AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 8(7): 15844–15875.
DOI: 10.3934/math.2023809
Received: 12 January 2023
Revised: 02 April 2023
Accepted: 11 April 2023
Published: 04 May 2023

Research article

kNN local linear estimation of the conditional density and mode for functional spatial high dimensional data

Fatimah Alshahrani[1], **Wahiba Bouabsa**[2], **Ibrahim M. Almanjahie**[3] **and Mohammed Kadi Attouch**[2,*]

[1] Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

[2] Laboratory of Statistics and Stochastic Processes, University of Djillali Liabes, BP 89, Sidi Bel Abbes 22000, Algeria

[3] Department of Mathematics, College of Science, King Khalid University, Abha 62223, Saudi Arabia

* **Correspondence:** Email: attou_kadi@yahoo.fr.

Abstract: Traditionally, regression problems are examined using univariate characteristics, including the scale function, marginal density, regression error, and regression function. When the correlation between the response and the predictor is reasonably straightforward, these qualities are helpful and instructive. Given the predictor, the response's conditional density provides more specific information regarding the relationship. This study aims to examine a nonparametric estimator of a scalar response variable's function of a density and mode, given a functional variable when the data are spatially dependent. The estimator is then derived and established by combining the local linear and the k nearest neighbors methods. Next, the suggested estimator's uniform consistency in the number of neighbors (UNN) is proved. Finally, to demonstrate the efficacy and superiority of the acquired results, we applied our new estimator to simulated and real data and compared it to the existing competing estimator.

Keywords: k nearest neighbors; spatial functional data analysis; local linear estimation; conditional density; conditional mode
Mathematics Subject Classification: 62H12, 62G07, 62G35, 62G20

1. Introduction

The most prominent topic of covariance in mathematical statistics is the statistical investigation between two random variables. Typically, regression modeling is employed to illustrate this relationship. However, there are several situations in which this strategy is not particularly beneficial (for example, see Collomb et al. [1]). In this study, we investigate the conditional mode, an alternative strategy that maximizes conditional density. Precisely, we focus on estimating the conditional density and mode function using the local linear approach weighted by the k nearest neighbors (kNN) smoothing methodology. The functional local linear estimating (LLE) problem has been the subject of numerous studies since the publication of the monograph paper (Ferraty and Vieu [2]) in non-parametric functional data analysis (NFDA). The first response to this query occurred from Ballo and Grané [3], who showed the $L2$-consistency of an LLE's Hilbertian regression function. Barrientos et al. [4] then offered a different response that could be used as a more versatile functional regressor. On the other hand, Demongeot et al. [5] developed the preliminary results on the LLE of conditional density using functional variables and established the almost-complete consistency (a.co.) of an LLE of the conditional density. Then, the LLE's mean quadratic error of the conditional density in Rachdi et al. [6] was explicitly mentioned as the leading term. In addition, Zhou and Lin [7] established the local linear asymptotic normality of the regression function. Recently, Almanjahie et al. [8] treated the kNN LLE of the conditional density in a scalar-on-function regression framework. For further information on the LLE in NFDA, check, for instance, Belarbi et al. [9], Chahad et al. [10] and Attouch et al. [11].

It should be mentioned that the LLE's significant advantage over the classic kernel technique is that it serves as the primary inspiration for all the research on functional LLE that has been referenced. In particular, as we know, the local linear approach reduces the kernel method's bias error. (For additional information on the benefits of this approach, see, for example, Fan and Gijbels [12] and Rachdi et al. [6]). The estimate of the conditional density/mode for functional and spatial data using the LLE approach paired with the kNN smoothing process is the main novelty of the current study, which we will refer to as spatial-kNN-LEE. In contrast, most research in this field employs the kernel estimation method to estimate nonparametric functional models. In addition, combining these two techniques with spatial data makes it possible to build an estimator that is less biased and more attractive. Indeed, it is generally known that the kNN technique allows for selecting a parameter for a smoothing bandwidth that is more appropriate due to the data's local nature; in fact, this estimate may be upgraded to account for any more additional notes. This estimator form is similarly robust, rapidly converges, and is simple to create and apply in reality.

The spatial data issue has attracted a lot of interest. The local linear approach is thoroughly discussed by Hallin et al. [13,14]. The quantile estimates and the local linear regression estimation were given illustrations, together with the asymptotic normality results. To our knowledge, the functional spatial local linear estimating scenario has only a few outcomes. For the regression model, we remember that the contribution of the most original was made by Chouaf and Laksaci [15], which focused on spatial LLE functional data. After that, the estimation of the conditional models utilizing the local linear approach was demonstrated by Laksaci et al. [16] when the data are spatially dependent and functional in structure.

Gheriballah et al. [17] presented the M-estimation spatial form of the regression function and attained the asymptotic normality result. After that, the asymptotic normality spatial relative error regression was proved by Attouch et al. [18]. Abeidallah et al. [19] then addressed the issue of estimating specific conditional models using the LLE approach. More recent advances and references in functional nonparametric estimation can be found in Rachdi et al. [20], who proved the parametric and nonparametric conditional quantile regression modelization, which focused on dependent spatial functional; then, the expectile regression for spatial functional data analysis was made by Rachdi et al. [21].

The current study, motivated by previous research, focuses on estimating the functional conditional density using dependent spatial data and combining the LLE methodology and the *k*NN smoothing method. Hence, for our proposed estimator, we examine the almost complete convergence rate.

There are various studies on estimating conditional densities; we quote, for instance, Amiri and Dabo-Niang [22], who analyzed the recursive density kernel structure. They investigated the spatial asymptotic behavior of the built estimator. We refer to Giraldo and Dabo-Niang [23] for a complete discussion of spatial functional data analysis. The nonparametric functional modal regression's uniform consistency is shown by Chaouch and Laib [24], taking into account the ergodicity structure. For functional space, the continuous time processes of the conditional density's asymptotic properties were established by Maillot and Chesneau [25] and determined the expectation and convergence of the conditional mode. In order to develop a useful cross-validation approach for the selection of the bandwidth parameter, Kirkby and Leitao [26] studied an alternative estimate of the conditional density.

The research on the estimation of the *k*NN approach in NFDA is still sparse and focuses on the regression model. However, because of their adaptability and effectiveness, nonparametric *k*NN techniques obtained much interest in the statistical literature. The functional *k*NN smoothing approach has recently received more interest due to its favorable properties. Since Cover [27], which pioneered work in this area, many publications have been published in a variety of estimating contexts, including discrimination, density, and mode estimates, in addition to clustering analysis. The most recent references are listed in the citations: Kudraszow and Vieu [28], Kara et al. [29], Attouch [30], Almanjahie et al. [31], Bouabsa [32], Almanjahie et al. [8], Bouabsa [33].

In this study, we establish the almost complete convergence under a few general assumptions (with rate). Our current work is the connection between Almanjahie et al. [8] study and Kara et al. study's [29]. Remember, for instance, that a random variable serves as the bandwidth parameter in the *k*NN technique; this shows the difficulty and interest of our study. The distance amidst the random functional variables is precisely used to construct the bandwidth parameter. It enables the topological and spectral components of the data to be investigated. To create complex details of the conditional density and mode functional estimators, we merge all advantages, techniques, and methods in this paper. Therefore, the goal is to mix the two strategies with the functional spatial dependency data, starting with their respective benefits. As we stated earlier, the innovation lies in the combination of the technique of LLE with the smoothing method *k*NN to give a brand new estimator of the functional conditional density with spatial data. With this combination, we can solve the bandwidth selection issue and achieve a desirable estimator with a slight bias. However, our new estimator has two significant difficulties: (1) On the other hand, in the strategy of the kernel, a fixed scalar is the deterministic parameter, and in the *k*NN method, the bandwidth parameter is a random variable, making it more difficult to examine its asymptotic characteristics; (2) While the optimal number of neighbors varied depending on the data in practice, the bandwidth parameter's number of neighbors is deterministic. To cover the practical case, it is insufficient to focus only on determining a standard asymptotic characteristic with a fixed number of neighbors. Hence and to incorporate the actual case, the primary goal of this work, with respect to the number of neighbors, is to establish the almost complete convergence of our new estimator.

In order to favor connections between the FDA and high-dimensional statistics, We have included a few references on this last topic, including the work by Bodnar et al. [34]. However, other papers are at the cross-roads between the two fields and are therefore highly representative of the wide variety of connections between FDA and "Big Data". In the papers by Gao et al. [35] and Berrendero et al. [36], high-dimensional methods are adapted to FDA. The relationship in Aaron et al. [37] comes from a general modeling technique that gives a unified perspective of the two fields. The high-dimensional characteristic of the work, in the contribution by Gao et al. [35], lies in the fact that the statistical samples are formed of a large number of variables, each of which is possibly infinite-dimensional. The suggested technique depends on the FDA as well as the high-dimensional statistical literature. The infinite-dimensional variables are reduced to vectors using functional principal component analysis (FPCA) (see for more details and explication [38, 39]), which are then processed using high-dimensional factor analysis approach. Asymptotics are defined for this two-stage process and simulation studies compare its behavior in finite samples to alternative multi-functional dependent analysis methods.

The content of the article is as described in the following. In the next section, we discuss our estimator. Then, we present the density and mode estimators' hypotheses and asymptotic outcomes in Section 3. Then, we exhibit our findings as well as implementation using a real-world data sample in Section 4, Then, in Section 5, an appendix describing the proofs of the auxiliary results are provided. Finally, our conclusion is stated in the last section.

2. Models and estimator

2.1. Spatial data presentation

To introduce the spatial functional form of the kNN conditional density estimator with local linear estimation, assuming denote with $X_{\mathbf{i}} = (Z_{\mathbf{i}}, W_{\mathbf{i}})$ the measurable process and strictly stationary technique, with $\mathbf{i} \in \mathbb{Z}^M$ and $M \geq 1$ selected from $X = (Z, W)$ random variables, which is valued in $(\mathcal{F} \times \mathbb{R})$, where we define a semi-metric space by $(\mathcal{F}, d)$. We denote by z a fixed curve in $\mathcal{F}$, and we indicate to the neighborhood of z by N_z, and to a fixed compact subset of $\mathbb{R}$ by ϱ. Then, we suppose that the $X_{\mathbf{i}}$'s are (i.d.), which means identically distributed to $X = (Z, W)$ and that the conditional probability of W given Z has a regular form. Furthermore, suppose that $\mathbf{i} = (i_1, ..., i_M) \in \mathbb{Z}^M$ is a point known as a site.

We suppose that the process is accessible on the set

$$\mathcal{L}_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_M) \in \mathbb{Z}^M, \quad 1 \leq i_\vartheta \leq n_\vartheta, \quad \vartheta = 1, ..., M\}, \text{ where } \mathbf{n} = (n_1, ..., n_M) \in \mathbb{Z}^M.$$

When there is no possibility of confusion, we indicate any generic positive constants by C and C', and $\mathcal{S}$ a compact of $\mathbb{R}$ throughout the paper, and we write

$$\text{if } \min_{\vartheta=1,...,M} \{n_\vartheta\} \to \infty \text{ and } |n_{\mathbf{j}}/n_\vartheta| < C \ \forall \mathbf{j}, \vartheta \in \{1, ..., M\}, \text{ we have } \mathbf{n} \to \infty.$$

Since it permits increasing the region of observations while keeping the distance between observation positions to a minimum, this is regarded as an asymptotically increasing area.

When the functional random field $X_{\mathbf{i}} = (Z_{\mathbf{i}}, W_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{Z}^M$ fulfills the consequent mixing suppositions, this study objective is to examine the kNN spacial conditional density f^z with LLE:

$$\begin{cases} \text{when } \upsilon \to \infty, \text{ there is a function } \Phi(\upsilon) \downarrow 0, \text{as well as } \forall G, G' \text{a part of } \mathbb{Z}^M \text{cardinal finite} \\ \alpha\left(\mathcal{B}(G), \mathcal{B}(G')\right) = \sup\limits_{H \in \mathcal{B}(G), E \in \mathcal{B}(G')} |\mathbb{P}(H \cap E) - \mathbb{P}(H)\mathbb{P}(E)| \leq \Psi\left(\mathrm{Card}(G), \mathrm{Card}\,(G')\right) \Phi\left(\mathrm{dist}\,(G, G')\right), \end{cases}$$

such that, $\mathcal{B}(G)$ (resp. $\mathcal{B}(G')$) denoted the Borelian tribe produced by $(X_{\mathbf{i}}, \mathbf{i} \in X)$; then, $\mathrm{Card}(G)$ and respectively Card (G') is the number of the cardinal of G (resp. Z'), $\mathrm{dist}\,(Z, Z')$ assign the Euclidean distance among G and G' and Ψ a symmetric nondecreasing positive function: $\mathbb{Z}^2 \to \mathbb{R}^+$, in every variable in such a way that:

$$\Psi(m_1, m_2) \leq C \min(m_1, m_2), \quad m_1, m_2 \in \mathbb{Z}, \tag{2.1}$$

or

$$\Psi(m_1, m_2) \leq C(m_1 + m_2 + 1)^{\tilde{\vartheta}}, \quad m_1, m_2 \in \mathbb{Z}. \tag{2.2}$$

For certain $\widetilde{\vartheta} \geq 1$ and $C > 0$, note that Tran [40] employed these requirements, and that many spatial models can satisfy them (see Guyon [41]).

Take into account that when Eq (2.1) maintains with $\Psi \equiv 1$ or $M = 1$, the random area $X_{\mathbf{i}} = (Z_{\mathbf{i}}, W_{\mathbf{i}})$ is described as strongly mixing (for further details and illustrations of mixing properties see Doukhan [42]). Additionally, let's suppose that the process X fulfills the next mixing criterion:

$$\sum_{\mathbf{i}=1}^{\infty} i^{\vartheta} \Phi(\mathbf{i}) < \infty, \vartheta > 0. \tag{2.3}$$

We point out that the conditions (2.2) and (2.3) are the same as the mixing conditions used by Carbon et al. [43] and Tran [40], and mixing conditions utilized by them are identical to the conditions (2.2) and (2.3).

2.2. *Methodology: notes & required general background*

With the spatial-kNN-LLE technique, this study aims to determine and estimate the conditional probability density. The purpose of this method is to approximate the function $f(.|z)$ linearly. Then, the local linear method is utilized using the Taylor expansion of $f(.|z)$ in the vectorial situation. Such extension is not obtainable in the functional data analysis. Instead, we assume that

$$\forall z_0 \in N_z, \forall w \in \mathbb{R}, f(w|z_0) = a(w|z) + b(w|z) d(z_0, z) + o(d(z_0, z)). \tag{2.4}$$

Now, the spatial-kNN-LLE of the functionals $\hat{a}(w|z)$ and $\hat{b}(w|z)$ is derived by the optimization rule

$$(\hat{a}, \hat{b}) = \arg\min_{(a,b) \in \mathbb{R}} \sum_{\mathbf{i} \in \mathcal{L}_{\mathbf{n}}} (b_\delta^{-1} \delta(b_\delta^{-1}(w - W_{\mathbf{i}})) - a - b\chi(Z_{\mathbf{i}}, z))^2 L(h_L^{-1} \wp(z, Z_{\mathbf{i}})), \tag{2.5}$$

where the locating functions $\chi(.;.)$ and $\wp(.;.)$ defined from $\mathcal{F}^2$ into $\mathbb{R}$, as follows:

$$d(.;.) = |\wp(.;.)| \; \textit{and} \; \chi(\xi; \xi) = 0, \; \forall \xi \in \mathcal{F}.$$

The distribution function L, the kernel function δ, and the (h_L, b_δ) are the kNN smoothing parameters, defined by

$$h_L = \min\{h \in \mathbb{R}^+, \text{ such that } \sum_{\mathbf{i}\in\mathcal{L}_\mathbf{n}} \mathbf{1}_{B(z,h_L)}(Z_\mathbf{i}) = L\},$$

and

$$b_\delta = \min\{h \in \mathbb{R}^+, \text{ such that } \sum_{\mathbf{i}\in\mathcal{L}_\mathbf{n}} \mathbf{1}_{w-b,w+b}(W_\mathbf{i}) = \delta\}.$$

When take $z_0 \in N_z$, we obtain $f(.|z) = a(w|z)$. Then, with respect to z, $b(w|z)$ is the derivative form of $f(.|z)$. So, $f(.|z)$ and $b(w|z)$ are estimated by (2.5). Both estimators are explicit. Indeed, we have

$$^{t}\mathfrak{L} = \begin{pmatrix} 1, 1, \ldots, 1 \\ \wp(Z_1, z) \ldots, \wp(Z_\mathbf{n}, z) \end{pmatrix}, \quad \mathbb{L} = \begin{pmatrix} L\left(h_L^{-1}\wp(z, Z_1)\right), 0, \ldots, 0 \\ 0, \ldots, 0, L\left(h_L^{-1}\wp(z, Z_\mathbf{n})\right) \end{pmatrix},$$

and

$$\mathbb{W} = \begin{pmatrix} b_\delta^{-1}\delta\left(b_\delta^{-1}(w - W_1)\right) \\ \vdots \\ b_\delta^{-1}\delta\left(b_\delta^{-1}(w - W_\mathbf{n})\right) \end{pmatrix}.$$

