote also by $\mathbf{Y}_i = \left(Y_{i,j_1\ldots j_D}\right)^{\top}_{1\le j_d\le N_d,1\le d\le D}$, $\boldsymbol{m} = \left(m\left(\boldsymbol{x}_{j_1\ldots j_D}\right)\right)^{\top}_{1\le j_d\le N_d,1\le d\le D}$, $\boldsymbol{e}_i = \left(\sigma\left(\boldsymbol{x}_{j_1\ldots j_D}\right)\varepsilon_{i,j_1\ldots j_D}\right)^{\top}_{1\le j_d\le N_d,1\le d\le D}$, $\boldsymbol{\phi}_k = \left(\phi_k\left(\boldsymbol{x}_{j_1\ldots j_D}\right)\right)^{\top}_{1\le j_d\le N_d,1\le d\le D}$ and $\mathbf{R}_i = \sum_{k=1}^{\infty}\xi_i\boldsymbol{\phi}_k$,

Elementary algebra expresses estimators $\widehat{\eta}_i(\cdot)$ in (2.4) and $\widehat{m}_p(\cdot)$ in (2.3) in matrix form

$$\widehat{\eta}_i(\cdot) = \mathbf{B}_p(\cdot)^{\top}\left(\mathbf{X}^{\top}\mathbf{X}\right)^{-1}\mathbf{X}^{\top}\mathbf{Y}_i, \tag{3.11}$$

$$\widehat{m}_p(\cdot) = n^{-1}\sum_{i=1}^{n}\widehat{\eta}_i(\cdot). \tag{3.12}$$

Karhunen-Loève representation (1.1) decomposes estimator $\widehat{\eta}_i(\cdot)$ as

$$\widehat{\eta}_i(\cdot) = \widetilde{m}_p(\cdot) + \widetilde{R}_{ip}(\cdot) + \widetilde{\boldsymbol{e}}_{ip}(\cdot), 1\le i\le n \tag{3.13}$$

where $\widetilde{m}_p(\cdot)$, $\widetilde{R}_{ip}(\cdot)$, $\widetilde{\boldsymbol{e}}_{ip}(\cdot)$ respectively

$$\begin{aligned}
\widetilde{m}_p(\cdot) &= \mathbf{B}_p(\cdot)^\top \left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\mathbf{m}, \\
\widetilde{R}_{ip}(\cdot) &= \mathbf{B}_p(\cdot)^\top \left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\mathbf{R}_i, \\
\widetilde{\boldsymbol{e}}_{ip}(\cdot) &= \mathbf{B}_p(\cdot)^\top \left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\mathbf{e}_i.
\end{aligned}$$

Two main results are stated below.

Theorem 1. *Under Assumptions (A1)- (A6), as* $n \to \infty$,

$$\max_{1\le i\le n} \|\widehat{\eta}_i - \eta_i\|_\infty = \mathcal{O}_{a.s.}\left(N_s^{-p}(n\log n)^{2/\omega_0} + N^{D\beta_2-1}N_s\right).$$

Theorem 2. *Under Assumptions (A1) - (A6), as* $n \to \infty$,

$$\|\overline{m} - \widehat{m}_p\|_\infty = o_p\left(n^{-1/2}\right).$$

The above Theorem establish uniform closeness of each spline smoother of $\eta_i(\cdot)$ in (3.11) to the corresponding true trajectory, and oracle efficiency of the proposed estimator $\widehat{m}_p(\cdot)$ in (2.3) in the sense that the difference between tensor product spline estimator $\widehat{m}_p$ and $\overline{m}$ is of order $o_p\left(n^{-1/2}\right)$, much smaller than $\mathcal{O}_p\left(n^{-1/2}\right)$, the order of difference between $\overline{m}$ and m.

Let $\Xi(\boldsymbol{x}) = G(\boldsymbol{x},\boldsymbol{x})^{-1/2}\sum_{k=1}^{\infty} Z_k\phi_k(\boldsymbol{x})$ be the continuous sample path Gaussian process over $\boldsymbol{x} \in [0,1]^D$ with $\mathbb{E}\Xi(\cdot) \equiv 0, \mathbb{E}\Xi^2(\cdot) \equiv 1$ and covariance function

$$\mathbb{E}\Xi(\boldsymbol{x})\Xi(\boldsymbol{x}') \equiv G(\boldsymbol{x},\boldsymbol{x}')\left\{G(\boldsymbol{x},\boldsymbol{x})G(\boldsymbol{x}',\boldsymbol{x}')^{1/2}\right\}, \quad \boldsymbol{x},\boldsymbol{x}' \in [0,1]^D.$$

This Ξ is the weak limit of $|\overline{m}(\cdot) - m(\cdot)|\,G(\cdot,\cdot)^{-1/2}$ according to Theorem 3, used in the same way as a pivotal quantity. Simultaneous inference on $m(\cdot)$ is therefore based on distribution function of $\|\Xi\|_\infty = \sup_{\boldsymbol{x}\in[0,1]^D}|\Xi(\boldsymbol{x})|$, $\sup_{\boldsymbol{x}\in[0,1]^D}\Xi(\boldsymbol{x})$ or $\inf_{\boldsymbol{x}\in[0,1]^D}\Xi(\boldsymbol{x})$. For any real random variable M with continuous and strictly monotone distribution function, an *exact quantile* $q_{1-\alpha}$ exists at any level $1-\alpha \in (0,1)$, i.e., $\mathbb{P}[M \le q_{1-\alpha}] = 1-\alpha$. Unlike the counter-example in Tsirel'Son (1976), $\Xi(\cdot)$ satisfies the assumptions of Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b), thus random variables $\|\Xi\|_\infty$, $\sup_{\boldsymbol{x}\in[0,1]^D}\Xi(\boldsymbol{x})$ and $\inf_{\boldsymbol{x}\in[0,1]^D}\Xi(\boldsymbol{x})$ all have continuous and strictly monotone distribution functions, so exact quantiles exist for them at all levels.