To demonstrate that the minimizing of (2.5) using the partial derivative are zeros of

$$^{t}\mathfrak{L}\left(\mathbb{L}\mathbb{W} - \mathbb{L}\mathfrak{L}\begin{pmatrix} \widehat{a}(w \mid z) \\ \widehat{b}(w \mid z) \end{pmatrix}\right) = 0,$$

it follows

$$\begin{pmatrix} \hat{a}(w \mid z) \\ \widehat{b}(w \mid z) \end{pmatrix} = \left(^{t}\mathcal{L}\mathbb{L}\mathfrak{L}\right)^{-1}\left(^{t}\mathfrak{L}\mathbb{L}\mathbb{W}\right).$$

Consequently,

$$\widehat{a}(w \mid z) = (1, 0)\left(^{t}\mathcal{L}\mathbb{L}\mathcal{L}\right)^{-1}\left(^{t}\mathfrak{L}\mathbb{L}\mathbb{W}\right).$$

As a result, the spatial-kNN-LLE estimator for $f(w|z)$ is written as

$$\tilde{f}(w \mid z) = \frac{\sum_{\substack{\mathbf{i},\mathbf{j}\in\mathcal{L}_\mathbf{n} \\ \mathbf{i}\neq\mathbf{j}}} Y_{\mathbf{ij}}(z)\delta\left(h_\delta^{-1}\left(w - W_\mathbf{j}\right)\right)}{b_\delta \sum_{\substack{\mathbf{i},\mathbf{j}\in\mathcal{L}_\mathbf{n} \\ \mathbf{i}\neq\mathbf{j}}} Y_{\mathbf{ij}}(z)}, \tag{2.6}$$

where

$$Y_{\mathbf{ij}}(z) = \chi(Z_\mathbf{i}, z)\left(\chi(Z_\mathbf{i}, z) - \chi\left(Z_\mathbf{j}, z\right)\right) \times L\left(h_L^{-1}\wp(z, Z_\mathbf{i})\right)L\left(h_L^{-1}\wp\left(z, Z_\mathbf{j}\right)\right).$$

3. kNN estimators consistency

With regard to the number of neighbors $(L, \delta) \in (L_{1,\mathbf{n}}, L_{2,\mathbf{n}}) \times (\delta_{1,\mathbf{n}}, \delta_{2,\mathbf{n}})$, we aim to establish the uniformly a.co. consistency of $\tilde{f}(w \mid z)$. To do this, we define the minimum number of open balls required to cover the set of functions $\mathcal{D}$ with regard to $M_2(S)$ as $\mathcal{N}(\varepsilon, \mathcal{D}, \|.\|_{M_2(S)}$. Additionally, we define $\|\varphi\|_\mathcal{D} = \sup_{q\in\mathcal{D}}|\varphi(q)|$.

3.1. The conditional density estimator consistency

Our main finding is linked to the almost complete convergence rate of the spatial-kNN-LLE conditional density's. To be able to accomplish our outcomes, we present the hypotheses listed below that are needed to validate the results of this section.

- **(A1):**
  - **(A1a)** $\forall g > 0, \mathbb{P}(z \in B(Z, g)) := \zeta_z(g) > 0$.
  - **(A1b)** $\lim_{g\to 0} \frac{\zeta_z(eg)}{\zeta_z(g)} = \tau_z(e), \forall e \in (0, 1)$ such that

$$L\left(\frac{1}{2}\right)\tau_z\left(\frac{1}{2}\right) - \int_0^{\frac{1}{2}} L'(e)\tau_z(e)de > 0,$$

and

$$\left(\frac{1}{4}\right)L\left(\frac{1}{2}\right)\tau_z\left(\frac{1}{2}\right) - \int_0^{\frac{1}{2}} \left(e^2L(e)\right)' \tau_z(e)ds > 0.$$

- **(A2):** For some $0 < \nu < M^{-1}$,

$$0 < \sup_{\mathbf{i}\neq\mathbf{j}} \mathbb{P}\left[\left(Z_{\mathbf{i}}, Z_{\mathbf{j}}\right) \in B(z, h_L) \times B(z, h_L)\right] \leq C\left(\zeta_z(h_L)\right)^{(\nu+1)/\nu}.$$

- **(A3):** The conditional $f(.|z)$ is the probability density such that $\exists\varpi_1, \varpi_2 \in (0, 1], \forall (z_1, z_2) \in N_z \times N_z$ and $(w_1, w_2) \in \mathbb{R}^2$,

$$|f(w_1 \mid z_1) - f(w_2 \mid z_2)| \leq C\left(\wp^{\varpi_1}(z_1, z_2) + |w_1 - w_2|^{\varpi_2}\right).$$

- **(A4):** $d(\cdot, \cdot)$ is the function written in such a way that

$$\forall z_1, z_2 \in \mathcal{F}, C'|\wp(z_1, z_2)| \leq |\wp(z_1, z_2)| \leq C|\wp(z_1, z_2)|.$$

- **(A5):** We have a positive kernel L and differentiable function with support for $(-1, 1)$.
- **(A6):** δ is a differentiable function with a continuous derivative, δ' is a bounded function, such that

$$\int |r|^{b_2}\delta'(r)dr < \infty, \int \delta'^2(r)dr < \infty.$$

Additionally,

$$\forall (w_1, w_2) \in \mathbb{R}^2, |\delta'(w_1) - \delta'(w_2)| \leq C|w_1 - w_2|.$$

- **(A7):** $\alpha < (k - 5M)/2M$ and η_0 are positive numbers, expressed as

$$\lim_{\mathbf{n}\to\infty} b_\delta = 0, \quad \lim_{\mathbf{n}\to\infty} \hat{\mathbf{n}}^\alpha b_\delta = \infty,$$

and

$$C\hat{\mathbf{n}}^{\frac{(5+2\alpha)M-k}{k}+\eta_0} \leq \zeta_z(h_L),$$

where $\hat{\mathbf{n}} = n_1 \times \cdots \times n_M$.

- **(A8):**
  - **(A8a)** $\forall \imath = 0, 1, 2$ and $\ell = 0, 1$,

$$\mathcal{K}^{\imath,\ell} = \{(\cdot_1, \cdot_2) \mapsto \gamma^{-\imath} L\left(\gamma^{-1} d(z, \cdot_1)\right) d^{\imath}(\cdot, z)\gamma^{-\ell}\delta^{\ell}\left(\gamma^{-1}(w - \cdot_2)\right), \ \lambda > 0, \ \gamma > 0\}$$

are pointwise measurable classes.

  - **(A8b)** $\mathcal{K}^{\ell,\kappa}$ is such that

$$\sup_{S} \int_0^1 \sqrt{1 + \log \mathcal{N}\left(\epsilon \left\| |F^{\imath,\ell}\right\| |_{M_2(S)}, \mathcal{K}^{\imath,\ell}, \| \cdot \|_{M_2(S)}\right)} d\epsilon < \infty.$$

We indicate the envelope function of the class $\mathcal{K}^{\imath,\ell}$ by $F^{\imath,\ell}$, and call $\| \cdot \|_{M_2(S)}$ the $M_2(S)$ norm.

- **(A9):** The next sets $(L_{1,\mathbf{n}}), (L_{2,\mathbf{n}}), (\delta_{1,\mathbf{n}})$ and $(\delta_{2,\mathbf{n}})$, check: $\frac{\delta_{2,\mathbf{n}}}{\mathbf{n}} \to 0, \zeta_z^{-1}\left(\frac{L_{2,\mathbf{n}}}{\mathbf{n}}\right) \to 0$ and for some $\vartheta > 0$, we have

$$\frac{\mathbf{n} \log \mathbf{n}}{\delta_{1,\mathbf{n}} \min\left(\mathbf{n}\zeta_z^{-1}\left(\frac{L_{1,\mathbf{n}}}{\mathbf{n}}\right), L_{1,\mathbf{n}}\right)}, \mathbf{n}^{\vartheta-1}\delta_{1,\mathbf{n}} \to \infty.$$

Remark 3.1. Comments on the hypotheses: *In the asymptotic theory of nonparametric functional statistics, assumption (A1) is a prerequisite. The first part of this hypothesis is the explanatory variable's concentration in small balls. Numerous publications on nonparametric functional statistics also consider the second part. To achieve the same convergence rate as in the independence case indicated by condition (A2), the local dependency between the observations must hold. The hypothesis of Lipschitz (A3) is applied to the conditional density function, demonstrating that each variable has a continuous relationship with the function. The assumption (A4) is identical to the one employed by Barrientos-Marin et al. [4]. Nonparametric function estimation frequently uses condition (A5). For specific examples of kernels satisfying (A5) and (A6), we can see Ferraty and Vieu [2]. Hypothesis (A6) imposes certain regularity constraints on the kernels δ and δ' used in our findings. According to Almanjahie et al. [8], conditions (A7) and (A9) are similar. An easy prerequisite is the assumption (A8) on the bandwidth parameters and certain technical requirements that make the demonstration simpler to comprehend.*

Theorem 3.1. *Under the assumptions (A1–A9), we obtain*

$$\sup_{L_{1,\boldsymbol{n}} \le L \le L_{2,\boldsymbol{n}}} \sup_{\delta_{1,\boldsymbol{n}} \le \delta \le \delta_{2,\boldsymbol{n}}} \sup_{w \in S} | \widetilde{f}(w \mid z) - f(w \mid z) | = O\left(\zeta_z^{-1}\left(\frac{L_{2,\boldsymbol{n}}}{\boldsymbol{n}}\right)^{\varpi_1}\right) + O\left(\frac{\delta_{2,\boldsymbol{n}}}{\boldsymbol{n}}\right)^{\varpi_2} + O_{a.co}\left(\sqrt{\frac{\boldsymbol{n} \log \boldsymbol{n}}{\delta_{1,\boldsymbol{n}} L_{1,\boldsymbol{n}}}}\right).$$

Proof. For a fixed $\alpha \in]0, 1[$, we have

$$y_{\mathbf{n}} = \zeta_z^{-1}\left(\frac{L_{2,\mathbf{n}}}{\mathbf{n}}\right)^{\varpi_1} + \left(\frac{\delta_{2,\mathbf{n}}}{\mathbf{n}}\right)^{\varpi_2} + \left(\sqrt{\frac{\mathbf{n} \log \mathbf{n}}{\delta_{1,\mathbf{n}} L_{1,\mathbf{n}}}}\right),$$

where

$$\hat{a}_{\mathbf{n}} = \zeta_z^{-1}\left(\frac{\alpha L_{1,\mathbf{n}}}{\mathbf{n}}\right), \ \hat{b}_{\mathbf{n}} = \zeta_z^{-1}\left(\frac{L_{2,\mathbf{n}}}{\mathbf{n}\alpha}\right), \ \hat{c}_{\mathbf{n}} = \frac{\alpha \delta_{1,\mathbf{n}}}{\mathbf{n}} \text{ and } \hat{d}_{\mathbf{n}} = \frac{\delta_{2,\mathbf{n}}}{\mathbf{n}\alpha}.$$

Then, $\forall \varepsilon > 0$, we get

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{L_{1,\mathbf{n}}\leq L\leq L_{2,\mathbf{n}}}\ \sup_{\delta_{1,\mathbf{n}}\leq\delta\leq\delta_{2,\mathbf{n}}}\ \sup_{w\in S}|\tilde{f}(w\mid z)-f(w\mid z)|\geq \varepsilon y_{\mathbf{n}}\right\}\\
\leq&\mathbb{P}\left\{\sup_{L_{1,\mathbf{n}}\leq L\leq L_{2,\mathbf{n}}}\ \sup_{\delta_{1,\mathbf{n}}\leq\delta\leq\delta_{2,\mathbf{n}}}\ \sup_{w\in S}|\tilde{f}(w\mid z)-f(w\mid z)|\times\mathbf{I}_{\left\{\hat{a}_{\mathbf{n}}\leq h_L\leq\zeta_z^{-1}\left(\frac{L_{2,\mathbf{n}}}{\mathbf{n}\alpha}\right),\frac{\alpha L_{1,\mathbf{n}}}{\mathbf{n}}\leq b_\delta\leq\frac{\delta_{2,\mathbf{n}}}{\mathbf{n}\alpha}\right\}}\geq\frac{\varepsilon y_{\mathbf{n}}}{2}\right\}\\
&+\mathbb{P}\left\{h_L\notin\left(\zeta_z^{-1}\left(\frac{\alpha L_{1,\mathbf{n}}}{\mathbf{n}}\right),\zeta_z^{-1}\left(\frac{L_{2,\mathbf{n}}}{\mathbf{n}\alpha}\right)\right)\right\}+\mathbb{P}\left\{b_\delta\notin\left(\frac{\alpha\delta_{1,\mathbf{n}}}{\mathbf{n}},\frac{\delta_{2,\mathbf{n}}}{\mathbf{n}\alpha}\right)\right\}.
\end{aligned}$$

Kara et al. [29] studied the latest two probabilities. As a result, just the first one has to be examined. This latter is mostly the result of the following proposition.

Proposition 3.1. *Based on the same Theorem 3.1 hypotheses, we have*

$$\sup_{\hat{a}_n\leq h_L\leq\hat{b}_n}\ \sup_{\hat{c}_n\leq b_\delta\leq\hat{d}_n}|\tilde{f}(w\mid z)-f(w\mid z)|=O\left(\hat{b}_{\boldsymbol{n}}^{\varpi_1}\right)+O\left(\hat{d}_{\boldsymbol{n}}^{\varpi_2}\right)+O_{a.co}\left(\sqrt{\frac{\log\hat{\boldsymbol{n}}}{\hat{\boldsymbol{n}}\hat{c}_{\boldsymbol{n}}\zeta_z(\hat{a}_{\boldsymbol{n}})}}\right).\tag{3.1}$$

Proof. The demonstration is focused on decomposition that follows and the corollary and lemmas below:

$$\tilde{f}(w\mid z)-f(w\mid z)=\tilde{B}_{\mathbf{n}}(w\mid z)+\frac{\tilde{R}_{\mathbf{n}}(w\mid z)}{\tilde{f}_D(z)}+\frac{\tilde{Q}_{\mathbf{n}}(w\mid z)}{\hat{f}_D(z)},$$

where

$$\tilde{Q}_{\mathbf{n}}(w\mid z)=\left(\tilde{f}_N(w\mid z)-\mathbb{E}\tilde{f}_N(w\mid z)\right)-f(w\mid z)\left(\tilde{f}_D(z)-\mathbb{E}\tilde{f}_D(z)\right),$$

$$\tilde{B}_{\mathbf{n}}(w\mid z):=\frac{\mathbb{E}\tilde{f}_N(w\mid z)}{\mathbb{E}\tilde{f}_D(z)}-f^z(w)\quad\text{ and }\quad\hat{R}_{\mathbf{n}}(w\mid z):=-\widehat{B}_{\mathbf{n}}(w\mid z)\left(\hat{f}_D(z)-\mathbb{E}\tilde{f}_D(z)\right),$$

with

$$\tilde{f}_N(w\mid z)=\frac{1}{\hat{\mathbf{n}}(\hat{\mathbf{n}}-1)b_\delta\mathbb{E}[Y_{12}(z)]}\sum_{\mathbf{i}\neq\mathbf{j}}Y_{\mathbf{ij}}(z)\delta'\left(b_\delta^{-1}(w-W_{\mathbf{j}})\right),$$

$$\tilde{f}_D(z):=\frac{1}{\hat{\mathbf{n}}(\hat{\mathbf{n}}-1)\mathbb{E}[Y_{12}(z)]}\sum_{\mathbf{i}\neq\mathbf{j}}Y_{\mathbf{ij}}(z).$$

Thus, Proposition 3.1 is a consequence of the following corollary and lemmas, where their proofs are included in the Appendix.