Proposition 1. *For any $\alpha \in (0,1)$, there exist unique $(1-\alpha)$-th exact quantiles $Q_{1-\alpha}$, $Q_{+,1-\alpha}$ and $Q_{-,1-\alpha}$ such that*

$$\mathbb{P}\left\{\|\Xi\|_\infty \le Q_{1-\alpha}\right\} = \mathbb{P}\left\{\sup_{\boldsymbol{x}\in[0,1]^D}\Xi(\boldsymbol{x}) \le Q_{+,1-\alpha}\right\} = \mathbb{P}\left\{\inf_{\boldsymbol{x}\in[0,1]^D}\Xi(\boldsymbol{x}) \le Q_{-,1-\alpha}\right\} = 1-\alpha. \tag{3.14}$$

Let $z_{1-\alpha}$ denote the $100(1-\alpha)\%$-tile of $N(0,1)$.

Theorem 3. *Under Assumptions (A1)-(A6), $\forall \alpha \in (0,1)$, as $n \to \infty$, the "infeasible estimator" $\overline{m}(\boldsymbol{x})$ in (2.2) converges at rate $n^{1/2}$,*

$$\sqrt{n}\left\{\overline{m}(\cdot) - m(\cdot)\right\} G(\cdot,\cdot)^{-1/2} \to_d \Xi(\cdot).$$

Hence

$$\begin{aligned}
\mathbb{P}\left\{\sup_{\boldsymbol{x}\in[0,1]^D} \sqrt{n}\left|\overline{m}(\boldsymbol{x}) - m(\boldsymbol{x})\right| G(\boldsymbol{x},\boldsymbol{x})^{-1/2} \le Q_{1-\alpha}\right\} &\to 1-\alpha, \\
\mathbb{P}\left\{\sup_{\boldsymbol{x}\in[0,1]^D} \sqrt{n}\left\{\overline{m}(\boldsymbol{x}) - m(\boldsymbol{x})\right\} G(\boldsymbol{x},\boldsymbol{x})^{-1/2} \le Q_{+,1-\alpha}\right\} &\to 1-\alpha, \\
\mathbb{P}\left\{\inf_{\boldsymbol{x}\in[0,1]^D} \sqrt{n}\left\{\overline{m}(\boldsymbol{x}) - m(\boldsymbol{x})\right\} G(\boldsymbol{x},\boldsymbol{x})^{-1/2} \le Q_{-,1-\alpha}\right\} &\to 1-\alpha,
\end{aligned}$$

and for all $\boldsymbol{x} \in [0,1]^D$,

$$\begin{aligned}
\mathbb{P}\left\{\sqrt{n}\left|\overline{m}(\boldsymbol{x}) - m(\boldsymbol{x})\right| G(\boldsymbol{x},\boldsymbol{x})^{-1/2} \le z_{1-\alpha/2}\right\} \to 1-\alpha, \\
\mathbb{P}\left\{\sqrt{n}\left\{\overline{m}(\boldsymbol{x}) - m(\boldsymbol{x})\right\} G(\boldsymbol{x},\boldsymbol{x})^{-1/2} \le z_{1-\alpha}\right\} \to 1-\alpha.
\end{aligned}$$

Theorems 2 and 3 provide an infeasible simultaneous confidence region (SCR) of mean function $m(\cdot)$.

Corollary 1. *Under Assumptions (A1)-(A6), as $N \to \infty$, an asymptotic* $100(1-$

α)% correct SCR for $m(\cdot)$ is

$$\widehat{m}_p(\cdot) \pm n^{-1/2} G(\cdot,\cdot)^{1/2} Q_{1-\alpha},$$

while an asymptotic $100(1-\alpha)\%$ confidence interval for $m(\boldsymbol{x}), \boldsymbol{x} \in [0,1]^D$ is $\widehat{m}_p(\boldsymbol{x}) \pm n^{-1/2} G(\boldsymbol{x},\boldsymbol{x})^{1/2} z_{1-\alpha/2}$.

The above SCR is infeasible as it makes use of unknown $G(\cdot,\cdot)^{1/2}$ and quantile $Q_{1-\alpha}$. A method-of-moment estimator for covariance function $G(\boldsymbol{x},\boldsymbol{x}')$ is:

$$\widehat{G}_p(\boldsymbol{x},\boldsymbol{x}') = n^{-1} \sum_{i=1}^{n} \left(\widehat{\eta}_i(\boldsymbol{x}) - \widehat{m}_p(\boldsymbol{x})\right)\left(\widehat{\eta}_i(\boldsymbol{x}') - \widehat{m}_p(\boldsymbol{x}')\right), \boldsymbol{x},\boldsymbol{x}' \in [0,1]^D, \tag{3.15}$$

which is uniformly consistent.

Theorem 4. *Under Assumptions (A1) -(A6), the covariance estimator in (3.15) is uniformly consistent: for some $\varrho > 0$, as $N \to \infty$,*

$$\left\|\widehat{G}_p - G\right\|_\infty = \mathcal{O}_p\left(n^{-\varrho}\right).$$

As $\widehat{G}_p(\boldsymbol{x},\boldsymbol{x}')$ is an $n^{-\varrho}$-consistent for $G(\boldsymbol{x},\boldsymbol{x}')$, and it is of rank $K_n = \prod_{d=1}^{D}(N_{s_d}+p)$, an approximation of $\Xi(\cdot)$ is $\widehat{\Xi}_{K_n}(\cdot) = \widehat{G}_p(\cdot,\cdot)^{-1/2} \sum_{k=1}^{K_n} Z_k \widehat{\phi}_k(\cdot)$ where $\left\{\widehat{\phi}_k(\cdot)\right\}_{k=1}^{K_n}$ are the rescaled FPCs of $\widehat{G}_p(\boldsymbol{x},\boldsymbol{x}')$, Z_k are i.i.d. $N(0,1)$ variables $1 \le k \le K_n$, independent of $\mathbf{Y}_i, 1 \le i \le n$. Denote $\tilde{T}_D = n_1 \times \cdots \times n_D$ equally spaced grid points in $[0,1]^D$ as $\boldsymbol{t}_{l_1 \ldots l_D} = (l_1/n_1, \ldots, l_D/n_D)$, $1 \leq l_d \leq n_d, 1 \leq d \leq D$ over which $\left|\widehat{\Xi}_{K_n}(\cdot)\right|$ is computed and maximized, and denote the $(1-\alpha)$-th quantile of $\max_{1 \leq l_d \leq n_d, 1 \leq d \leq D} \left|\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1 \ldots l_D})\right|$ by $\widehat{Q}_{1-\alpha}$, the unique existence of which is ensured by Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b). The next theorem ensures that $\widehat{Q}_{1-\alpha}$ is a consistent estimator of quantile $Q_{1-\alpha}$.