Corollary 3.1. *There exists $u>0$ under the assumptions of Theorem 3.1, where*

$$\sum_{\mathbf{n}\in\mathbb{Z}^M}\mathbb{P}\left\{\sup_{\hat{a}_{\mathbf{n}}\leq h_L\leq\hat{b}_{\mathbf{n}}}\ \sup_{\hat{c}_{\mathbf{n}}\leq b_\delta\leq\hat{d}_{\mathbf{n}}}\left|1-\tilde{f}_D(z)\right|<u\right\}<\infty.$$

Lemma 3.1. *Similar to the condition of Theorem 3.1, we have*

$$\sup_{w\in S}\ \sup_{\hat{a}_n\leq h_L\leq\hat{b}_n}\ \sup_{\hat{c}_n\leq b_\delta\leq\hat{d}_n}\left|\frac{\mathbb{E}\left[\widetilde{f_N}(w\mid z)\right]-f(w\mid z)\mathbb{E}\left[\tilde{f}_D(z)\right]}{\mathbb{E}\left[\widetilde{f_D}(z)\right]}\right|=O\left(\hat{b}_{\boldsymbol{n}}^{\varpi_1}\right)+O\left(\hat{d}_{\boldsymbol{n}}^{\varpi_2}\right).$$

Lemma 3.2. *Using the same conditions as in Theorem 3.1, we have*

$$\sup_{w\in S}\ \sup_{\hat{a}_n\leq h_L\leq\hat{b}_n}\ \sup_{\hat{c}_n\leq b_\delta\leq\hat{d}_n}\frac{1}{\tilde{f}_D(z)}\left|\tilde{f}_N(w\mid z)-\mathbb{E}\left[\tilde{f}_N(w\mid z)\right]\right|=O_{a.co}\left(\sqrt{\frac{\log\hat{\boldsymbol{n}}}{\hat{\boldsymbol{n}}\hat{c}_{\boldsymbol{n}}\zeta_z(\hat{a}_{\boldsymbol{n}})}}\right).$$

3.2. kNN conditional mode estimator consistency

The prior results have a direct effect on the spatial-kNN-LLE of the conditional mode estimator convergence rate, defined as

$$\tilde{\theta}(z) = \arg\max_{w \in \mathcal{S}} \widetilde{f}(w \mid z), \tag{3.2}$$

of the conditional mode given by $\theta(z) = \arg\max_{w \in \mathcal{S}} f(w \mid z)$. Therefore, we will add this to the conditions we assumed in Section 3.

- **(A10)**
  - (i1) The $f(. \mid z)$ function is j-times continuous and differentiable with regard to w on the topological interior of $\mathcal{S}$ if $\exists j > 1$ and $\forall z \in \mathcal{N}_z$.
  - (i2) $f(w \mid z(l)) = 0$ when $1 \leq l < j$.
  - (i3) Since $0 < |f(w \mid z(j))| < \infty$, $f(w \mid z(j))$ signifies thej^{th}-the conditional density function order derivative of $f(w \mid z)$, is uniformly continuous on $\mathcal{S}$.

Remark 3.2. *Under the minimal condition in the functional estimation (A10)(i1), (A7) and (A8), hypotheses in finite or infinite dimension spaces are classical assumptions, and the convergence of the mode estimator can be established.*

In the theorem that follows, the asymptotic behavior of $\tilde{\theta}(z)$ is described.

Theorem 3.2. *Under conditions (A1)–(A10), we get*

$$\sup_{L_{1,\boldsymbol{n}} \leq L \leq L_{2,\boldsymbol{n}}} \sup_{\delta_{1,\boldsymbol{n}} \leq \delta \leq \delta_{2,\boldsymbol{n}}} \sup_{w \in \mathcal{S}} |\tilde{\theta}(z) - \theta(z)| = O\zeta_z^{-1} \left(\frac{L_{2,\boldsymbol{n}}}{\boldsymbol{n}}\right)^{\frac{\varpi_1}{j}} + O\left(\frac{\delta_{2,\boldsymbol{n}}}{\boldsymbol{n}}\right)^{\frac{\varpi_2}{j}} + O_{a.co}\left(\sqrt{\frac{\boldsymbol{n} \log \boldsymbol{n}}{\delta_{1,\boldsymbol{n}} L_{1,\boldsymbol{n}}}}\right)^{\frac{1}{j}}.$$

The proof of this theorem is relegated to the Appendix.

4. Applications

In this section, we will compare the kNN local linear technique to the kernel one in a functional spatial setting using both simulated and real data. First, we present the conditional mode as a predictive tool that strongly correlates with conditional density estimation. Second, we suggest a more straightforward and faster approach for implementing the estimator given by (2.6) and demonstrate how its performance is affected by (3.2), the selection of smoothing parameters, and the locating functions. Finally, a practical example using real data is shown to demonstrate the advantages of the local linear estimating method over the kernel estimation approach.

4.1. A simulation study

In this section, we first demonstrate the finite sample behavior of the proposed estimator $\widetilde{\theta}$ by simulating the observations $(Z_{\mathbf{i}}, Y_{\mathbf{i}}) \in (\mathcal{F} \times \mathbb{R})$. For simplicity, we will assume that $N = 2$ and write $\mathbf{i} = (i_1, i_2)$ with $1 \leq i_1 \leq n_1$, $1 \leq i_2 \leq n_2$ and $\forall \mathbf{i} \in \mathbb{Z}^2$. The model was generated using

$$Z_{\mathbf{i}}(t) = \cos(2\pi A_{\mathbf{i}} t) + B_{\mathbf{i}} t, \quad t \in [0, 1],$$

and

$$Y_i = r(Z_{\mathbf{i}}) + \varepsilon_{\mathbf{i}}, \tag{4.1}$$

where $r(Z) = 5.\dfrac{1}{\int_0^1 |Z(t)|dt}$. Then, we designate by $GRF\left(m, \sigma^2, s\right)$ a stationary Gaussian random field with mean m and covariance function defined by

$$C(l) = \sigma^2 \exp\left(-\left(\frac{||l||}{s}\right)^2\right),\ l \in \mathbb{R}^2,\ s > 0.$$

Next, we simulated model (4.1) using the following parameters:

$$A = D * \sin\left(\frac{G}{2} + .5\right),\ B = GRF(2.5, 5, 3),\ \varepsilon = GRF(0, .1, 5),$$
$$G = GRF(0, 5, 3),\ D_{\mathbf{i}} = \frac{1}{n_1 \times n_2} \sum_{\mathbf{j}} \exp\left(-\frac{||\mathbf{i} - \mathbf{j}||}{a}\right).$$

Specifically, we use

$$D_{(\mathbf{i},\mathbf{j})} = \frac{1}{n_1 \times n_2} \sum_{1 \le j_1, j_2 \le 25} \exp\left(-\frac{||(i_1, i_2) - (j_1, j_2)||}{a}\right).$$

The spatial mixing condition is ensured and controlled by the function D (even if using the Gaussian Random Fields also brings some spatial dependency).

The utilized semi-metric is the first sample curves derivative, which is determined by

$$d\left(Z_{\mathbf{i}}, Z_{\mathbf{j}}\right) = \sqrt{\int_0^1 \left(Z'_{\mathbf{i}}(t) - Z'_{\mathbf{j}}(t)\right)^2 dt}, \quad \text{for } \forall Z_{\mathbf{i}}, Z_{\mathbf{j}} \in \mathcal{F}.$$

Furthermore, the kernel function is selected: $K(u) = \frac{3}{2}\left(1 - u^2\right)\mathbb{1}_{[0,1]}(u)$.

Observations of sites $\mathbf{i}$ and $\mathbf{j}$ with $||\mathbf{i} - \mathbf{j}|| < 15$ are spatially dependent and nearly independent from $||\mathbf{i} - \mathbf{j}|| \ge 15$ because the model (under these conditions) is based on Gaussian random fields with covariance function C and scale $s = 5$. As a result, our observations are a mix of dependent and i.i.d. observations (see, Figures 1–4). Therefore, reducing the value of a is necessary to leave from independence (our research is based on $a = 0.5$).

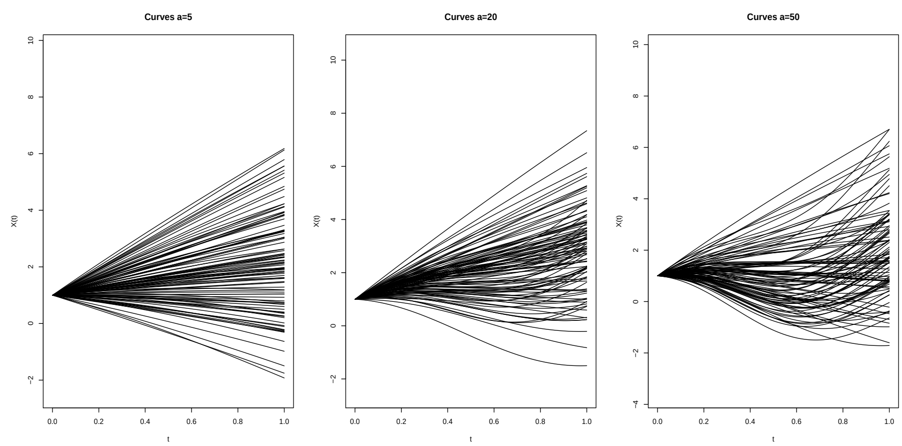

Figure 1. The curves $Z_{\mathbf{i}},\ t \in [0, 1]$ for $a = 5, 20, 50$.

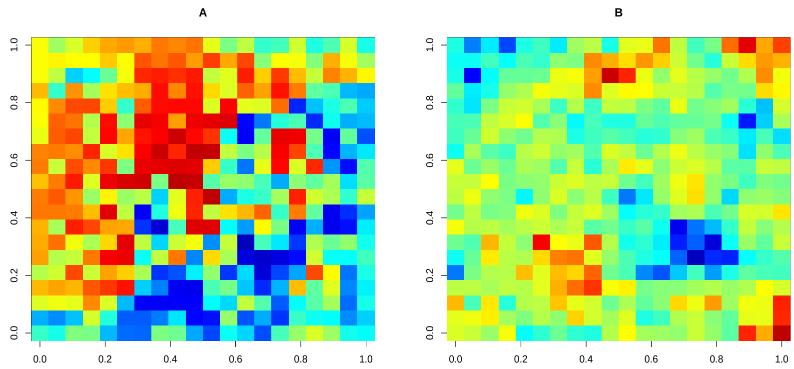

Figure 2. Random field simulation for $a = 5$.

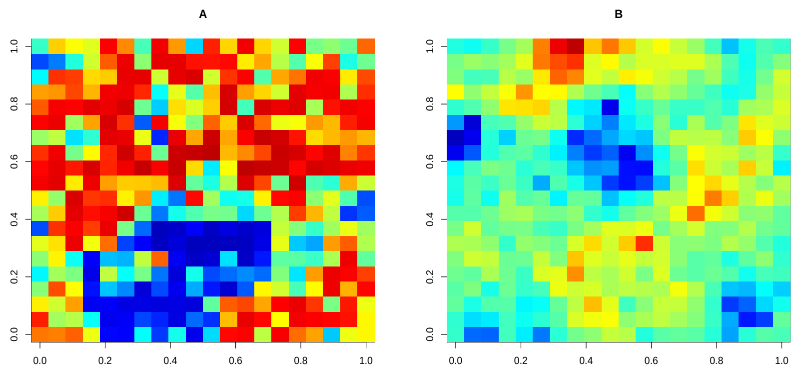

Figure 3. Random field simulation for $a = 20$.

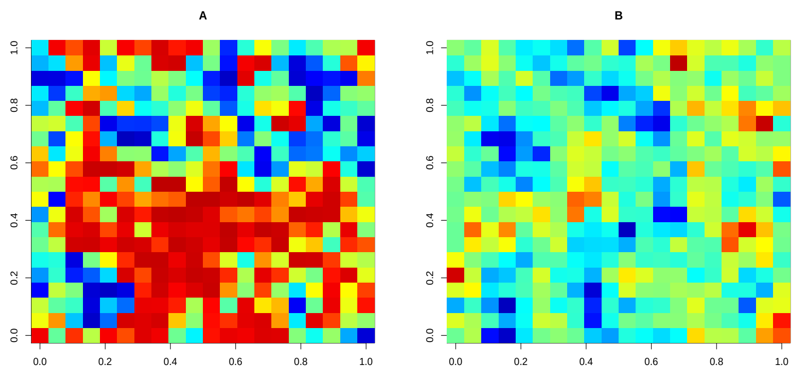

Figure 4. Random field simulation for $a = 50$.

Recall that the main goal is to compare the introduced estimator (spatial-kNN-LLE, $\widetilde{\theta}_{\mathbf{n}}(\omega)$) with the local linear without kNN (spatial-LLE, $\widehat{\theta}_{\mathbf{n}}(\omega)$), and the kernel estimator case (KNW, $\bar{\theta}_{\mathbf{n}}(\omega)$) defined by Dabo-Niang et al. [44]. To test the performance of the proposed estimator, we randomly divided our data $(Z_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}}$ into two subsets: the training sample $(Z_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}\in I}$ and the test sample $(Z_{\mathbf{i}}, Y_{\mathbf{i}}) \in I'$. The training sample was used to calculate the $h_{k_{opt}}$ smoothing parameters for the kNN cross-validation procedures:

$$h_k = \min\left\{ h \in \mathbb{R}^+ \text{such that } \sum_{\mathbf{i}\in I} \mathbb{1}_{B(z,h)}\left(Z_{\mathbf{i}}\right) = k \right\},$$

where $k_{opt} = \arg\min_k CV(k)$ with $CV(k) = \sum_{\mathbf{i}\in I}\left(\theta_i - \check{\theta}_{\mathbf{n}}^{(-\mathbf{i})}\left(Z_{\mathbf{i}}\right)\right)^2$ and $\check{\theta}_{\mathbf{n}}^{(-\mathbf{i})}$ are the leave one out of $\check{\theta}_{\mathbf{n}}$ (see Ferraty and Vieu [2]). The $\check{\theta}_{\mathbf{n}}(\cdot)$ of $\theta(\cdot)$ accuracy was quantified via mean square errors (MSE):

$$\text{MSE} = \frac{1}{\#\left(I'\right)} \sum_{\mathbf{i}\in I'} \left(\check{\theta}_{\mathbf{n}}\left(Z_{\mathbf{i}}\right) - \theta\left(Z_{\mathbf{i}}\right)\right)^2,$$

where $\#(I')$ is the length of testing sample I', and $\check{\theta}_{\mathbf{n}}(\cdot)$ can mean either $\widetilde{\theta}_{\mathbf{n}}(\omega)$, $\widehat{\theta}_{\mathbf{n}}(\omega)$ and $\bar{\theta}_{\mathbf{n}}(\omega)$. Figure 5 depicts the obtained results for the three models, shown as the predicted values against the real values.

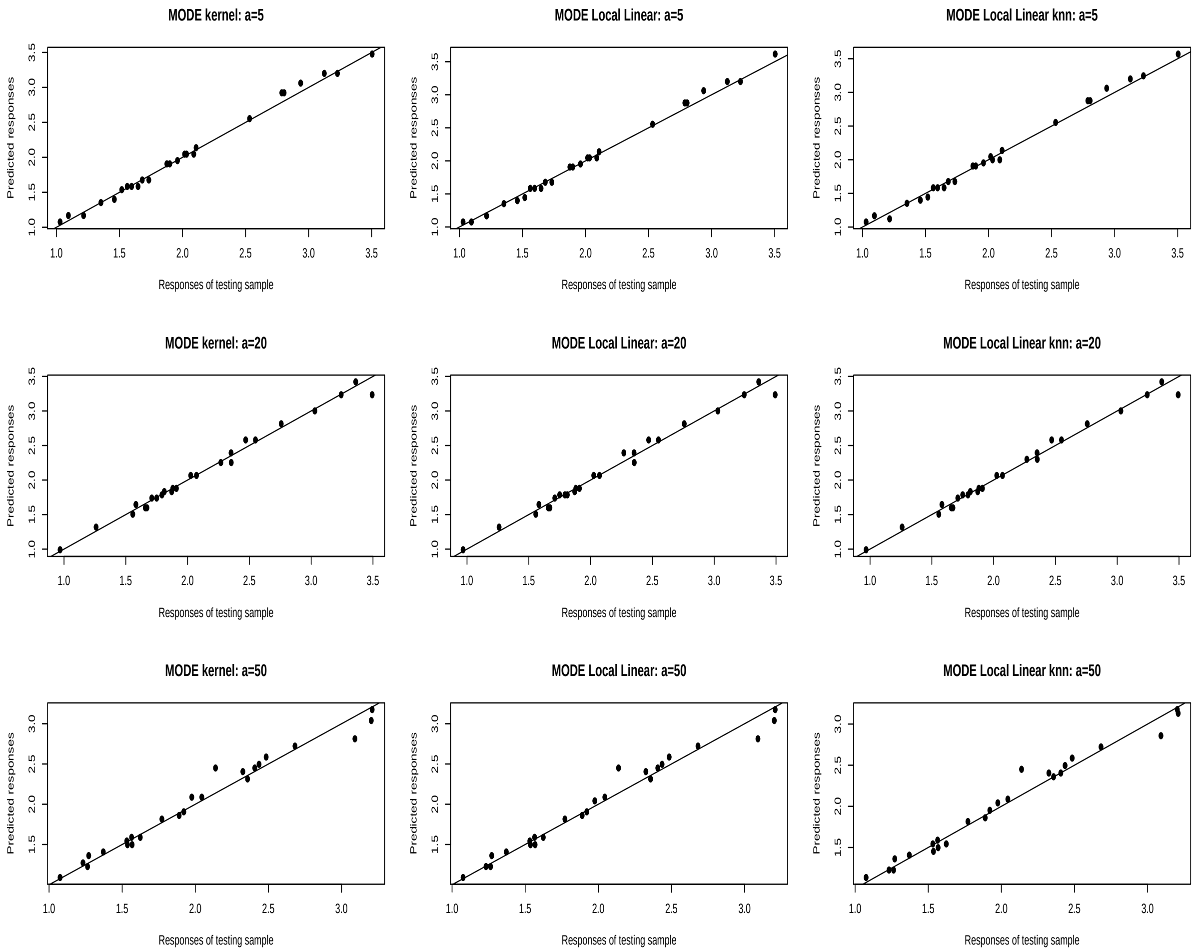

Figure 5. Predictions of the 3 model.

For these three models, we conduct 100 simulations with $a = 5, 20$ and 50 with different sample sizes (see Figure 6, 7 and 8 respectively).

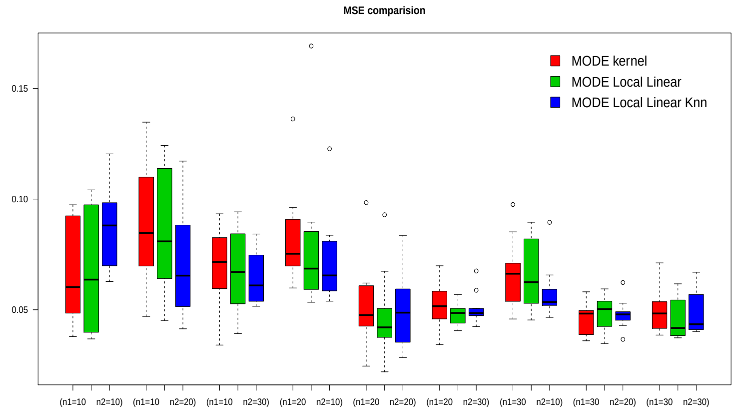

Figure 6. Mean squared error of the 3 models with different values of n_1, n_2 and $a = 5$.

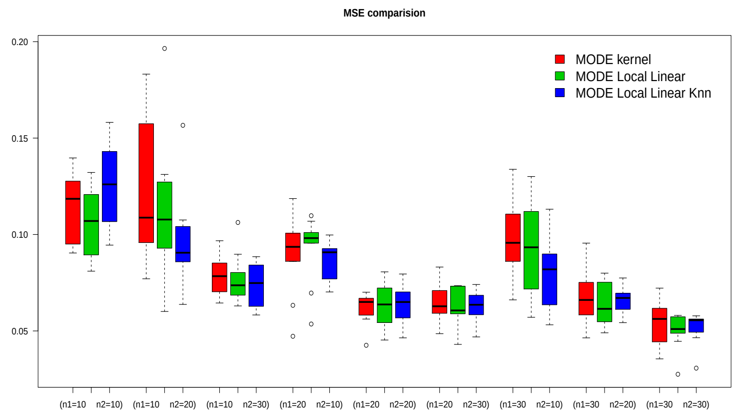

Figure 7. Mean squared error of the 3 models with different values of n_1, n_2 and $a = 20$.

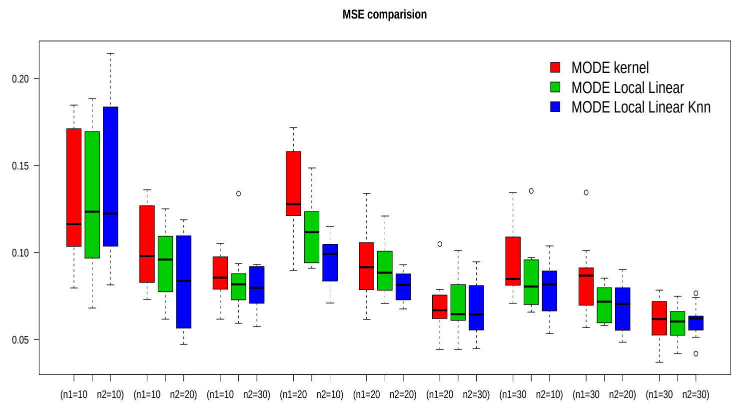

Figure 8. Mean squared error of the 3 models with different values of n_1, n_2 and $a = 50$.

Based on Figures 6–8, we can observe that the values of MSE decrease to 0 as n_1 and n_2 increases.

The MSE under the KNW, LLE, and kNN-LLE are reported in Table 1 for various values of a, n_1 and n_2.

According to Table 1, the kNN-LLE model exhibits superior prediction effects in comparison to the other models.

Table 1. Mean squared error for KNW, LLE and kNN-LLE models respectively.

a		5			20			50		
n_1	n_2	KNW	LLE	kNN-LLE	KNW	LLE	kNN-LLE	KNW	LLE	kNN-LLE
	10	0.0668	0.0677	0.0869	0.1144	0.1073	0.1252	0.1322	0.1292	0.1374
10	20	0.0862	0.0844	0.0725	0.1224	0.1136	0.0948	0.1030	0.0948	0.0831
	30	0.0679	0.0673	0.0636	0.0789	0.0768	0.0733	0.0866	0.0842	0.0789
	10	0.0819	0.0795	0.0716	0.0899	0.0927	0.0862	0.1319	0.1127	0.0951
20	20	0.0501	0.0468	0.0498	0.0618	0.0636	0.0643	0.0929	0.0918	0.0805
	30	0.0519	0.0481	0.0507	0.0650	0.0630	0.0622	0.0688	0.0698	0.0679
	10	0.0658	0.0660	0.0573	0.0996	0.0921	0.0801	0.0938	0.0862	0.0790
30	20	0.0461	0.0479	0.0478	0.0676	0.0634	0.0661	0.0853	0.0711	0.0686
	30	0.0498	0.0455	0.0488	0.0552	0.0501	0.0513	0.0608	0.0595	0.0610

4.2. Real data study

Our aim in this section is to apply the theoretical results from the preceding section to real data. More specifically, we investigate the performance of the proposed estimator using the k nearest neighbors local linear estimating (kNN-LLE) technique in the context of spatial functional prediction using real data application that emphasizes the usefulness of controlling for the spatial locations of the data.

Our real example is related to the chemical concentrations, like ozone O_3, that cause air pollution. Undoubtedly, air pollution is caused by various chemicals influencing air quality and human health. There are a variety of sources for these substances, some of them naturally and others through human industrial activities. In this example, we are interested in predicting the future concentration of O_3 using the past concentration curve. We considered hourly O_3 concentration for this application between January 1st, 2021 and May 2nd, 2021. The data was provided by 131 stations in the United States. Figure 9 displays the locations of 131 stations throughout the United States. This information is accessible at the following website: `https://www.epa.gov/outdoor-air-quality-data`.

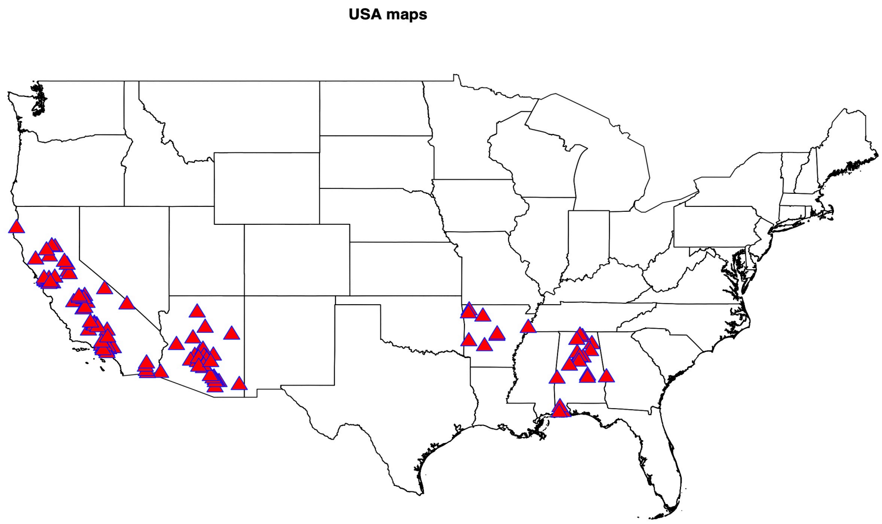

Figure 9. The 131 location stations in the USA.

The following regression equation is assumed to be a link between the observations:

$$Y_s = r(Z_s) + \epsilon, \quad s = 1, 2, ...24,$$

where the response variable can be taken as : $Y_s = O_{3;(s)}$ (for the hour s in the day: May 2^{nd} 2021). For the functional variable, we take : $Z_s = O_{3;(s)}(t); t = s, ..., s + 24$ (for the day: May 1^{st} 2021). According to the notation introduced in the previous section, the functional predictor $Z_{\mathbf{i}}$ is the daily ozone curve in the $\mathbf{i}$th station (identified by its geographic coordinates $\mathbf{i} = (Latitude; Longitude)$), and $Y_{\mathbf{i}}$ is the predicted ozone concentration in the same station. In order to see the spatial stationarity hypothesis in the data, the starting data must be prepared in advance for this form of spatial modeling, as mentioned in Hallin et al. [14]. The latter regulates the spatial heterogeneity connected to the differentiation of space's impacts on the sample units. For the multivariate situation in finite dimension, where the geographical heterogeneity of the two variables (explanatory and response) is modeled by the following regression, we adopt the approach given by Hallin et al. [14] and Rachdi et al. [21] to regulate this aspect:

$$\widetilde{Z_{\mathbf{i}}} = r_1(\mathbf{i}) + Z_{\mathbf{i}}, \;\; \tilde{Y}_{\mathbf{i}} = r_2(\mathbf{i}) + Y_{\mathbf{i}}.$$

So, rather than the initial observations $(Z_{\mathbf{i}}, Y_{\mathbf{i}})$, the conditional mode estimator $\widetilde{\theta_{\mathbf{n}}}$ is computed from the statistics $(Z_{\mathbf{i}}, Y_{\mathbf{i}})$. These results are achieved by

$$\widehat{Z_{\mathbf{i}}} = \widetilde{Z_{\mathbf{i}}} - r_1(\mathbf{i}), \;\; \widehat{Y_{\mathbf{i}}} = \widetilde{Y_{\mathbf{i}}} - r_2(\mathbf{i}),$$

and $\widehat{r_1}(.)$ (resp. $\widehat{r_2}(.)$) is the kernel estimator of the $r_1(.)$ (respectively $r_2(.)$) regression function, which is denoted by

$$\widehat{r_1}\,(\mathbf{i}) = \frac{\sum_{\mathbf{j}\in\mathbf{I_n}} H\left(l_{\mathbf{n}}^{-1} \|\mathbf{i} - \mathbf{j}\|\right) X_{\mathbf{j}}}{\sum_{\mathbf{j}\in\mathbf{I_n}} H\,(l_{\mathbf{n}}^{-1} \|\mathbf{i} - \mathbf{j}\|)} \quad \left(\text{resp}\,. \quad \widehat{r_2}\,(\mathbf{j}) = \frac{\sum_{\mathbf{i}\in\mathbf{I_n}} H\left(d_{\mathbf{n}}^{-1} \|\mathbf{j} - \mathbf{i}\|\right) Y_{\mathbf{i}}}{\sum_{\mathbf{i}\in\mathbf{I_n}} H\,(d_{\mathbf{n}}^{-1} \|\mathbf{j} - \mathbf{i}\|)}\right),$$

where H is a kernel function, and l_n and d_n are the real regression bandwidth parameters. Such a step, termed "detrending step", is crucial to non-parametric geographical data analysis. For our real data set, we emphasize the effects of this detrending procedure. To accomplish this, we evaluate the performance of the conditional mode regression in both instances (with and without detrending).

The functional curves $Z_{\mathbf{i}}$ are depicted in Figure 10.

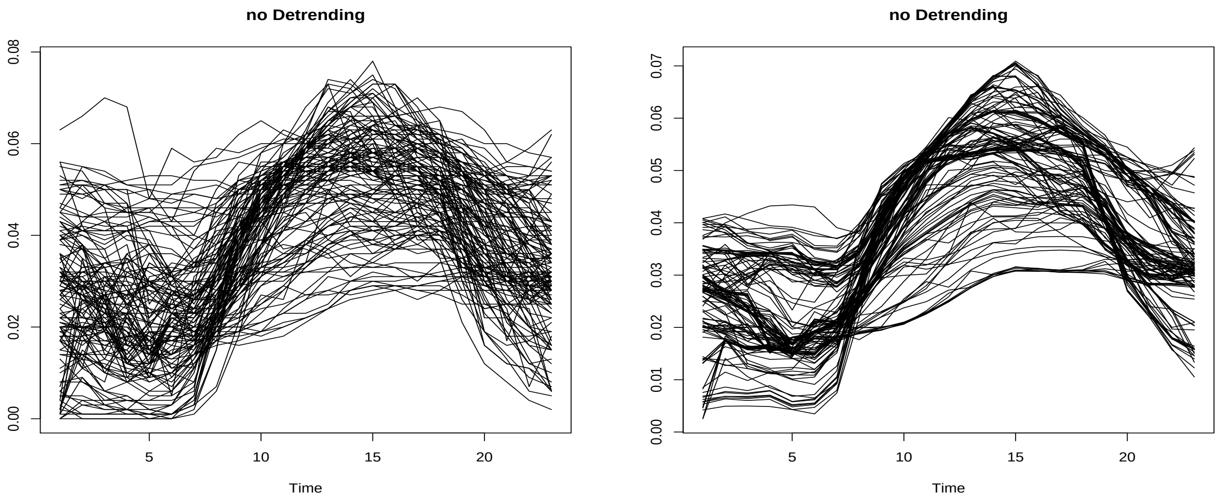

Figure 10. Ozone levels daily at 131 different US monitoring sites for both cases (detrending and no-detrending).

In particular, we compare the prediction of ozone pollution in a station using the three described models in the simulation part. We divided the observations as follows: Learning sample $(Z_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}\in I}$ (104 stations), and test sample $(Z_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}\in I'}$ (27 stations).

The quadratic kernel K is chosen as $K(u) = \frac{3}{4}\left(\frac{12}{11} - u^2\right)\mathbb{I}_{[0,1]}(u)$. The selection of the bandwidth parameter h is an important subject in nonparametric estimation, we propose utilizing a cross-validation approach to select the optimal bandwidth for the other methods. We use the cross-validation selection rule suggested by Ferraty and Vieu [2].

The standard *PCA* semi-metrics are used, where

$$d_q^{PCA}\left(Z_{\mathbf{i}}, Z_{\mathbf{j}}\right) = \sqrt{\sum_{k=1}^{q}\left(\int\left[Z_{\mathbf{i}}(t) - Z_{\mathbf{j}}(t)\right]v_k(t)dt\right)^2}.$$

In this case, $q = 4$ is used, and the v_k is chosen from the eigenfunctions of the empirical covariance operator:

$$\Gamma_Z^{\hat{n}}(s,t) = \frac{1}{\hat{n}}\sum_{\mathbf{i}\in I} Z_{\mathbf{i}}(s)Z_{\mathbf{i}}(t).$$

About the real regressions $\widehat{r_1}(.)$ and $\widehat{r_2}(.)$, we utilized the R-package *np* routine code *npreg*.

The three estimators $\widetilde{\theta}_{\mathbf{n}}(\omega)$, $\widehat{\theta}_{\mathbf{n}}(\omega)$ and $\bar{\theta}_{\mathbf{n}}(\omega)$ performance and behavior is expressed by the mean square error (MSE), defined by

$$\mathrm{MSE} = \frac{1}{\#(I')}\sum_{\mathbf{i}\in I'}\left(Y_{\mathbf{i}} - \theta\left(Z_{\mathbf{i}}\right)\right)^2, \tag{4.2}$$

where $(Z_{\mathbf{i}}, Y_{\mathbf{i}})$ represents the output observations (testing sample) and θ refers to either $\widetilde{\theta}_{\mathbf{n}}$, $\widehat{\theta}_{\mathbf{n}}$ and $\bar{\theta}_{\mathbf{n}}(\omega)$. The outcome of the prediction indicates that the kNN-LLE method is significantly superior to other given estimators. We plot the predicted values versus the real values for both methods in Figure 11 to illustrate the outcomes.