Theorem 5. *Under Assumptions (A1) -(A6), if* $(\max_{1 \leq d \leq D} n_d)(\min_{1 \leq d \leq D} n_d)^{-1} \leq c$ *for some* $c \in (0, +\infty)$ *and* $\log \tilde{T}_D \propto \log n$*., then as* $N \to \infty$

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}\left(\sup_{\boldsymbol{x} \in [0,1]^D} \sqrt{n} \left| \frac{\widehat{m}_p(\boldsymbol{x}) - m(\boldsymbol{x})}{\widehat{G}_p^{1/2}(\boldsymbol{x}, \boldsymbol{x})} \right| \leq z \right) - \mathbb{P}\left(\max_{\substack{1 \leq l_d \leq n_d, \\ 1 \leq d \leq D}} \left|\widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1 \ldots l_D})\right| \leq z \left| \{\mathbf{Y}_i\}_{i=1}^n \right.\right) \right| = \mathcal{O}_p(1),$$

therefore, as $N \to \infty$*, an asymptotic* $100(1-\alpha)\%$ *correct SCR for* $m(\cdot)$ *is*

$$\widehat{m}_p(\cdot) \pm n^{-1/2} \widehat{G}_p(\cdot, \cdot)^{1/2} \widehat{Q}_{1-\alpha},$$

while an asymptotic $100(1-\alpha)\%$ *confidence interval for* $m(\boldsymbol{x})$, $\boldsymbol{x} \in [0,1]^D$ *is* $\widehat{m}_p(\boldsymbol{x}) \pm n^{-1/2} \widehat{G}_p(\boldsymbol{x}, \boldsymbol{x})^{1/2} z_{1-\alpha/2}$*. Furthermore,* $\widehat{Q}_{1-\alpha}$ *consistently estimates the quantile* $Q_{1-\alpha}$*, i.e., as* $N \to \infty$*,*

$$\left|\widehat{Q}_{1-\alpha} - Q_{1-\alpha}\right| \to_p 0,$$

$$\sup_{\boldsymbol{x} \in [0,1]^D} \left| \widehat{G}_p(\boldsymbol{x}, \boldsymbol{x})^{1/2} \widehat{Q}_{1-\alpha} - G(\boldsymbol{x}, \boldsymbol{x})^{1/2} Q_{1-\alpha} \right| \to_p 0.$$

Hence, $\forall \varepsilon > 0$*, with probability approaching* 1*, the width of SCR* $\widehat{m}_p(\boldsymbol{x}) \pm n^{-1/2} \widehat{G}_p(\boldsymbol{x},\boldsymbol{x})^{1/2} \widehat{Q}_{1-\alpha}$ *falls in the range*

$$n^{-1/2} G(\boldsymbol{x},\boldsymbol{x})^{1/2} (Q_{1-\alpha} - \varepsilon, Q_{1-\alpha} + \varepsilon)$$

uniformly for $\boldsymbol{x} \in [0,1]^D$.

3.3 One-sided methods

The proposed SCRs in Theorem 5 are two-sided, but there are situations where one-sided SCRs and hypothesis testing are more appropriate. For instance, one might test whether the surface temperatures of seawater in a specified region globally exceed a predefined threshold. Consider the following hypotheses concerning an uniform upper bound function $m_0(\cdot) \in \mathcal{C}([0,1]^D)$:

$$H_0 : m(\boldsymbol{x}) \leq m_0(\boldsymbol{x}), \forall \boldsymbol{x} \in [0,1]^D \text{ vs. } H_1 : m(\boldsymbol{x}) > m_0(\boldsymbol{x}), \exists \boldsymbol{x} \in [0,1]^D. \tag{3.16}$$

Quantile $Q_{+,1-\alpha}$ exists uniquely for the one-sided extreme $\sup_{\boldsymbol{x}\in[0,1]^D} \Xi(\boldsymbol{x})$ which is a continuous random variable with strictly monotone distribution function by Proposition 1. The corresponding $\widehat{Q}_{+,1-\alpha}$ represents the $(1-\alpha)$-th quantile of $\max_{1\leq l_d \leq n_d, 1\leq d\leq D} \widehat{\Xi}_{K_n}(\boldsymbol{t}_{l_1\ldots l_D})$, with uniqueness guaranteed by Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b). The decision rule is to reject the null hypothesis at significance level α if $m_0(\boldsymbol{x})$ is less than $\widehat{m}_p(\boldsymbol{x}) - n^{-1/2} \widehat{Q}_{+,1-\alpha} \widehat{G}^{1/2}(\boldsymbol{x},\boldsymbol{x})$

for some $\boldsymbol{x} \in [0,1]^D$. The test statistic is then

$$T_n = I\left\{m_0(\boldsymbol{x}) < \widehat{m}_p(\boldsymbol{x}) - n^{-1/2}\widehat{Q}_{+,1-\alpha}\widehat{G}^{1/2}(\boldsymbol{x},\boldsymbol{x}), \exists \boldsymbol{x} \in [0,1]^D\right\}.$$

The level and power of the proposed decision rule are justified as follows.

Theorem 6. *For $\alpha \in (0,1)$, $\widehat{Q}_{+,1-\alpha}$ consistently estimates the quantile $Q_{+,1-\alpha}$. As $N \to \infty$, both*

$$\left|\widehat{Q}_{+,1-\alpha} - Q_{+,1-\alpha}\right| \to_p 0,$$

and

$$\sup_{\boldsymbol{x}\in[0,1]^D}\left|\widehat{G}_p\left(\boldsymbol{x},\boldsymbol{x}\right)^{1/2}\widehat{Q}_{+,1-\alpha}, G\left(\boldsymbol{x},\boldsymbol{x}\right)^{1/2}Q_{+,1-\alpha}\right| \to_p 0.$$

Under the null hypothesis specified in (3.16),

$$\mathbb{P}\left(m_0(\boldsymbol{x}) < \widehat{m}_p(\boldsymbol{x}) - n^{-1/2}\widehat{Q}_{+,1-\alpha}\widehat{G}^{1/2}(\boldsymbol{x},\boldsymbol{x}), \exists \boldsymbol{x} \in [0,1]^D\right) \to \alpha,$$

and under the alternative hypothesis specified in (3.16),

$$\mathbb{P}\left(m_0(\boldsymbol{x}) < \widehat{m}_p(\boldsymbol{x}) - n^{-1/2}\widehat{Q}_{+,1-\alpha}\widehat{G}^{1/2}(\boldsymbol{x},\boldsymbol{x}), \exists \boldsymbol{x} \in [0,1]^D\right) \to 1.$$

Analogous test statistic can be formulated for the following with similar properties,

$$H_0 : m(\boldsymbol{x}) \geq m_0(\boldsymbol{x}), \forall \boldsymbol{x} \in [0,1]^D \text{ vs } H_1 : m(\boldsymbol{x}) < m_0(\boldsymbol{x}), \exists \boldsymbol{x} \in [0,1]^D.$$

3.4 Two sample extension

A two-sample extension of Theorem 3 is described in the Subsection. Denote two samples indicated by $s = 1, 2$, for any $1 \leq i \leq n_s, 1 \leq j_d \leq N_d, 1 \leq d \leq D$,

$$Y_{i,j_1\ldots j_D}^{(s)} = m^{(s)}\left(\boldsymbol{x}_{j_1\ldots j_D}\right) + R_i^{(s)}\left(\boldsymbol{x}_{j_1\ldots j_D}\right) + \sigma_i^{(s)}\left(\boldsymbol{x}_{j_1\ldots j_D}\right)\varepsilon_{i,j_1\ldots j_D}^{(s)}.$$

with covariance functions $G_s\left(\boldsymbol{x}, \boldsymbol{x}'\right) = \sum_{k=1}^{\infty} \phi_{sk}\left(\boldsymbol{x}\right)\phi_{sk}\left(\boldsymbol{x}'\right)$, respectively. One denotes the ratio of two-sample sizes as $\widehat{r} = n_1/n_2$ and assumes that $\lim_{n_1\to\infty} \widehat{r} = r > 0$.

For both groups, let $\widehat{m}_{1p}\left(\cdot\right)$ and $\widehat{m}_{2p}\left(\cdot\right)$ be the tensor product spline estimates of mean function by (2.3) or (3.12). Also denote by $\Xi_{12}\left(\cdot\right)$ a standardised Gaussian random field such that $\mathbb{E}\Xi_{12}\left(\cdot\right) \equiv 0, \mathbb{E}\Xi_{12}^2\left(\cdot\right) \equiv 1$ with covariance function

$$\mathbb{E}\Xi_{12}\left(\boldsymbol{x}\right)\Xi_{12}\left(\boldsymbol{x}'\right) = \frac{G_1\left(\boldsymbol{x}, \boldsymbol{x}'\right) + rG_2\left(\boldsymbol{x}, \boldsymbol{x}'\right)}{\{G_1\left(\boldsymbol{x}, \boldsymbol{x}\right) + rG_2\left(\boldsymbol{x}, \boldsymbol{x}\right)\}^{1/2}\{G_1\left(\boldsymbol{x}', \boldsymbol{x}'\right) + rG_2\left(\boldsymbol{x}', \boldsymbol{x}'\right)\}^{1/2}}.$$

Again, as Ξ_{12} satisfies the assumptions of Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b), $\|\Xi_{12}\|_\infty$ is continuous random variable with strictly monotone distribution function. Similar with Proposition 1, for any $\alpha \in (0, 1)$, there exists unique $Q_{12,1-\alpha}$, such that

$$\mathbb{P}\{\|\Xi_{12}\|_\infty \leq Q_{12,1-\alpha}\} = 1-\alpha.$$

One mimics the two-sample t-test and state the following theorem whose proof is analogous to that of Theorem 3

Theorem 7. *If Assumptions (A1)-(A6) are modified for each group accordingly, then as* $n_1 \to \infty, \widehat{r} \to r > 0$,

$$\sup_{\boldsymbol{x}\in[0,1]^D} \frac{\sqrt{n_1}\left\{\left(\widehat{m}_{1p} - \widehat{m}_{2p} - m_1 + m_2\right)(\boldsymbol{x})\right\}}{\left\{\left(G_1(\boldsymbol{x},\boldsymbol{x}) + rG_2(\boldsymbol{x},\boldsymbol{x})\right)\right\}^{1/2}} \to_d \Xi_{12}(\boldsymbol{x}).$$

Hence for any $\alpha \in (0,1)$,

$$\mathbb{P}\left\{\sup_{\boldsymbol{x}\in[0,1]^D} \frac{\sqrt{n_1}\left|\left(\widehat{m}_{1p} - \widehat{m}_{2p} - m_1 + m_2\right)(\boldsymbol{x})\right|}{\left\{\left(G_1(\boldsymbol{x},\boldsymbol{x}) + rG_2(\boldsymbol{x},\boldsymbol{x})\right)\right\}^{1/2}} \leq Q_{12,1-\alpha}\right\} \to 1-\alpha.$$

Analogously, denote the estimators of $G_1(\boldsymbol{x},\boldsymbol{x}'), G_2(\boldsymbol{x},\boldsymbol{x}')$ *as* $\widehat{G}_{1p}(\boldsymbol{x},\boldsymbol{x}'), \widehat{G}_{2p}(\boldsymbol{x},\boldsymbol{x}')$ *and estimated quantile as* $\widehat{Q}_{12,1-\alpha}$*. As* $N \to \infty$*, an asymptotic* $100(1-\alpha)\%$ *correct SCR for* $m_1(\cdot) - m_2(\cdot)$ *is*

$$\widehat{m}_{1p}(\cdot) - \widehat{m}_{2p}(\cdot) \pm n^{-1/2}\left(\widehat{G}_{1p} + \widehat{r}\widehat{G}_{2p}\right)(\cdot,\cdot)^{1/2}\widehat{Q}_{12,1-\alpha},$$

while an asymptotic $100(1-\alpha)\%$ *confidence interval for* $m_1(\boldsymbol{x}) - m_2(\boldsymbol{x}), \boldsymbol{x} \in [0,1]^D$ *is* $\widehat{m}_{1p}(\boldsymbol{x}) - \widehat{m}_{2p}(\boldsymbol{x}) \pm n^{-1/2}\left(\widehat{G}_{1p} + \widehat{r}\widehat{G}_{2p}\right)(\boldsymbol{x},\boldsymbol{x})^{1/2} z_{1-\alpha/2}$. *Furthermore,* $\widehat{Q}_{12,1-\alpha}$ *consistently estimates the* $(1-\alpha)$*-th quantiles, i.e., as* $N \to \infty$,

$$\left|\widehat{Q}_{12,1-\alpha} - Q_{12,1-\alpha}\right| \to_p 0,$$

$$\sup_{\boldsymbol{x}\in[0,1]^D}\left|\left(\widehat{G}_{1p} + \widehat{r}\widehat{G}_{2p}\right)^{1/2}(\boldsymbol{x},\boldsymbol{x})\,\widehat{Q}_{12,1-\alpha} - (G_1 + rG_2)(\boldsymbol{x},\boldsymbol{x})^{1/2} Q_{12,1-\alpha}\right| \to_p 0.$$