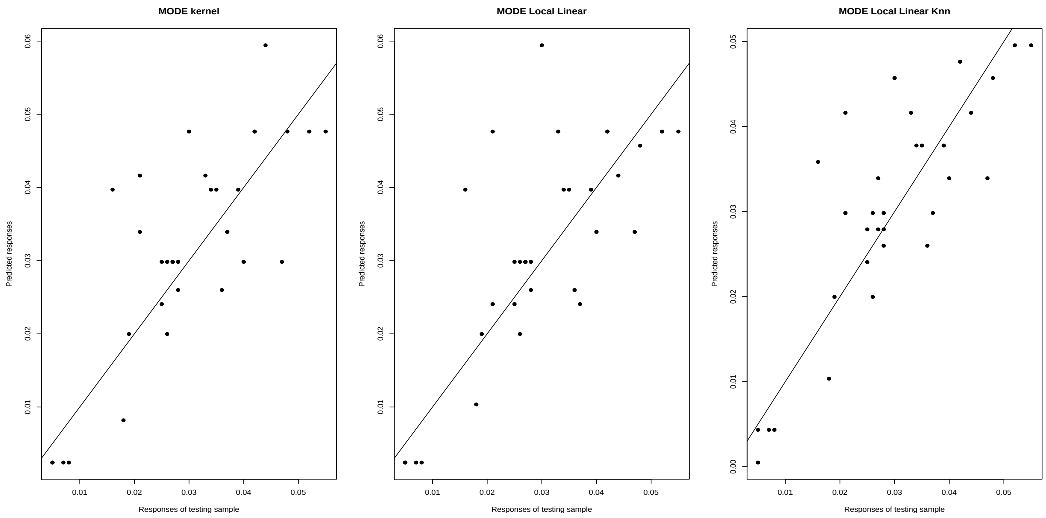

Figure 11. The predicted results for both tree methods.

The kernel method results are given on the right part. The center one gives the local linear method, while the left part of Figure 11 presents the kNN local linear method. Then, we remark that the performance of the prediction is controlled by the continuous line, in the sense that the efficiency of the prediction method is quantified by the closeness of the dark point to this continuous line. However, when the initial data are utilized without detrending, it is evident that there is a significant difference between the detrending case and the non-stationary case. In particular, detrending permits the MSE to be decreased. Boxplot (see Figure 12) reveals that the median error in the detrending scenario is 0.00196 while in the alternative situation it is 0.00467. In conclusion, we may state that the stationarity hypothesis is crucial to the non-parametric analysis of spatio-functional data, and that the proposed detrending method is an ideal instrument for verifying this hypothesis.

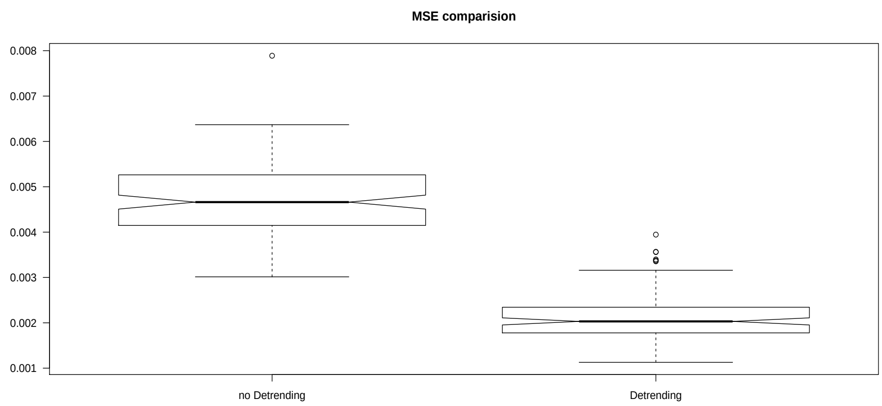

Figure 12. Comparison of the MSE values between cases with and without detrending.

The result of the kNN-LLE method's conditional mode prediction reveals that the trending case is much superior to the other non-trending case. Figure 13 illustrates the results by plotting the forecasted values against the true numbers for both data types.

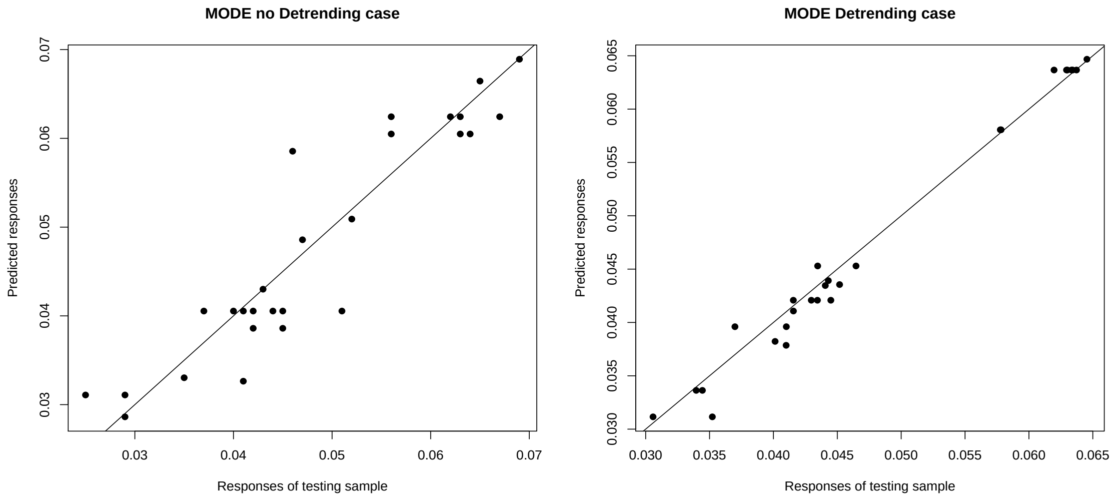

Figure 13. The predicted results for conditional mode in both cases with and without detrending.

Clearly, the comparison results in Figures 11–17 indicate that the method based on the kNN local linear polynomial estimation is significantly superior and more effective than the other methods (Predictions were taken in 4 different spatial areas). Furthermore, when the initial data are used without detrending, there is a significant superiority between the detrending case and the non-stationary case; this is confirmed by the mean squared errors MSE(kNN-LLE)detrending= 0.0027814, MSE(kNN-LLE)= 0.006185477 whereas MSE(LLE)= 0.0065786 and MSE(KNW)= 0.006699258.

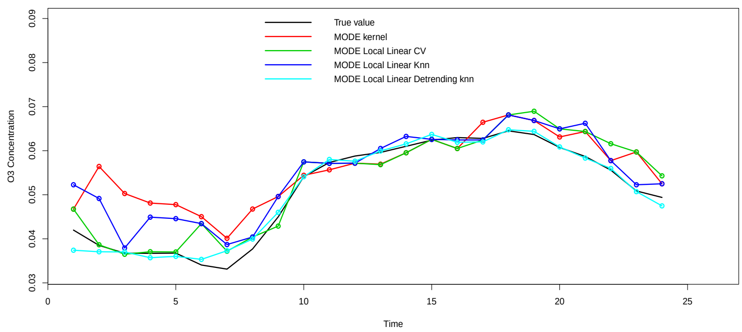

Figure 14. Prediction for zone 1 on May 2nd, 2021. The true values are linked by the solid black curve.

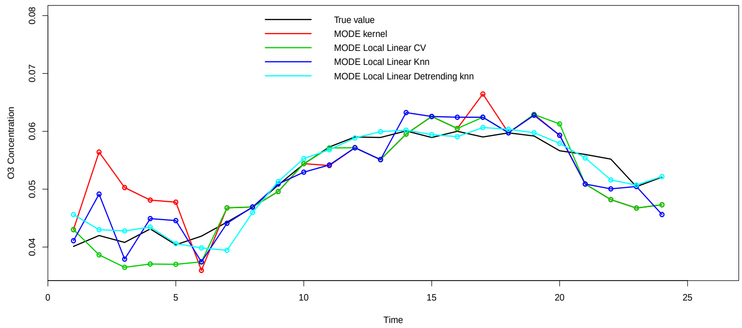

Figure 15. Prediction for zone 2 on May 2nd, 2021. The true values are linked by the solid black curve.

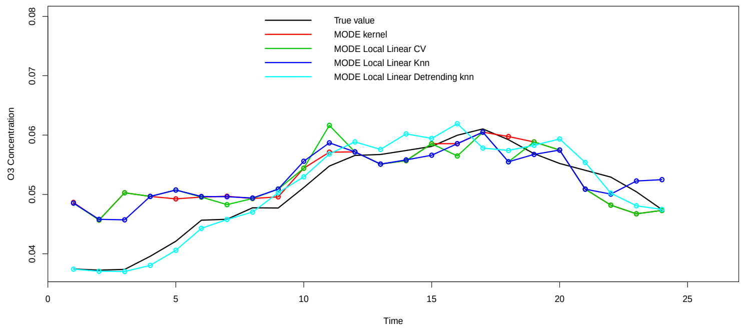

Figure 16. Prediction for zone 3 on May 2nd, 2021. The true values are linked by the solid black curve.

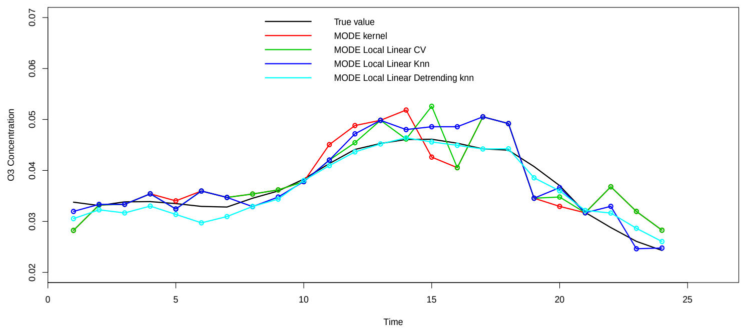

Figure 17. Prediction for zone 4 on May 2nd, 2021. The true values are linked by the solid black curve.

5. Conclusions

The functional kNN local linear technique (kNN-LLE), which combines the kNN algorithm and the local linear smoothing method to generate an estimate of the conditional mode, is a very interesting way to get a different estimator that benefits from the advantages of both strategies. To estimate the conditional density and mode for functional spatial data, this work proposes conditional distribution function and mode estimators. Using kernel estimations and several usual assumptions, we were able to establish the uniform, almost complete convergence of those models. To show the effectiveness of our suggested estimator and the theoretical conclusions, we conducted simulations and analyses of real data.

The estimation of conditional density is extremely valuable in practice. It provides more insightful details on the relationship between the input and output variables. In fact, the conditional density explains both the behavior of the center and the behavior extremes of the data, in contrast to the standard regression, which analyzes the central tendency of the data. Additionally, it plays a crucial part in statistical modeling, including prediction by conditional mode function, interval prediction by shortest conditional modal interval, risk analysis by shortfall expectation model, and association testing. The study presented in this paper provides some promising possibilities for future research. For instances:

- The functional response case: The nonparametric conditional mode has been presented in the nonparametric FDA where both variables (outcome and regressors) are of a functional sort (see, for example, Tadj et al. [45] for the kernel approach). However, there is still a lack of literature on this topic. Indeed, a natural extension is to apply our contribution results to this case.
- The spatial data: Other research questions, such as extensions to the semiparametric linear regression model, can also be investigated using the spatial data.

Acknowledgments

This research project was funded by the Deanship of Scientific R esearch, P rincess N ourah bint Abdulrahman University, through the Program of Research Project Funding After Publication, grant No (43- PRFA-P-23).

Conflict of interest

The authors declare no conflicts of interest.

References

1. G. Collomb, W. Härdle, S. Hassani, A note on prediction via estimation of the conditional mode function, *J. Stat. Plan. Infer.*, **15** (1987), 227–236. https://doi.org/10.1016/0378-3758(86)90099-6
2. F. Ferraty, P. Vieu, *Nonparametric functional data analysis: theory and practice*, New York: Springer, 2006. https://doi.org/10.1007/0-387-36620-2
3. A. Baíllo, A. Grané, Local linear regression for functional predictor and scalar response, *J. Multivariate Anal.*, **100** (2009), 102–111. https://doi.org/10.1016/j.jmva.2008.03.008
4. J. Barrientos-Marin, F. Ferraty, P. Vieu, Locally modelled regression and functional data, *J. Nonparametr. Stat.*, **22** (2010), 617–632. https://doi.org/10.1080/10485250903089930

5. J. Demongeot, A. Laksaci, F. Madani, M. Rachdi, Functional data: local linear estimation of the conditional density and its application, *Statistics*, **47** (2013), 26–44. https://doi.org/10.1080/02331888.2011.568117
6. M. Rachdi, A. Laksaci, J. Demongeot, A. Abdali, F. Madani, Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data, *Comput. Stat. Data Anal.*, **73** (2014), 53–68. https://doi.org/10.1016/j.csda.2013.11.011
7. Z. Y. Zhou, Z. Y. Lin, Asymptotic normality of locally modelled regression estimator for functional data, *J. Nonparametr. Stat.*, **28** (2016), 116–131. https://doi.org/10.1080/10485252.2015.1114112
8. I. M. Almanjahie, Z. Kaid, A. Laksaci, M. Rachdi, Estimating the conditional density in scalar-on-function regression structure: k-N-N local linear approach, *Mathematics*, **10** (2022), 1–16. https://doi.org/10.3390/math10060902
9. F. Belarbi, S. Chemikh, A. Laksaci, Local linear estimate of the nonparametric robust regression in functional data, *Stat. Probabil. Lett.*, **134** (2018), 128–133. https://doi.org/10.1016/j.spl.2017.11.003
10. A. Chahad, L. Ait-Hennani, A. Laksaci, Functional local linear estimate for functional relative-error regression, *J. Stat. Theory Pract.*, **11** (2017), 771–789. https://doi.org/10.1080/15598608.2017.1321071
11. M. Attouch, A. Laksaci, F. Rafaa, Local linear estimate of the regression operator by the *k*NN method, *C. R. Math.*, **355** (2017), 824–829. https://doi.org/10.1016/j.crma.2017.05.007
12. J. Fan, I. Gijbels, *Local polynomial modelling and its applications*, London: Chapman & Hall, 1996. https://doi.org/10.1201/9780203748725
13. M. Hallin, Z. D. Lu, L. T. Tran, Local linear spatial regression, *Ann. Statist.*, **32** (2004), 2469–2500. https://doi.org/10.1214/009053604000000850
14. M. Hallin, Z. D. Lu, K. M. Yu,Local linear spatial quantile regression, *Bernouilli*, **15** (2009), 659–686. https://doi.org/10.3150/08-BEJ168
15. A. Chouaf, A. Laksaci, On the functional local linear estimate for spatial regression, *Statist. Risk Model.*, **29** (2012), 189–214. https://doi.org/10.1524/strm.2012.1114
16. A. Laksaci, M. Rachdi, S. Rahmani, Spatial modelization: local linear estimation of the conditional distribution for functional data, *Spatial Stat.*, **6** (2013), 1–23. https://doi.org/10.1016/j.spasta.2013.04.004
17. A. Gheriballah, A. Laksaci, R. Rouane, Robust nonparametric estimation for spatial regression, *J. Stat. Plan. Infer.*, **140** (2010), 1656–1670. https://doi.org/10.1016/j.jspi.2010.01.042
18. M. Attouch, A. Laksaci, N. Messabihi, Nonparametric relative error regression for spatial random variables, *Stat. Papers*, **58** (2017), 987–1008. https://doi.org/10.1007/s00362-015-0735-6
19. M. Abeidallah, B. Mechab, T. Merouan, Local linear estimate of the point at high risk: spatial functional data case, *Commun. Stat. Theory Methods*, **49** (2020), 2561–2584. https://doi.org/10.1080/03610926.2019.1580735
20. M. Rachdi, A. Laksaci, F. A. Al-Awhadi, Parametric and nonparametric conditional quantile regression modelisation for dependent spatial functional data, *Spatial Stat.*, **43** (2021), 100498. https://doi.org/10.1016/j.spasta.2021.100498

21. M. Rachdi, A. Laksaci, N. M. Al-Kandari, Expectile regression for spatial functional data analysis (sFDA), *Metrika*, **85** (2022), 627–655. https://doi.org/10.1007/s00184-021-00846-x

22. A. Amiri, S. Dabo-Niang, M. Yahaya, Nonparametric recursive density estimation for spatial data, *C. R. Math.*, **354** (2016), 205–210. https://doi.org/10.1016/j.crma.2015.10.010

23. R. Giraldo, S. Dabo-Niang, S. Martínez, Statistical modeling of spatial big data: an approach from a functional data analysis perspective, *Stat. Probabil. Lett.*, **136** (2018), 126–129. https://doi.org/10.1016/j.spl.2018.02.025

24. M. Chaouch, N. Laïb, D. Louani, Rate of uniform consistency for a class of mode regression on functional stationary ergodic data, *Stat. Methods Appl.*, **26** (2017), 19–47. https://doi.org/10.1007/s10260-016-0356-9

25. B. Maillot, C. Chesneau, On the conditional density estimation for continuous time processes with values in functional in spaces, *Stat. Probabil. Lett.*, **178** (2021), 109179. https://doi.org/10.1016/j.spl.2021.109179

26. J. L. Kirkby, A. Leitao, D. Nguyen, Nonparametric density estimation and bandwidth selection with B-spline bases: a novel Galerkin method, *Comput. Stat. Data Anal.*, **159** (2021), 107202. https://doi.org/10.1016/j.csda.2021.107202

27. T. Cover, Estimation by the nearest neighbor rule, *IEEE Trans. Inform. Theory*, **14** (1968), 50–55. https://doi.org/10.1109/TIT.1968.1054098

28. N. L. Kudraszow, P. Vieu, Uniform consistency of kNN regressors for functional variables, *Stat. Probabil. Lett.*, **83** (2013), 1863–1870. https://doi.org/10.1016/j.spl.2013.04.017

29. L. Kara, A. Laksaci, M. Rachdi, P. Vieu, Data-driven kNN estimation in nonparametric functional data analysis, *J. Multivariate Anal.*, **153** (2017), 176–188. https://doi.org/10.1016/j.jmva.2016.09.016

30. M. Attouch, W. Bouabsa, Z. Chiker el Mozoaur, The k-nearest neighbors estimation of the conditional mode for functional data under dependency, *Int. J. Stat. Econ.*, **19** (2018), 48–60.