Hence $\forall\varepsilon > 0$, *with probability approaching* 1, *the width of SCR* $\widehat{m}_{1p}(\boldsymbol{x}) - \widehat{m}_{2p}(\boldsymbol{x}) \pm n^{-1/2}\left(\widehat{G}_{1p} + \widehat{r}\widehat{G}_{2p}\right)(\boldsymbol{x},\boldsymbol{x})^{1/2}\widehat{Q}_{12,1-\alpha}$ *falls in the range*

$$n^{-1/2}(G_1 + rG_2)(\boldsymbol{x},\boldsymbol{x})^{1/2}(Q_{12,1-\alpha} - \varepsilon, Q_{12,1-\alpha} + \varepsilon)$$

uniformly for $\boldsymbol{x} \in [0,1]^D$.

Furthermore, for one-sided hypothesis tests on two-sample problems,

$$H_0 : m^{(1)}(\boldsymbol{x}) \equiv m^{(2)}(\boldsymbol{x}), \forall \boldsymbol{x} \in [0,1]^D, \text{v.s. } H_1 : m^{(1)}(\boldsymbol{x}) > m^{(2)}(\boldsymbol{x}), \exists \boldsymbol{x} \in [0,1]^D,$$

one can establish the corresponding testing statistics and procedures based on Theorems 6 and 7 and omit them to save the manuscript space.

4. Implementation

This Section describes the computing of $\widehat{G}_p(\boldsymbol{x}, \boldsymbol{x}')$ and $\widehat{Q}_{1-\alpha}$ used in the SCR of Theorem 5. The SCR of Theorem 7 is computed similarly.

4.1 Estimating the covariance function

With the spline trajectories in (3.11), the estimator $\widehat{G}_p(\boldsymbol{x}, \boldsymbol{x}')$ is a tensor product spline:

$$\widehat{G}_p(\boldsymbol{x}, \boldsymbol{x}') = \mathbf{B}_p(\boldsymbol{x})^\top \widehat{g}_p \mathbf{B}_p(\boldsymbol{x}'),$$

where

$$\widehat{g}_p = n^{-1} \sum_{i=1}^{n} (\boldsymbol{\beta}_i - \boldsymbol{\beta})(\boldsymbol{\beta}_i - \boldsymbol{\beta})^\top,$$

$$\boldsymbol{\beta}_i = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y}_i, \boldsymbol{\beta} = n^{-1} \sum_{i=1}^{n} \boldsymbol{\beta}_i. \tag{4.17}$$

4.2 FPC analysis

Denote by $K_n = \prod_{d=1}^{D}(N_{s_d} + p)$ dimension of the tensor product spline space. For $k = 1, \ldots, K_n$, a tensor product spline estimate $\widehat{\psi}_k(\boldsymbol{x}) = \widehat{\gamma}_k^\top \mathbf{B}_p(\boldsymbol{x})$ is defined for $\psi_k(\boldsymbol{x})$, where the $\widehat{\gamma}_k$'s satisfy eigen-equations:

$$\int_{[0,1]^D} \widehat{G}_p(\boldsymbol{x}, \boldsymbol{x}') \widehat{\psi}_k(\boldsymbol{x}') d\boldsymbol{x}' = \widehat{\lambda}_k \widehat{\psi}_k(\boldsymbol{x}), k = 1 \ldots, K_n, \tag{4.18}$$

under the following constraints, with T_D as given in Assumption (A3),

$$T_D^{-1}\widehat{\gamma}_k^{\top}\mathbf{X}^{\top}\mathbf{X}\widehat{\gamma}_k = 1, k = 1\ldots, K_n. \tag{4.19}$$

There are only K_n eigenvalues and eigenfunctions can be solved from (4.18), since the integral operator induced by $\widehat{G}_p(\boldsymbol{x},\boldsymbol{x}')$ is a rank-K_n operator. Then, plugging in (4.17), equations (4.18) simplify to a matrix form:

$$T_D^{-1}\widehat{g}_p\mathbf{X}^{\top}\mathbf{X}\widehat{\gamma}_k = \widehat{\lambda}_k\widehat{\gamma}_k, k = 1\ldots, K_n. \tag{4.20}$$

The Cholesky decomposition $T_D^{-1}\mathbf{X}^{\top}\mathbf{X} = L_{\mathbf{X}}^{\top}L_{\mathbf{X}}$ is used to solve equations (4.20) subject to constraints (4.19). It is obvious that for $k = 1,\ldots,K_n$, the pair $\left(\widehat{\lambda}_k, L_{\mathbf{X}}\widehat{\gamma}_k\right)$ are the k-th eigenvalue and unit eigenvector of $L_{\mathbf{X}}\widehat{g}_pL_{\mathbf{X}}^{\top}$. Consequently, $\widehat{\gamma}_k$ equals $L_{\mathbf{X}}^{-1}\times$ the k-th unit eigenvector of $L_{\mathbf{X}}\widehat{g}_pL_{\mathbf{X}}^{\top}$ and $\widehat{\psi}_k(\boldsymbol{x}) = \widehat{\gamma}_k^{\top}\mathbf{B}_p(\boldsymbol{x}), \widehat{\phi}_k(\cdot) = \widehat{\lambda}_k^{1/2}\widehat{\psi}_k(\cdot)$. Matrix equations (4.20) are computationally much more expedient to solve than the integral equations (4.18).