31. I. M. Almanjahie, K. A. Assiri, A. Laksaci, Z. Chikr Elmezouar, The k nearest neighbors smoothing of the relative-error regression with functional regressor, *Commun. Stat. Theory Methods*, **51** (2022), 4196–4209. https://doi.org/10.1080/03610926.2020.1811870

32. W. Bouabsa, Nonparametric relative error estimation via functional regressor by the k nearest neighbors smoothing under truncation random data, *Appl. Appl. Math.*, **16** (2021), 97–116.

33. W. Bouabsa, Unform in bandwith of the conditional distribution function with functional explanatory variable: the case of spatial data with the k nearest neighbour method, *Econometrics*, **26** (2022), 30–46. https://doi.org/10.15611/eada.2022.2.03

34. T. Bodnar, O. Okhrin, N. Parolya, Optimal shrinkage estimator for high-dimensional mean vector, *J. Multivariate Anal.*, **170** (2019), 63–79. https://doi.org/10.1016/j.jmva.2018.07.004

35. Y. Gao, H. L. Shang, Y. R. Yang, High-dimensional functional time series forecasting: an application to age-specific mortality rates, *J. Multivariate Anal.*, **170** (2019), 232–243. https://doi.org/10.1016/j.jmva.2018.10.003

36. J. R. Berrendero, B. Bueno-Larraz, A. Cuevas, An RKHS model for variable selection in functional linear regression, *J. Multivariate Anal.*, **170** (2019), 25–45. https://doi.org/10.1016/j.jmva.2018.04.008
37. C. Aaron, A. Cholaquidis, R. Fraiman, B. Ghattas, Multivariate and functional robust fusion methods for structured big data, *J. Multivariate Anal.*, **170** (2019), 149–161. https://doi.org/10.1016/j.jmva.2018.06.012
38. P. Hall, M. Hosseini-Nasab, On properties of functional principal components analysis, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **68** (2006), 109–126. https://doi.org/10.1111/j.1467-9868.2005.00535.x
39. P. Hall, H. G. Muller, J. L. Wang, Properties of principal component methods for functional and longitudinal data analysis, *Ann. Statist.*, **34** (2006), 1493–1517. https://doi.org/10.1214/009053606000000272
40. L. T. Tran, Kernel density estimation on random fields, *J. Multivariate Anal.*, **34** (1990), 37–53. https://doi.org/10.1016/0047-259X(90)90059-Q
41. X. Guyon, Estimation d'un champ par pseudo-vraisemblance conditionnelle: Etude asymptotique et application au cas Markovien, In: *Proceedings of the sixth Franco-Belgian meeting of statisticians*, 1987.
42. P. Doukhan, *Mixing: properties and examples*, New York: Springer, 1994. https://doi.org/10.1007/978-1-4612-2642-0
43. M. Carbon, L. T. Tran, B. Wu, Kernel density estimation for random fields (density estimation for random fields), *Stat. Probabil. Lett.*, **36** (1997), 115–125. https://doi.org/10.1016/S0167-7152(97)00054-0
44. S. Dabo-Niang, M. Rachdi, A. F. Yao, Kernel regression estimation for spatial functional random variables, *Far East J. Theor. Stat.*, **37** (2011), 77–113.
45. A. Tadj, *Sur les modèles non paramétriques conditionnels en statistique fonctionnelle*, Thèse de doctorat, Université de Toulouse, 2011.
46. G. Biau, B. Cadre, Nonparametric spatial prediction, *Stat. Infer. Stoch. Pro.*, **7** (2004), 327–349. https://doi.org/10.1023/B:SISP.0000049116.23705.88

Appendix

Proof of Corollary 3.1

Remember that for any $h_L \in \left(\hat{a}_{\mathbf{n}}, \hat{b}_{\mathbf{n}}\right)$ and $b_\delta \in \left(\hat{c}_{\mathbf{n}}, \hat{d}_{\mathbf{n}}\right)$, we have

$$\left|1 - \tilde{f}_D(z)\right| \leq \frac{\left(1 - \tilde{f}_D(z)\right)}{2} \quad \text{implies} \quad \left|\tilde{f}_D(z) - f_D(z)\right| \geq \frac{\left(1 - \tilde{f}_D(z)\right)}{2}.$$

Therefore, we have

$$\begin{aligned} &\sum_{\mathbf{n}\in\mathbb{Z}^M} \mathbb{P}\left(\inf_{h_L\in\left(\hat{a}_{\mathbf{n}},\hat{b}_{\mathbf{n}}\right),b_\delta\in\left(\hat{c}_{\mathbf{n}},\hat{d}_{\mathbf{n}}\right)} \left|1 - \tilde{f}_D(z)\right| \leq \frac{1 - \tilde{f}_D(z)}{2}\right) \\ &\leq \sum_{\mathbf{n}\in\mathbb{Z}^M} \mathbb{P}\left(\sup_{h_L\in\left(\hat{a}_{\mathbf{n}},\hat{b}_{\mathbf{n}}\right),b_\delta\in\left(\hat{c}_{\mathbf{n}},\hat{d}_{\mathbf{n}}\right)} \left|\tilde{f}_D(z) - f_D(z)\right| \geq \frac{1 - \tilde{f}_D(z)}{2}\right) < \infty. \end{aligned}$$

Proof of Lemma 3.1

Due to the expectation's linearity, we have $\forall h_L \in (\hat{a}_{\mathbf{n}}, \hat{b}_{\mathbf{n}})$ and $b_\delta \in (\hat{c}_{\mathbf{n}}, \hat{d}_{\mathbf{n}})$,

$$\mathbb{E}\left[\tilde{f}_N(w \mid z)\right] - f(w \mid z)\mathbb{E}\left[\tilde{f}_D(z)\right] = \frac{1}{b_\delta Y_{12}}\mathbb{E}\left[Y_{12}\left[\mathbb{E}\left[\delta_2(w) \mid Z_2\right] - f(w \mid z)\right]\right],$$

where $\delta_2(w) = b_\delta^{-1}\delta(b_\delta^{-1}(w - W_2))$.

Then, we apply (A3), uniformly on $w \in \mathcal{S}$, $\forall h_L \in (\hat{a}_{\mathbf{n}}, \hat{b}_{\mathbf{n}})$ and $b_\delta \in (\hat{c}_{\mathbf{n}}, \hat{d}_{\mathbf{n}})$, to obtain

$$\mathbb{I}_{\mathbf{B}(z,h_L)}(Z_2)\left|\mathbb{E}\left[\delta_2(w) \mid Z_2\right] - f(w \mid z)\right| \leq C\left[h_L^{\varpi_1} + b_\delta^{\varpi_2}\int_{\mathbb{R}}|t|^{\varpi_2}\delta_2(t)dt\right]dt.$$

Finally, assumption (A6) and Corollary 3.1 lead in the demonstration of this lemma.

Proof of Lemma 3.2

To begin, we take into account the subset's $\mathcal{S}$ compactness condition, permitting us to create a sequence $(r_1, r_2, \cdots r_{s_{\mathbf{n}}}) \in \mathcal{S}$. Hence, we have $\mathcal{S} \in \bigcup_{j=1}^{s_{\mathbf{n}}}(r_j - t_{\mathbf{n}}, r_j + t_{\mathbf{n}})$, with $t_{\mathbf{n}} = \hat{\mathbf{n}}^{(-1/2-\frac{3\alpha}{2})}$ and $s_{\mathbf{n}} \leq \hat{\mathbf{n}}^{1/2+\alpha}$. Then, $\forall w \in \mathcal{S}$, we place $r_w = \arg\min_{t\in\{r_1,\ldots,r_{s_{\mathbf{n}}}\}}|w - r_j|$, and consider the following decomposition:

$$\left|\tilde{f}_N(w \mid z) - \mathbb{E}\tilde{f}_N(w \mid z)\right| \leq \underbrace{\tilde{f}_N(w \mid z) - \tilde{f}_N\left(r_{j(z)}\right)|}_{\Gamma_1} + \underbrace{\tilde{f}_N\left(r_{j(z)}\right) - \mathbb{E}\tilde{f}_N\left(r_j|z\right)|}_{\Gamma_2} + \underbrace{\left|\mathbb{E}\tilde{f}_N\left(r_{j(z)}\right) - \mathbb{E}\tilde{f}_N(w \mid z)\right|}_{\Gamma_3}.$$

- For the terms Γ_1 and Γ_3, we use (A6) to get successively

$$\begin{aligned}
&\sup_{\hat{c}_{\mathbf{n}}\leq b_\delta\leq\hat{d}_{\mathbf{n}}}\ \max_{w\in\{r_j-t_{\mathbf{n}},r_j+t_{\mathbf{n}}\}}\left|\tilde{f}_N(w \mid z) - \tilde{f}_N\left(r_{j(z)}\right)\right| \\
\leq&\frac{1}{\hat{\mathbf{n}}(\hat{\mathbf{n}}-1)b_\delta\mathbb{E}\left[Y_{12}(z)\right]}\sum_{\mathbf{i}\neq\mathbf{j}}Y_{\mathbf{ij}}(z)\left|\delta_{\mathbf{i}}'(z) - \delta_{\mathbf{i}}'\left(r_{j(z)}\right)\right| \\
\leq&C\frac{t_{\mathbf{n}}}{b_\delta^2}\tilde{f}_D(z).
\end{aligned}$$

  Under (A7) and by using the definition of $t_{\mathbf{n}}$, we deduce that $\dfrac{t_{\mathbf{n}}}{b_\delta^2} = o\left(\sqrt{\dfrac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}b_\delta\zeta_z(h_L)}}\right)$.

  Next, we get

$$\sup_{w\in\mathcal{S}}\ \sup_{\hat{a}_{\mathbf{n}}\leq h_L\leq\hat{b}_{\mathbf{n}}}\ \sup_{\hat{c}_{\mathbf{n}}\leq b_\delta\leq\hat{d}_{\mathbf{n}}}\left|\tilde{f}_N(w|z) - \tilde{f}_N\left(r_{j(z)}\right)\right| = o\left(\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}d_{\mathbf{n}}\zeta_z(a_{\mathbf{n}})}}\right), \tag{a1}$$

  and

$$\sup_{w\in\mathcal{S}}\ \sup_{\hat{a}_{\mathbf{n}}\leq h_L\leq\hat{b}_{\mathbf{n}}}\ \sup_{\hat{c}_{\mathbf{n}}\leq b_\delta\leq\hat{d}_{\mathbf{n}}}\left|\mathbb{E}\tilde{f}_N\left(r_{j(z)}\right) - \mathbb{E}\tilde{f}_N(w|z)\right| = o\left(\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}\hat{d}_{\mathbf{n}}\zeta_z(\hat{a}_{\mathbf{n}})}}\right). \tag{a2}$$

- For the term Γ_2, for any $\eta > 0$, we must assess the quantity

$$\mathbb{P}\left(\sup_{w\in\mathcal{S}}\ \sup_{\hat{a}_{\mathbf{n}}\leq h_L\leq\hat{b}_{\mathbf{n}}}\ \sup_{\hat{c}_{\mathbf{n}}\leq b_\delta\leq\hat{d}_{\mathbf{n}}}\ \max_{w\in\{r_j-t_{\mathbf{n}},r_j+t_{\mathbf{n}}\}}\ \max_{\mathfrak{x}\in\{1,2,\ldots,s_{\mathbf{n}}\}}|\tilde{f}_N\left(r_{j(z)}\right) - \mathbb{E}\tilde{f}_N\left(r_{j(z)}\right)| > \eta\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}d_{\mathbf{n}}\zeta_z(a_{\mathbf{n}})}}\right)$$

$$= \mathbb{P}\left(\sup_{w\in\mathcal{S}} \sup_{\hat{a}_{\mathbf{n}}\le h_L\le\hat{b}_{\mathbf{n}}} \sup_{\hat{c}_{\mathbf{n}}\le b_\delta\le\hat{d}_{\mathbf{n}}} \max_{w\in\left\{r_j-t_{\mathbf{n}},r_j+t_{\mathbf{n}}\right\}} \max_{\mathfrak{x}\in\{1,2,\dots,s_{\mathbf{n}}\}} |\tilde{f}_N^z\left(r_{j(z)}\right) - \mathbb{E}\tilde{f}_N^z\left(r_{j(z)}\right)| > \eta\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}d_{\mathbf{n}}\zeta_z(a_{\mathbf{n}})}} \right)$$

$$\le s_{\mathbf{n}} \sup_{w\in\mathcal{S}} \max_{w\in\left\{r_j-t_{\mathbf{n}},r_j+t_{\mathbf{n}}\right\}} \max_{\mathfrak{x}\in\{1,2,\dots,s_n\}} \mathbb{P}\left(\sup_{\hat{a}_{\mathbf{n}}\le h_L\le\hat{b}_{\mathbf{n}}} \sup_{\hat{c}_{\mathbf{n}}\le b_\delta\le\hat{d}_{\mathbf{n}}} |\tilde{f}_N^z\left(r_{j(z)}\right) - \mathbb{E}\tilde{f}_N^z\left(r_{j(z)}\right)| > \eta\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}d_{\mathbf{n}}\zeta_z(a_{\mathbf{n}})}} \right),$$

for some $\eta > 0$.

Additionally, highlight that this latter condition may be handled using the same methods as in Barrientos-Marin et al. [4]. In effect, we can use the breakdown described below:

$$\begin{aligned} \hat{f}_N\left(r_{j(z)}\right) = \underbrace{\frac{\hat{\mathbf{n}}^2 h_L^2\zeta_z^2(h_L)}{\hat{\mathbf{n}}(\hat{\mathbf{n}}-1)\mathbb{E}[Y_{12}]}}_{U_0}[\underbrace{\left(\frac{1}{\hat{\mathbf{n}}}\sum_{\mathbf{j}\in t_{\mathbf{n}}}\frac{L_{\mathbf{j}}\delta'_{\mathbf{j}}\left(r_j(w)\right)}{b_\delta\zeta_z(h_L)}\right)}_{D_1}\underbrace{\left(\frac{1}{\hat{\mathbf{n}}}\sum_{\mathbf{i}\in t_{\mathbf{n}}}\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^2}{h_L^2\zeta_z(h_L)}\right)}_{D_2} \\ -\underbrace{\left(\frac{1}{\hat{\mathbf{n}}}\sum_{\mathfrak{x}\in t_{\mathbf{n}}}\frac{L_{\mathfrak{x}}\chi_{\mathfrak{x}}\delta'_{\mathfrak{x}}\left(r_j(w)\right)}{h_L b_\delta\zeta_z(h_L)}\right)}_{D_3}\underbrace{\left(\frac{1}{\hat{\mathbf{n}}}\sum_{\mathbf{i}\in t_{\mathbf{n}}}\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}}{h_L\zeta_z(h_L)}\right)}_{D_4}\Bigg]. \end{aligned} \tag{a3}$$

The result demonstrates the following equations for some positive numbers, $\hat{b}_{\mathbf{0}}$ and $\hat{d}_{\mathbf{0}}$,

$$\sum_{\mathbf{n}} s_{\mathbf{n}}\mathbb{P}\left\{ \sup_{\hat{a}_{\mathbf{n}}\le h_L\le\hat{b}_{\mathbf{0}}} \sup_{\hat{c}_{\mathbf{n}}\le b_\delta\le\hat{d}_{\mathbf{0}}} |D_{\mathfrak{x}} - \mathbb{E}[D_{\mathfrak{x}}]| > \eta_0\left(\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}\zeta_z(h_L)}\right)^{1/2} \right\} < \infty, \text{ for } \mathfrak{x} = 2, 4. \tag{a4}$$

$$\sum_{\mathbf{n}} s_{\mathbf{n}}\mathbb{P}\left\{ \sup_{\hat{a}_{\mathbf{n}}\le h_L\le\hat{b}_{\mathbf{0}}} \sup_{\hat{c}_{\mathbf{n}}\le b_\delta\le\hat{d}_{\mathbf{0}}} |D_{\mathfrak{x}} - \mathbb{E}[D_{\mathfrak{x}}]| > \eta_0\left(\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}b_\delta\zeta_z(h_L)}\right)^{1/2} \right\} < \infty, \text{ for } \mathfrak{x} = 1, 3. \tag{a5}$$

$$\mathrm{Cov}(D_1, D_2) = o\left(\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}b_\delta\zeta_z(h_L)}}\right), \tag{a6}$$

and

$$\mathrm{Cov}(D_3, D_4) = o\left(\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}b_\delta\zeta_z(h_L)}}\right). \tag{a7}$$

Let's now show the (a4) and (a5) results. It is sufficient to show that using the same method in Barrientos-Marin et al. [4] as in the (i.i.d.) case. We are going to evaluate U_0. To illustrate the findings of (a4) and (a5), we use the two variables $\Omega_{\mathbf{i}}$ and $\Upsilon_{\mathbf{i}}$, defined by

$$\Omega_{\mathbf{i}}^{\mathfrak{x}} = \frac{1}{h_L^{\mathfrak{x}}}L_{\mathbf{i}}\chi_{\mathbf{i}}^{\mathfrak{x}} - \frac{1}{h_L^{\mathfrak{x}}}\mathbb{E}\left[L_{\mathbf{i}}\chi_{\mathbf{i}}^{\mathfrak{x}}\right], \text{ for } \mathfrak{x} = 1, 2,$$

and

$$\Upsilon_{\mathbf{i}}^{\mathfrak{x}} = \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{\mathfrak{x}}\delta'_{\mathbf{i}}\left(r_j \mid w\right)}{h_L^{\mathfrak{x}}b_\delta} - \mathbb{E}\left[\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{\mathfrak{x}}\delta'_{\mathbf{i}}\left(r_j \mid w\right).}{h_L^{\mathfrak{x}}b_\delta}\right], \quad \text{for} \quad \mathfrak{x} = 0, 1.$$

It can then be observed that for $\mathbf{i} = 2, 4, \exists \mathfrak{X}$ as well as $D_i - \mathbb{E}[D_i] = \dfrac{1}{\hat{\mathbf{n}}\zeta_z(h_L)} \sum_{\mathbf{i}\in \iota_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}}$.

Next, by considering the spatial breakdown of Tran [40] on the $\Omega_{\mathbf{i}}$ and $\Upsilon_{\mathbf{i}}$ variables, expressed, for a constant integer $\rho_{\mathbf{n}}$, shown below:

$$\begin{aligned}
A(1, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_\rho = 2j_\rho q_{\mathbf{n}+1} \\ \mathfrak{X}=1,\ldots,M}}^{2j_\rho q_{\mathbf{n}} + q_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}}, \\
A(2, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_\rho = 2j_\rho q_{\mathbf{n}}+1 \\ \mathfrak{X}=1,\ldots,M-1}} \sum_{i_M = 2j_M q_{\mathbf{n}} + q_{\mathbf{n}}+1}^{2(j_M+1)q_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}}, \\
A(3, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_\rho = 2j_\rho q_{\mathbf{n}+1} \\ \mathfrak{X}=1,\ldots,M-2}}^{2j_\rho q_{\mathbf{n}} + q_{\mathbf{n}}} \sum_{i_{M-1} = 2j_{M-1} q_{\mathbf{n}} + q_{\mathbf{n}}+1}^{2(j_{M-1}+1)q_{\mathbf{n}}} \sum_{i_M = 2j_M q_{\mathbf{n}}+1}^{2j_M q_{\mathbf{n}} + q_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}}, \\
A(4, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_\rho = 2j_\rho q_{\mathbf{n}+1} \\ \mathfrak{X}=1,\ldots,M-2}}^{2j_\rho q_{\mathbf{n}}} \sum_{i_{M-1} = 2j_{M-1} q_{\mathbf{n}} + q_{\mathbf{n}}+1}^{2(j_{M-1}+1)q_{\mathbf{n}}} \sum_{i_M = 2j_M q_{\mathbf{n}} + q_{\mathbf{n}}+1}^{2(j_M+1)\sigma_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}},
\end{aligned}$$

the final two terms, in this series, are

$$A\left(2^{M-1}, \mathbf{n}, \mathbf{j}\right) = \sum_{\substack{i_\rho = 2j_\rho q_{\mathbf{n}} + q_{\mathbf{n}}+1 \\ \mathfrak{X}=1,\ldots,M-1}}^{2\left(j_\rho+1\right)q_{\mathbf{n}}} \sum_{i_M = 2j_M q_{\mathbf{n}}+1}^{2j_M q_{\mathbf{n}} + q_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}} \text{ and } A\left(2^{M}, \mathbf{n}, \mathbf{j}\right) = \sum_{\substack{i_\rho = 2j_\rho q_{\mathbf{n}} + q_{\mathbf{n}}+1 \\ \mathfrak{X}=1,\ldots,M}}^{2\left(j_\rho+1\right)q_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}},$$

for $i = 1, \ldots, M, \varsigma_i = 2^{-1}\mathbf{n}_i q_{\mathbf{n}}^{-1}$, and $Q = \{0, \ldots, \varsigma_1 - 1\} \times \cdots \times \{0, \ldots, \varsigma_M - 1\}$. We indicate

$$\mathcal{H}(\mathbf{n}, i) = \sum_{\mathbf{j}\in Q} A(i, \mathbf{n}, \mathbf{j}, \mathfrak{X}) \text{ for } 1 \leq i \leq 2^M,$$

in which it is simple and easy to see that

$$\frac{1}{\hat{\mathbf{n}}\zeta_z(h_L)} \sum_{\mathbf{i}\in \mathcal{I}_{\mathbf{n}}} \Omega \mathbf{i}^{\mathfrak{X}} - \mathbb{E}\left[\frac{1}{\hat{\mathbf{n}}\zeta_z(h_L)} \sum_{\mathbf{i}\in \mathcal{I}_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}}\right] = \frac{1}{\hat{\mathbf{n}}\zeta_z(h_L)} \sum_{i=1}^{2^M} \mathcal{H}(\mathbf{n}, i, \mathfrak{X}).$$

It is crucial to take remark that larger blocks are preferred, the summation of the random variables $\Omega_{\mathbf{i}}^{\mathfrak{X}}$ is $\mathcal{H}(\mathbf{n}, 1)$; the other terms, however, are sums over small blocks, which is $\mathcal{H}(\mathbf{n}, i)$, for $2 \leq i \leq 2^M$.

We mainly emphasize that, like previously stated in Biau and Cadre [46], the term $\mathcal{H}\left(\mathbf{n}, 2^M + 1\right)$ (that includes the $\Omega_{\mathbf{i}}^{\mathfrak{X}}$ is at the end and is not part of the above blocks) can be added if we don't have the equality $n_i = 2\varsigma_i q_{\mathbf{n}}$. Since this expression has minimal effects on the proof, we can express, for every $\eta > 0$,

$$\mathbb{P}\left(\left|\frac{1}{\hat{\mathbf{n}}\zeta_z(h_L)} \sum_{\mathbf{i}\in \mathcal{I}_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}} - \mathbb{E}\left[\frac{1}{\hat{\mathbf{n}}\zeta_z(h_L)} \sum_{\mathbf{i}\in \mathcal{I}_{\mathbf{n}}} \Omega_{\mathbf{i}}^{\mathfrak{X}}\right]\right| \geq \eta\right) \leq 2^M \max_{i=1,\ldots,2^M} \mathbb{P}\left(\mathcal{H}(\mathbf{n}, i) \geq \eta \hat{\mathbf{n}}\zeta_z(h_L)\right).$$

Consequently, when the following quantities are evaluated, the desired outcome follows.

$$\mathbb{P}\left(\mathcal{H}(\mathbf{n}, i) \geq \eta_0 \hat{\mathbf{n}} \zeta_z\left(h_L\right)\right), \text{ for all } i=1, \ldots ., 2^M .$$

Due to the similarity of the other situations, we only discuss the situation when $i = 1$, just for shortness. To accomplish that, we list the random variables

$$\beta=\prod_{\mathfrak{X}=1}^M \varsigma_{\mathfrak{X}}=2^{-M} \hat{\mathbf{n}} q_{\mathbf{n}}^{-M},$$

$A(1, \mathbf{n}, \mathbf{j}, \mathfrak{X})$ for every $\mathbf{j} \in Q$ in an arbitrary way $H_1^{\mathfrak{X}}, \ldots, H_\beta^{\mathfrak{X}}$. This means that there are specific $\mathbf{j}$ in β for each $H_{\mathbf{j}}^{\mathfrak{X}}$ such that

$$H_{\mathbf{j}}^{\mathfrak{X}}=\sum_{\mathbf{i} \in \mathcal{I}(1, \mathbf{n}, \mathbf{j})} \Omega_{\mathbf{i}}^{\mathfrak{X}},$$

with $\mathcal{I}(1, \mathbf{n}, \mathbf{j})=\left\{\mathbf{i}: 2 j_\rho q_{\mathbf{n}}+1 \leq i_{\mathfrak{X}} \leq 2 j_\rho p_{\mathbf{n}}+q_{\mathbf{n}}\right.$ for $\mathfrak{X}=1, \ldots, M\}$. Obviously, the ensemble $\mathcal{I}(1, \mathbf{n}, \mathbf{j})$ includes $q_{\mathbf{n}}^M$ domains; hence, there is a $q_{\mathbf{n}}$ distance between these sites. Additionally, it is implied by assumptions (A4) and (A5) that a positive constant named C exists, where

$$\frac{1}{h_L^{\mathfrak{X}}} L_{\mathbf{i}} \chi_{\mathbf{i}}^{\mathfrak{X}} \leq \frac{1}{h_L^{\mathfrak{X}}} L_{\mathbf{i}}\left|\Omega\left(Z_{\mathbf{i}}, z\right)\right|^{\mathfrak{X}} \leq \frac{1}{h_L^{\mathfrak{X}}} L\left(h_L^{-1} \wp\left(z, Z_{\mathbf{i}}\right)\right)\left|\wp\left(Z_{\mathbf{i}}, z\right)\right|^{\mathfrak{X}} \mathbb{I}\Big]-1,1\Big[\left(\left(h_L^{-1} \delta\left(z, Z_{\mathbf{i}}\right)\right) \leq L\left(h_L^{-1} \wp\left(z, Z_{\mathbf{i}}\right)\right) \leq C\right. .$$

Consequently, as indicated in (Carbon et al. [43], Lemma 4.5), for $j=1, \ldots, \beta$, that one may obtain independent random variables $H_1^*, \ldots, H_\beta^*$ that are i.i.d. as $H_j^{\mathfrak{X}}$ and such that

$$\sum_{j=1}^\beta \mathbb{E}\left|H_{\mathbf{j}}^{\mathfrak{X}}-H_{\mathbf{j}}^*\right| \leq 2 C \beta q_{\mathbf{n}}^M \Psi\left((\beta-1) q_{\mathbf{n}}^M, q_{\mathbf{n}}^M\right) \Phi\left(q_{\mathbf{n}}\right).$$

We are able to write

$$\mathbb{P}\left(\mathcal{H}(\mathbf{n}, i) \geq \eta_0 \hat{\mathbf{n}} \zeta_z\left(h_L\right)\right) \leq Q_1(\mathbf{n})+Q_2(\mathbf{n}),$$

where

$$Q_1(\mathbf{n})=\mathbb{P}\left(\| \sum_{j=1}^\beta H_{\mathbf{j}}^* \mid \geq \frac{\beta \eta_0 \hat{\mathbf{n}} \zeta_z\left(h_L\right)}{2 \beta}\right),$$

and

$$Q_2(\mathbf{n})=\mathbb{P}\left(\sum_{j=1}^\beta\left|H_{\mathbf{j}}^{\mathfrak{X}}-H_{\mathbf{j}}^*\right| \geq \frac{\eta_0 \hat{\mathbf{n}} \zeta_z\left(h_L\right)}{2}\right).$$

Regarding the expression $Q_1(\mathbf{n})$, the inequality of Bernstein implies that

$$Q_1(\mathbf{n})=\mathbb{P}\left(\left|\sum_{j=1}^\beta H_{\mathbf{j}}^*\right| \geq \frac{\beta \eta_0 \hat{\mathbf{n}} \zeta_z\left(h_L\right)}{2 \beta}\right) \leq 2 \exp \left(-\frac{\left(\eta_0 \hat{\mathbf{n}} \zeta_z\left(h_L\right)\right)^2}{\beta \operatorname{Var}\left[H_1^{\mathfrak{X} *}\right]+C q_{\mathbf{n}}^M \eta_0 \hat{\mathbf{n}} \zeta_z\left(h_L\right)}\right).$$

To understand how this expression behaves, asymptotically, we should compute $\operatorname{Var}\left[H_1^*\right]$. In fact,

$$\operatorname{Var}\left[H_1^{\mathfrak{X} *}\right]=\operatorname{Var}\left[\sum_{\mathbf{i} \in \mathcal{I}(1, \mathbf{n}, 1)} \Omega_{\mathbf{i}}^{\mathfrak{X}}\right]=\sum_{\mathbf{i}, \mathbf{j} \in \mathcal{I}(1, \mathbf{n}, 1)}\left|\operatorname{Cov}\left(\Omega_{\mathbf{i}}^{\mathfrak{X}}, \Omega_{\mathbf{j}}^{\mathfrak{X}}\right)\right| .$$

Let $R_{\mathbf{n}} = \sum_{\mathbf{i}\in\mathcal{I}(1,\mathbf{n},1)} \mathrm{Var}\left[\Omega \mathbf{i}^{\mathfrak{x}}\right]$ and $T_{\mathbf{n}} = \sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}(1,\mathbf{n},1)} \left|\mathrm{Cov}\left(\Omega_{\mathbf{i}}^{\mathfrak{x}}, \Omega_{\mathbf{j}}^{\mathfrak{x}}\right)\right|$. From hypothesis (A1) and Eq (a8), it follows

$$\mathrm{Var}\left[\Omega_{\mathbf{i}}^{\mathfrak{x}}\right] \leq C\left(\zeta_z(h_L) + (\zeta_z(h_L))^2\right).$$

As a result, we have

$$R_{\mathbf{n}} = O\left(q_{\mathbf{n}}^M \zeta_z(h_L)\right).$$

Then, we introduce the following sets for $T_{\mathbf{n}}$:

$$F_1 = \{0 < ||\mathbf{i}-\mathbf{j}|| \leq k_{\mathbf{n}}, \mathbf{i}, \mathbf{j} \in \mathcal{I}(1,\mathbf{n},1)\} \text{ and}$$
$$F_2 = \{||\mathbf{i}-\mathbf{j}|| > k_{\mathbf{n}}, \mathbf{i}, \mathbf{j} \in \mathcal{I}(1,\mathbf{n},1)\},$$

with the real sequences $k_{\mathbf{n}}$ diverge to $+\infty$ and that is going to be explained afterward. By dividing the sum $T_{\mathbf{n}}$ into two separate different summations over sites under F_1 and F_2, we are able to write

$$T_{\mathbf{n}} = \sum_{(\mathbf{i},\mathbf{j})\in F_1} \left|\mathrm{Cov}\left(\Omega_{\mathbf{i}}^{\mathfrak{x}}, \Omega \mathbf{j}^{\mathfrak{x}}\right)\right| + \sum_{(\mathbf{i},\mathbf{j})\in F_2} \left|\mathrm{Cov}\left(\Omega_{\mathbf{i}}^{\mathfrak{x}}, \Omega_{\mathbf{j}}^{\mathfrak{x}}\right)\right| = T_{\mathbf{n}}^1 + T_{\mathbf{n}}^2.$$

We have on the one side

$$T_{\mathbf{n}}^1 = C \sum_{(\mathbf{i},\mathbf{j})\in F_1} \left|\mathbb{E}\left[L_{\mathbf{i}} L_{\mathbf{j}}\right]\right| + \left|\mathbb{E}\left[L_{\mathbf{i}}\right]\mathbb{E}\left[L_{\mathbf{j}}\right]\right| \leq C q_{\mathbf{n}}^M k_{\mathbf{n}}^M \zeta_z(h_L)\left((\zeta_z(h_L))^{1/\nu} + \zeta_z(h_L)\right) \leq C q_{\mathbf{n}}^M k_{\mathbf{n}}^M \zeta_z(h_L)^{(\nu+1)/\nu}.$$

However, we also have

$$T_{\mathbf{n}}^2 = \sum_{(\mathbf{i},\mathbf{j})\in E_2} \left|\mathrm{Cov}\left(\Omega_{\mathbf{i}}^{\mathfrak{x}}, \Omega \mathbf{j}^{\mathfrak{x}}\right)\right|.$$