4.3 Quantile estimation

In practice, $\widehat{Q}_{1-\alpha}$ is based on (3.15) and (4.18) by Monte Carlo simulation of $\widehat{\Xi}_{K_n,b}(\cdot) = \widehat{G}_p(\cdot,\cdot)^{-1/2}\sum_{k=1}^{K_n} Z_{k,b}\widehat{\phi}_k(\cdot)$, where $Z_{k,b}$ are i.i.d. standard normal variables with $1 \leq k \leq K_n, 1 \leq b \leq B$ where $B \to \infty$ (in practice, B is set as a large number, default=

1000). Let $\widehat{Q}_{1-\alpha,B}=$ sample $(1-\alpha)$-th quantile of $\left\{\max_{1\leq l_d\leq n_d,1\leq d\leq D}\left|\widehat{\Xi}_{K_n,b}\left(\boldsymbol{t}_{l_1\ldots l_D}\right)\right|\right\}_{b=1}^{B}$. The SCR for the mean function $m\left(\cdot\right)$ is computed as

$$\widehat{m}_p\left(\cdot\right)\pm n^{-1/2}\widehat{G}_p\left(\cdot,\cdot\right)^{1/2}\widehat{Q}_{1-\alpha,B}. \tag{4.21}$$

As Theorem 1 in Yang (2025a) and Theorem 1 in Yang (2025b) ensure continuity and strict monotonicity of the distribution function of $\max_{1\leq l_d\leq n_d,1\leq d\leq D}\left|\widehat{\Xi}_{K_n}\left(\boldsymbol{t}_{l_1\ldots l_D}\right)\right|$, one has that for fixed N, $\widehat{Q}_{1-\alpha,B}\rightarrow_{a,s.}\widehat{Q}_{1-\alpha}$ as $B\rightarrow\infty$.

Therefore

$$\lim_{N\rightarrow\infty}\lim_{B\rightarrow\infty}\mathbb{P}\left[m\left(\boldsymbol{x}\right)\in\widehat{m}_p\left(\boldsymbol{x}\right)\pm n^{-1/2}\widehat{G}_p\left(\boldsymbol{x},\boldsymbol{x}\right)^{1/2}\widehat{Q}_{1-\alpha,B},\boldsymbol{x}\in\left[0,1\right]^D\right]=1-\alpha,$$

so an asymptotic $100(1-\alpha)\%$ correct SCR for $m\left(\cdot\right)$ is $\widehat{m}_p\left(\cdot\right)\pm n^{-1/2}\widehat{G}_p\left(\cdot,\cdot\right)^{1/2}\widehat{Q}_{1-\alpha,B}$ as $\widehat{Q}_{1-\alpha,B}$ also consistently estimates the quantile $Q_{1-\alpha}$:

$$\lim_{N\rightarrow\infty}\lim_{B\rightarrow\infty}\left|\widehat{Q}_{1-\alpha,B}-Q_{1-\alpha}\right|=0,$$

$$\lim_{N\rightarrow\infty}\lim_{B\rightarrow\infty}\sup_{\boldsymbol{x}\in[0,1]^D}\left|\widehat{G}_p\left(\boldsymbol{x},\boldsymbol{x}\right)^{1/2}\widehat{Q}_{1-\alpha,B}-G\left(\boldsymbol{x},\boldsymbol{x}\right)^{1/2}Q_{1-\alpha}\right|=0,$$

in probability. Hence, $\forall\varepsilon>0$, with probability approaching 1, the width of SCR

$\widehat{m}_p(\boldsymbol{x}) \pm n^{-1/2}\widehat{G}_p(\boldsymbol{x},\boldsymbol{x})^{1/2}\widehat{Q}_{1-\alpha,B}$ falls in the range

$$n^{-1/2}G(\boldsymbol{x},\boldsymbol{x})^{1/2}(Q_{1-\alpha}-\varepsilon, Q_{1-\alpha}+\varepsilon)$$

uniformly for $\boldsymbol{x} \in [0,1]^D$.

Similarly, with $\widehat{Q}_{12,1-\alpha,B}$ computed analogously, the SCR for $m_1(\cdot) - m_2(\cdot)$ is:

$$\widehat{m}_{1p}(\cdot) - \widehat{m}_{2p}(\cdot) \pm n^{-1/2}\left\{\left(\widehat{G}_{1p} + \widehat{r}\widehat{G}_{2p}\right)(\cdot,\cdot)\right\}^{1/2}\widehat{Q}_{12,1-\alpha,B}. \tag{4.22}$$

4.4 Simulation studies

For simplicity of presentation, one sets $D = 2$, $N_1 = N_2 = N$ and:

$$\begin{aligned}
m(\boldsymbol{x}) &= 2\sin\{\pi(x_1+x_2)/2\}e^{-(x_1+x_2)} + x_1\sin x_2,\\
\phi_1(\boldsymbol{x}) &= 2\sqrt{2}\sin(\pi x_1/2)\sin(\pi x_2/2),\\
\phi_2(\boldsymbol{x}) &= 2\sqrt{2}\sin(3\pi x_1/2)\sin(\pi x_2/2),\\
\phi_3(\boldsymbol{x}) &= 2\sin(3\pi x_1/2)\sin(\pi x_2/2),\\
\phi_4(\boldsymbol{x}) &= 2\sin(3\pi x_1/2)\sin(3\pi x_2/2),\\
\phi_5(\boldsymbol{x}) &= \sqrt{2}\sin(5\pi x_1/2)\sin(3\pi x_2/2),\\
\phi_6(\boldsymbol{x}) &= \sqrt{2}\sin(5\pi x_1/2)\sin(5\pi x_2/2),
\end{aligned}$$

$$\phi_k(\boldsymbol{x}) = 0 \qquad k \geq 7.$$

which implies $\lambda_1 = \lambda_2 = 2, \lambda_3 = \lambda_4 = 1, \lambda_5 = \lambda_6 = 0.5$ and $\lambda_k = 0$, $k \geq 7$. For all $1 \leq i \leq n, 1 \leq k < \infty, 1 \leq j_1, j_2 \leq N$ $\xi_{ik}, \varepsilon_{i,j_1j_2}$ are mutually independent and identically distributed. The ξ_{ik}'s follow one of three distributions: normal, uniform and Laplace, all with mean 0 and variance 1. The ε_{i,j_1j_2}'s are generated similarly. The standard deviation functions range from homoscedastic to strongly heteroscedastic, including $\sigma(x_1, x_2) \equiv 0.1$, 0.2 or $\sigma(x_1, x_2) = 0.15\,(5 - \exp(x_1 + x_2))\,/\,(5 + \exp(x_1 + x_2))$, $\sigma(x_1, x_2) = 0.3\,(5 - \exp(x_1 + x_2))\,/\,(5 + \exp(x_1 + x_2))$. The sample size and knots are $n = \left[0.4N^{1/2}(\log N)^2\right]$ and $N_{s_1} = N_{s_2} = \left[0.4N^{15/64}\log\log N\right]$, respectively. Cubic splines ($p = 4$) are used throughout this section. Empirical coverage rate is computed among the 1000 replications, which is the relative frequency that the true surface $m(\cdot, \cdot)$ is entirely covered by the SCR.