As a result of (Tran et al. [40], Lemma 2.1(ii)), we have

$$\left|\mathrm{Cov}\left(\Omega_{\mathbf{i}}^{\mathfrak{x}}, \Omega_{\mathbf{j}}^{\mathfrak{x}}\right)\right| \leq C\Phi(||\mathbf{i}-\mathbf{j}||).$$

Consequently,

$$T_{\mathbf{n}}^2 \leq C \sum_{(\mathbf{i},\mathbf{j})\in F_2} \Phi(||\mathbf{i}-\mathbf{j}||) \leq C q_{\mathbf{n}}^M \sum_{\mathbf{i}:||\mathbf{i}||\geq k_{\mathbf{n}}} \Phi(||\mathbf{i}||) \leq C q_{\mathbf{n}}^M k_{\mathbf{n}}^{-M\nu} \sum_{\mathbf{i}:||\mathbf{i}||\geq k_{\mathbf{n}}} ||\mathbf{i}||^{M\nu} \Phi(||\mathbf{i}||).$$

By taking $k_{\mathbf{n}} = (\zeta_z(h_L))^{-1/M\nu}$, we have

$$T_{\mathbf{n}}^2 \leq C q_{\mathbf{n}}^M k_{\mathbf{n}}^{-M\nu} \sum_{\mathbf{i}:||\mathbf{i}||\geq k_{\mathbf{n}}} ||\mathbf{i}||^{M\nu} \Phi(||\mathbf{i}||) \leq C q_{\mathbf{n}}^M \zeta_z(h_L) \sum_{\mathbf{i}:||\mathbf{i}||\geq k_{\mathbf{n}}} ||\mathbf{i}||^{M\nu} \Phi(||\mathbf{i}||).$$

As a result, it follows from Eq (2.3) that

$$T_{\mathbf{n}}^2 \leq C q_{\mathbf{n}}^N \zeta_z(h_L).$$

Additionally, using the same $k_{\mathbf{n}}$ choice as before, we get

$$T_{\mathbf{n}}^1 \leq C q_{\mathbf{n}}^M \zeta_z(h_L).$$

$$\mathrm{Var}\left[H_1^*\right] = O\left(q_{\mathbf{n}}^M \zeta_z\left(h_L\right)\right).$$

For any $h_L \in \left(\hat{a}_{\mathbf{n}}, \hat{b}_{\mathbf{0}}\right)$ and $b_\delta \in \left(\hat{c}_{\mathbf{n}}, \hat{d}_{\mathbf{0}}\right)$, this final outcome in combination with the definitions of $q_{\mathbf{n}}$,β and η_0 is sufficient to demonstrate

$$Q_1(\mathbf{n}) \leq e^{\left(-C(\eta_0)\log\hat{\mathbf{n}}\right)}.$$

Hence, a suitable selection of η_0 enables us to show that

$$\sum_{\mathbf{n}} s_{\mathbf{n}} Q_1(\mathbf{n}) < \infty.$$

Concerning the expression $Q_2(\mathbf{n})$, and due to the Markov inequality, we may create

$$\begin{aligned} Q_2(\mathbf{n}) =& \mathbb{P}\left(\sum_{j=1}^{\beta}\left|H_{\mathbf{j}}^{\mathfrak{X}} - H_{\mathbf{j}}^*\right| \geq \frac{\eta\hat{\mathbf{n}}\zeta_z\left(h_L\right)}{2}\right) \leq \frac{1}{\eta\hat{\mathbf{n}}\zeta_z\left(h_L\right)}\sum_{j=1}^{\beta}\mathbb{E}\left|H_{\mathbf{j}}^{\mathfrak{X}} - H_{\mathbf{j}}^*\right| \\ \leq& 2\frac{\beta q_{\mathbf{n}}^M}{\eta\hat{\mathbf{n}}\zeta_z\left(h_L\right)}\Psi\left((\beta-1)q_{\mathbf{n}}^M, q_{\mathbf{n}}^M\right)\Phi\left(q_{\mathbf{n}}\right). \end{aligned}$$

Afterwards, because $\hat{\mathbf{n}} = 2^M\beta q_{\mathbf{n}}^M$ and $\Psi\left((\beta-1)q_{\mathbf{n}}^M, q_{\mathbf{n}}^M\right) \leq q_{\mathbf{n}}^M$, if we consider $\eta = \eta_0\left(\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}\zeta_z\left(h_L\right)}\right)^{1/2}$, we get

$$Q_2(\mathbf{n}) \leq \hat{\mathbf{n}}q_{\mathbf{n}}^M(\log\hat{\mathbf{n}})^{-1/2}\left(\hat{\mathbf{n}}\zeta_z\left(h_L\right)\right)^{-1/2}\Phi\left(q_{\mathbf{n}}\right).$$

Furthermore, write $q_{\mathbf{n}} = C\left(\frac{\hat{n}\zeta_z\left(h_L\right)}{\log\hat{n}}\right)^{1/2M}$, to obtain

$$Q_2(\mathbf{n}) \leq \hat{\mathbf{n}}\Phi\left(q_{\mathbf{n}}\right). \tag{a8}$$

The hypothesis (A8) guarantees that

$$\sum_{\mathbf{n}} s_{\mathbf{n}} Q_2(\mathbf{n}) < \infty.$$

The demonstration of Eq (a5) is the same as to the one utilised for (a4) with $\Upsilon_{\mathbf{i}}^{\mathfrak{X}}$ variables and by taking $q_{\mathbf{n}} = C\left(\frac{\hat{n}b_\delta}{\log\hat{n}}\right)^{1/2M}$.

Now, let us demonstrate (a6) and (a7) results. It is sufficient to show that

$$\mathrm{Cov}\left(D_1, D_2\right) = \frac{1}{\hat{\mathbf{n}}^2\zeta_z\left(h_L\right)}\sum_{\mathbf{i}\in\mathcal{L}_{\mathbf{n}}}\mathrm{Cov}\left(L_{\mathbf{i}}\delta_{\mathbf{i}}', \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^2}{h_L^2}\right) + \frac{1}{\mathbf{n}^2\zeta_z^2\left(h_L\right)}\sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{L}_{\mathbf{n}}}\mathrm{Cov}\left(L_{\mathbf{j}}\delta_{\mathbf{j}}', \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^2}{h_L^2}\right),$$

and

$$\mathrm{Cov}\left(D_3, D_4\right) = \frac{1}{\hat{\mathbf{n}}^2\zeta_z^2\left(h_L\right)}\sum_{\mathbf{i}\in\mathcal{L}_{\mathbf{n}}}\mathrm{Cov}\left(\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}\delta_{\mathbf{i}}'}{h_L}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}}{h_L}\right) + \frac{1}{\hat{\mathbf{n}}^2\zeta_z^2\left(h_L\right)}\sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{L}_{\mathbf{n}}}\mathrm{Cov}\left(\frac{L_{\mathbf{j}}\chi_{\mathbf{j}}\delta_{\mathbf{j}}'}{h_L}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}}{h_L}\right).$$

Notice that, using (A4) and (A6) assumptions, for $(\mathfrak{X}\prime, l\prime) \in \{(0,2)\}$, and for $(\mathfrak{X}\prime, l\prime) \in \{(1,1)\}$, we obtain

$$\left|\mathrm{Cov}\left(\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{\mathfrak{X}\prime}\delta_{\mathbf{i}}'}{h_L^{\mathfrak{X}\prime}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l\prime}}{h_L^{l\prime}}\right)\right| = \left|\mathbb{E}\left[\frac{L_{\mathbf{i}}^2\chi_{\mathbf{i}}^{\mathfrak{X}\prime+l\prime}\delta_{\mathbf{i}}'}{h_L^{\mathfrak{X}\prime+l\prime}}\right] - \mathbb{E}\left[\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{\mathfrak{X}\prime}\delta_{\mathbf{i}}'}{h_L^{\mathfrak{X}\prime}}\right]\mathbb{E}\left[\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l\prime}}{h_L^{l\prime}}\right]\right|.$$

Through $Z_{\mathbf{i}}$ conditioning, we have that

$$\mathbb{E}\left[\delta'_{\mathbf{i}} \mid Z_{\mathbf{i}}\right] = O\left(b_{\delta}\right) \text{ and } \mathbb{E}\left[\delta'_{\mathbf{i}}\delta'_{\mathbf{j}} \mid \left(Z_{\mathbf{i}}, Z_{\mathbf{j}}\right)\right] = O\left(b_{\delta}^{2}\right), \text{ for all } \mathbf{i} \neq \mathbf{j}.$$

Consequently, we get

$$\sum_{\mathbf{i}\in\mathcal{L}_{\mathbf{n}}} \operatorname{Cov}\left(\frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{\mathfrak{X}'}\delta'_{\mathbf{i}}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l'}}{h_L^{l'}}\right) = O\left(\hat{\mathbf{n}} b_{\delta}\zeta_{Z}\left(h_L\right)\right).$$

Then, the following quantity must be established:

$$\sum_{\mathbf{j}\neq\mathbf{i}\in\mathcal{L}_{\mathbf{n}}} \operatorname{Cov}\left(\frac{L_{\mathbf{j}}\chi_{\mathbf{j}}^{\mathfrak{X}'}\delta'_{\mathbf{j}}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l'}}{h_L^{l'}}\right).$$

We take into consideration the following decomposition:

$$\sum_{\mathbf{j}\neq\mathbf{i}\in\mathcal{L}_{\mathbf{n}}} \operatorname{Cov}\left(\frac{L_{\mathbf{j}}\chi_{\mathbf{j}}^{\mathfrak{X}'}\delta'_{\mathbf{j}}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l'}}{h_L^{l'}}\right) = \sum_{\substack{\mathbf{i}\neq\mathbf{i}\\(\mathbf{i},\mathbf{j})\in F_1}} \operatorname{Cov}\left(\frac{L_{\mathbf{j}}\chi_{\mathbf{j}}^{\mathfrak{X}'}\delta'_{\mathbf{j}}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l'}}{h_L^{l'}}\right) + \sum_{\substack{\mathbf{j}\neq\mathbf{i}\\(\mathbf{i},\mathbf{j})\in F_2}} \operatorname{Cov}\left(\frac{L_{\mathbf{j}}\chi_{\mathbf{j}}^{\mathfrak{X}'}\delta'_{\mathbf{j}}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l'}}{h_L^{l'}}\right).$$

$$\sum_{(\mathbf{i},\mathbf{j})\in F_1} \operatorname{Cov}\left(\frac{L_{\mathbf{j}}\delta'_{\mathbf{j}}\chi_{\mathbf{j}}^{\mathfrak{X}'}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l'}}{h_L^{l'}}\right) \leq \mathbf{C}\mathbf{n}\, k_{\mathbf{n}}^{M} b_{\delta}^{2}\zeta_{z}\left(h_L\right)^{(\nu+1)/\nu},$$

and

$$\sum_{(\mathbf{i},\mathbf{j})\in F_2} \operatorname{Cov}\left(\frac{L_{\mathbf{j}}\chi_{\mathbf{j}}^{\mathfrak{X}'}H'_{\mathbf{j}}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}''}{h_L''}\right) \leq \mathrm{C}\hat{\mathrm{n}}\, k_{\mathbf{n}}^{-M\nu} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq k_{\mathbf{n}}} \|\mathbf{i}\|^{M\nu}\Phi(\|\mathbf{i}\|).$$

By selecting $k_{\mathbf{n}} = \left(b_{\delta}\zeta_{z}\left(h_L\right)\right)^{-1/M\nu}$, we get

$$\sum_{\mathbf{j}\neq\mathbf{i}\in\iota_{\mathbf{n}}} \operatorname{Cov}\left(\frac{L_{\mathbf{j}}\chi_{\mathbf{j}}^{\mathfrak{X}'}\delta'_{\mathbf{j}}}{h_L^{\mathfrak{X}'}}, \frac{L_{\mathbf{i}}\chi_{\mathbf{i}}^{l'}}{h_L''}\right) = O\left(\hat{\mathbf{n}} b_{\delta}\zeta_{z}\left(h_L\right)\right).$$

For $\mathfrak{X}' = 0$ and $l' = 2$ with the condition (A7), we conclude that

$$\operatorname{Cov}\left(D_1, D_2\right) = O\left(\frac{1}{\hat{\mathbf{n}} b_{\delta}\zeta_{z}\left(h_L\right)}\right) = o\left(\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}} b_{\delta}\zeta_{z}\left(h_L\right)}\right)^{1/2},$$

then, for $\mathfrak{X}' = 1$ and $\mathfrak{X}' = 1$, we deduce that

$$\operatorname{Cov}\left(D_3, D_4\right) = O\left(\frac{1}{\hat{\mathbf{n}} b_{\delta}\zeta_{z}\left(h_L\right)}\right) = o\left(\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}} b_{\delta}\zeta_{z}\left(h_L\right)}\right)^{1/2}.$$

Finally, we have

$$\sup_{w\in\mathcal{S}} \sup_{\hat{a}_{\mathbf{n}}\leq h_L\leq\hat{b}_{\mathbf{n}}} \sup_{\hat{c}_{\mathbf{n}}\leq b_{\delta}\leq\hat{d}_{\mathbf{n}}} \left|\hat{f}_N\left(r_{j(w)} \mid z\right) - \mathbb{E}\hat{f}_N\left(r_{j(w)} \mid z\right)\right| = O_{a.co.}\left(\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}} b_{\delta}\zeta_{z}\left(h_L\right)}}\right). \tag{a9}$$

Then, the proof of this lemma can be obtained from (a1), (a2) and (a9).

Proof of Theorem 3.2

By the unimodality of $f(. \mid z)$, the assumption (A10)(i2) allows us to indicate that $f(\theta(z) \mid z(l)) = f(\tilde{\theta}(z) \mid z(l)) = 0$. In addition, by expanding the function $f(. \mid z)$ at $\theta(z)$ using the Taylor procedure and for any $h_L \in \left(\hat{a}_{\mathbf{n}}, \hat{b}_{\mathbf{n}}\right)$ and $b_\delta \in \left(\hat{c}_{\mathbf{n}}, \hat{d}_{\mathbf{n}}\right)$, we have, with $\theta^\bullet(z)$ within $\theta(z)$ and $\tilde{\theta}(z)$, that

$$f(\tilde{\theta}(z) \mid z) = f(\theta(z) \mid z) + \frac{1}{j!} f(\theta^\bullet(z) \mid z(j))(\tilde{\theta}(z) - \theta(z))^j. \tag{a10}$$

Next, we obtain by easy manipulation that

$$\left|f(\tilde{\theta}(z) \mid z) - f(\theta(z) \mid z)\right| \leq 2 \sup_{w \in \mathcal{S}} \sup_{\hat{a}_{\mathbf{n}} \leq h_L \leq \hat{b}_{\mathbf{n}}} \sup_{\hat{c}_{\mathbf{n}} \leq b_\delta \leq \hat{d}_{\mathbf{n}}} \left|\tilde{f}(w \mid z) - f(w \mid z)\right|. \tag{a11}$$

In fact, the proof of the following assertion is needed to conclude the proof of Theorem 3.2.

$$\lim_{\mathbf{n} \to \infty} |\tilde{\theta}(z) - \theta(z)| = 0, \text{ a.co.}$$

Proof. Assuming the function $f(. \mid z)$ is continuous, it follows that

$$\forall \varepsilon > 0, \ \exists \delta(\varepsilon) > 0, \ \ | f(w \mid z) - f(\theta(z) \mid z)| \leq \delta(\varepsilon) \Rightarrow |w - \theta(z) | \leq \varepsilon.$$

This last result implies that

$$\forall \varepsilon > 0, \ \exists \delta(\varepsilon) > 0, \ \mathbb{P}(|\tilde{\theta}(z) - \theta(z)| > \varepsilon) \leq \mathbb{P}\left(\left|f(\tilde{\theta}(z) \mid z) - f(\theta(z) \mid z)\right| > \delta(\varepsilon)\right). \tag{a12}$$

Finally, the stated result can be determined by combining (a11), (a12) and Theorem 3.1.

Now, we will return to the demonstration of Theorem 3.2. Since

$$f(\theta(z) \mid z(j)) \to f(\theta(z) \mid z),$$

and by using (A10)(i3), we obtain

$$\exists c > 0, \ \sum_{n=1}^{\infty} \mathbb{P}\left(|f(\theta^\bullet(z) \mid z(j))| < c\right) < \infty.$$

Hence, we have

$$|\tilde{\theta}(z) - \theta(z)|^j = O\left(\sup_{w \in \mathcal{S}} \sup_{\hat{a}_{\mathbf{n}} \leq h_L \leq \hat{b}_{\mathbf{n}}} \sup_{\hat{c}_{\mathbf{n}} \leq b_\delta \leq \hat{d}_{\mathbf{n}}} \left|\tilde{f}(w \mid z) - f(w \mid z)\right|\right), \text{ a.co.}$$

This final result is obtained by the mixing statements (a10)–(a12). Thus, the proof of Theorem 3.2 is directly deduced from the Theorem 3.1 and the Proposition 3.1.

AIMS Press