Our setup allows for the sample $\{Y_{i,j_1j_2}\}_{i=1,j_1=1,j_2=1}^{n,N,N}$ to be generated by 36 combinations. Tables S.1-S.4 present the results by standard deviation functions. It is clear that the empirical coverage rate approaches the nominal confidence level as the sample sizes increase, a positive confirmation of Theorem 3.

Figures S.1 - S.4 depict the true mean function, the spline estimator of mean function and the 95% SCR of mean function respectively. They are all based on a typical run under the setting $N = 50, 100, 200, 400$ (correspondingly $n = 44, 85, 159, 288$).

As expected, when n increases, the tensor product estimator approximates the true surface and the simultaneous confidence region becomes narrower.

5. Real data example

We consider the data from the CMEMS global analysis and forecast product which contains 3D potential temperature, salinity and currents information from top to bottom and 2D sea surface level, potential bottom temperature, mixed layer thickness, sea ice thickness, sea ice fraction and sea ice velocities information. These globally observed data are defined on rectangular grid at approximately 8km and 50 standard vertical levels. The data are available on the Copernicus Marine Service website https://marine.copernicus.eu/access-data.

We focus on an area with longitude from -175 to -141 and latitude from -30 to -64, which is divided into a high-latitude domain $\Omega_2 = [-175, -141] \times [-47, -30]$ closer to the Equator a low-latitude domain $\Omega_1 = [-175, -141] \times [-64, -47]$ closer to the South Pole. Each sample is a surface over the rectangular area Ω_1 or Ω_2 with $409 \times 205(204)$ grids. A total of $n = 90$ daily observations are recorded between 2023-3-10 and 2023-8-10. The sea surface potential temperature and part of three-dimensional observations are shown in Figure 2

Tensor product cubic splines ($p = 4$) are used to estimate the mean function of sea surface potential temperature, which reflects the overall variation of the data

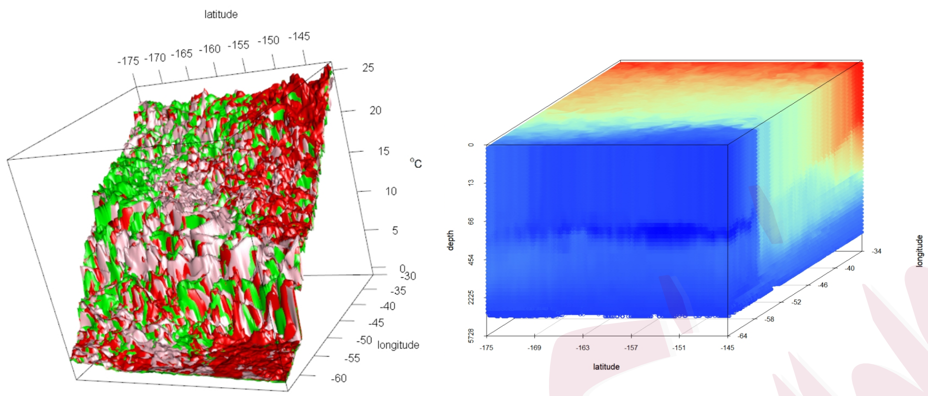

Figure 2: The 2D-image and 3D-image of some sample.

and its inner structure. Figure 3 depicts a one-sided 95% lower SCR, blown up by a factor of 20 for better view, as the SCR is too narrow for visualization.

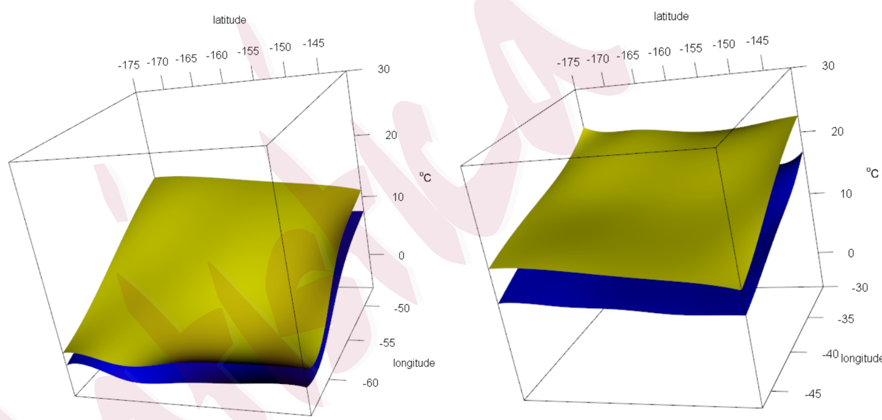

Figure 3: Mean function estimates and 95% lower SCRs over Ω_1 (left) and Ω_2 (eight).

The lower SCRs allow one to test the null hypothesis that the mean potential temperature $m(\cdot)$ is below some threshold across the entire rectangular domain. For instance, null and alternative hypotheses may be formalized as follows:

$$H_0 : m(\boldsymbol{x}) \leq 14.7, \forall \boldsymbol{x} \in \Omega, \quad \text{versus} \quad H_1 : m(\boldsymbol{x}) > 14.7, \exists \boldsymbol{x} \in \Omega, \tag{5.23}$$

Figure 4 shows distinct potential temperature differences in the high-latitude domain Ω_2 and less distinction in the low-latitude domain Ω_1. Subsequent hypothesis testing within domains Ω_1 and Ω_2 yields p-values of 0.061 and 0.001, respectively. Thus, the null hypothesis is retained for domain Ω_1 due to lack of evidence for the alternative but strongly rejected for Ω_2. These findings support the expectation of warmer sea surface in higher altitude regions with quantified uncertainty.

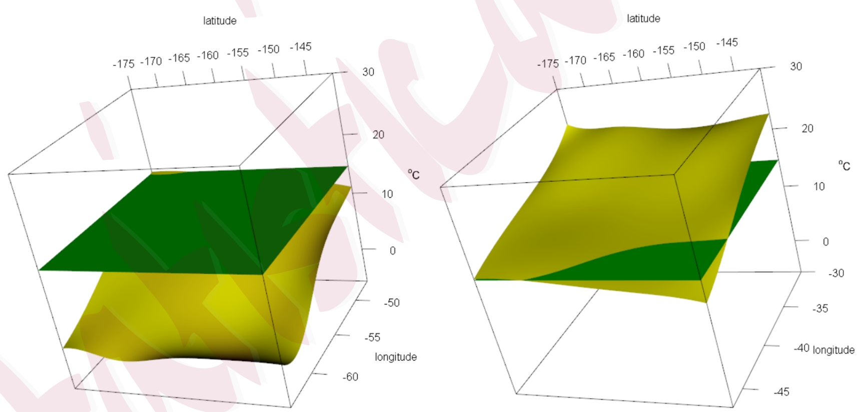

Figure 4: Mean function estimates and flat $m_0(\cdot) \equiv 14.7$ over Ω_1 (left) and Ω_2 (eight).

6. Concluding remarks

Statistical inference is developed for the mean function of functional data over a multi-dimensional domain. The "infeasible" estimator, i.e., the sample average of all true trajectories, satisfies a $\mathcal{C}\left([0,1]^D\right)$-Central Limit Theorem. Tensor product spline is used to recover each trajectory, leading to a two-step mean estimator that is oracally efficient, meaning that it is asymptotically indistinguishable from the infeasible estimator using unobservable trajectories. For the first time, consistent estimates are established for covariance function of multi-dimensional trajectory and exact quantile of a multi-dimensional maximal deviation Gaussian process. These consistent estimates are used to compute data-driven SCRs with preset asymptotic coverage and uniformly adaptive width of order $n^{-1/2}$. All theoretical results work for one-sided SCRs as for the more commonly used two-sided SCRs. One-sided lower SCRs for ocean surface temperature over two rectangular domains reveal geographical differences in global warming, and extensive Monte Carlo experiments illustrate the numerical performance of the proposed SCRs. Future works may zero in on statistical inference of covariance structure and eigen-systems of functional data over multi-dimensional domains.

Acknowledgements

This work is partially supported by National Natural Science Foundation of China award 12171269.

References

Adler, R. J., and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer, New York.

Cao, G., Wang, L., Li, Y., and Yang, L. (2016). Oracle-efficient confidence envelopes for covariance functions in dense functional data. *Statist. Sinica* **26** 359–383.

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

Chernozhukov, V., Chetverikov, D., and Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Relat. Fields* **162** 47–70.

Csőrgő, M., and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York-London.

De Boor, C. (2001). *A Practical Guide to Splines*. Springer, New York.

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

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

Götze, F., and Zaitsev, A. Y. (2010). Rates of approximation in the multidimensional invariance principle for sums of i.i.d. random vectors with finite moments. *J. Math. Sci.* **167** 495–500.

Gu, L., Wang, L., Härdle, W., and Yang, L. (2014). A simultaneous confidence corridor for varying coefficient regression with sparse functional data. *TEST* **23** 806–843.

Gu, L., and Yang, L. (2015). Oracally efficient estimation for single-index link function with simultaneous confidence band. *Electron. J. Statist.* **9** 1540–1561.

Hall, P., Müller, H. G., and Wang, J. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517.

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

Hsing, T., and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley, Chichester.

Hu, Q., and Li, J. (2024). Statistical inference for mean function of longitudinal imaging data over complicated domains. *Statist. Sinica* **34** 955–982.

Huang, K., Zheng, S., and Yang, L. (2022). Inference for dependent error functional data with application to event-related potentials. *TEST* **31** 1100–1120.

Klenke, A. (2014). *Probability Theory*, 2nd edn. Springer, London.

Li, J., and Yang, L. (2023). Statistical inference for functional time series. *Statist. Sinica* **33** 519–549.

Ma, S., Yang, L., and Carroll, R. (2012). A simultaneous confidence band for sparse longitudinal regression. *Statist. Sinica* **22** 95–122.

Ramsay, J., and Dalzell, C. (1991). Some tools for functional data analysis. *J. Roy. Statist. Soc. Ser. B* **53** 539–572.

Schumaker, L. (1981). *Spline Functions: Basic Theory*. Wiley.

Song, Q., and Yang, L. (2009). Spline confidence bands for variance functions. *J. Nonparametr. Stat.* **5** 589–609.

Tsirel'Son, V. S. (1976). The density of the distribution of the maximum of a Gaussian process. *Theory Probab. Appl.* **20** 847–856.

Wang, J., Cao, G., Wang, L., and Yang, L. (2020). Simultaneous confidence band for stationary covariance function of dense functional data. *J. Multivariate Anal.* **176** 104584.

Wang, Y., Wang, G., Wang, L., and Ogden, R. T. (2020). Simultaneous confidence corridors for mean functions in functional data analysis of imaging data. *Biometrics* **76** 427–437.

Wang, J., and Yang, L. (2009). Polynomial spline confidence bands for regression curves. *Statist. Sinica* **19** 325–342.

Yang, L. (2025a). Strict monotonicity of stochastic process extreme distributions. *Statist. Probab. Lett.* **217** 110292.

Yang, L. (2025b). Continuity of Gaussian extreme distributions. *Statist. Probab. Lett.* **215** 110274.

Yu, S., Wang, G., Wang, L., and Yang, L. (2021). Multivariate spline estimation and inference for image-on-scalar regression. *Statist. Sinica* **31** 1463–1487.

Zheng, S., Liu, R., Yang, L., and Härdle, W. (2016). Statistical inference for generalized additive models: Simultaneous confidence corridors and variable selection. *TEST* **25** 607–626.

Zheng, S., Yang, L., and Härdle, W. (2014). A smooth simultaneous confidence corridor for the mean of sparse functional data. *J. Amer. Statist. Assoc.* **109** 661–673.

Zhong, C., and Yang, L. (2023). Statistical inference for functional time series: Autocovariance function. *Statist. Sinica* **33** 2519–2543.

Department of Statistics and Data Science, Tsinghua University, Beijing 100084, China.

E-mail: hqr20@mails.tsinghua.edu.cn

Department of Statistics and Data Science, Tsinghua University, Beijing 100084, China.

E-mail: yanglijian@tsinghua.edu.cn