Mathematics

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

AIMS Mathematics, 9(3): 5550–5581.
DOI: 10.3934/math.2024269
Received: 12 December 2023
Revised: 17 January 2024
Accepted: 23 January 2024
Published: 29 January 2024

Research article

Strong consistency rate in functional single index expectile model for spatial data

Zouaoui Chikr Elmezouar[1], **Fatimah Alshahrani**[2], **Ibrahim M. Almanjahie**[1], **Salim Bouzebda**[3,*], **Zoulikha Kaid**[1] **and Ali Laksaci**[1]

[1] Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia; zchikrelemezouar@kku.edu.sa, imalmanjahi@kku.edu.sa, zqayd@kku.edu.sa, alikfa@kku.edu.sa

[2] Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P. O. Box 84428, Riyadh 11671, Saudi Arabia; fmalshahrani@pnu.edu.sa

[3] Département Génie Informatique, LMAC (Laboratoire de Mathématiques Appliquées de Compiègne), Université de Technologie de Compiègne, Avenue de Landshut 57, France

* **Correspondence:** Email: salim.bouzebda@utc.fr.

Abstract: Analyzing the real impact of spatial dependency in financial time series data is crucial to financial risk management. It has been a challenging issue in the last decade. This is because most financial transactions are performed via the internet and the spatial dependency between different international stock markets is not standard. The present paper investigates functional expectile regression as a spatial financial risk model. Specifically, we construct a nonparametric estimator of this functional model for the functional single index regression (FSIR) structure. The asymptotic properties of this estimator are elaborated over general spatial settings. More precisely, we establish Borel-Cantelli consistency (BCC) of the constructed estimator. The latter is obtained with the precision of the convergence rate. A simulation investigation is performed to show the easy applicability of the constructed estimator in practice. Finally, real data analysis about the financial data (Euro Stoxx-50 index data) is used to illustrate the effectiveness of our methodology.

Keywords: functional spatial data; complete convergence (a.co.); kernel estimator; expectile function; functional index; bandwidth parameter; financial risk management
Mathematics Subject Classification: 62G05, 62G08, 62R20

1. Introduction

Numerous fields of study, including econometrics, epidemiology, environmental science, image analysis, oceanography, meteorology, geostatistics, etc., frequently create spatial data. Typically, these data are gathered in numerous fields and analyzed statistically at measurement sites. Consult [1–5], as well as the references contained within, to identify credible sources of references to the research literature in this field and to learn about certain statistical applications. We highlight that modeling a spatio-temporal interaction of financial data is crucial in financial risk management. The momentousness of this matter is motivated by the digitalization of financial transactions. Actually, with technological development, most financial transactions are carried out by the internet allowing for an increase in the impact of the spatial interaction between financial institutes in financial movements. On the other hand with these new instruments in financial operations, the spatial correlation of the financial time series data is not standard. It is related to many additional factors than the geographical position. Indeed, the spatio-temporal feature in financial time is impacted by the economic exchange between countries, the relationship between economic sectors, and the mobility between countries, among others. Thus, analyzing the spatio-temporal features of financial data is tough work. Motivated by this challenge, we combine the ideas of functional single index modeling with the recent financial risk techniques to provide a statistical model that allows us to fit the spatio-temporal feature of financial data in risk management. Noting that it is well-recognized that spatio-temporal modeling is a particular case of spatio-functional data analysis, the single index model is more appropriate in econometrics and financial areas. Thus, it will be very interesting to utilize recent developments in spatio-functional statistics to introduce a new financial risk model based on a single index structure. Indeed, the FSIR model is one of the key tools in econometrics as well as in financial time series data. In particular, for financial areas, this model is often used to reduce the high number of factors of an investment or to determine the principal assets in a given portfolio. From a theoretical point of view, the single index model belongs to the semiparametric family of models, which many authors have studied from practical and theoretical point of view, for instance, see [6, 7]. We return to [8, 9] for the first results in the vectorial explanatory case. In the functional setting, the authors of [10] propose the Nadaraya-Watson-kernel-estimator (NWKE) for the nonparametric part in a functional single index regression (FSIR) structure. They proved the consistency of the NWKE when the covariate pertains to the Hilbertian subspace. In the last few decades, the popularity of these models has increased. We cite, for instance, [11] who introduced a new estimation in single index modeling. They proposed a multi-index fitting aproach adaptable to linear projections for functional data and show that their method makes it possible to predict with polynomial convergence rates. Recently, [12] generalized the functional parametric regression model to partially linear functional single index models. More recent advances in functional single index models were obtained by [13]. They used the k-nearest neighbors algorithm to estimate the nonparametric link function of the FSIR. They elaborated the Borel-Cantelli consistency (BCC) of the NWKE using a quasi-associated dependence structure.

In this paper, we investigate conditional exectiles, which is based on least asymmetrically weighted squares estimation, which was adopted from the econometrics literature and is a fundamental statistical application tool. This method frequently employs the [14] concept of expectiles, the least-squares equivalent of the conventional quantiles. They were given this name because they resemble the quantiles of a random variable, but, unlike quantiles, they are based on a quadratic loss function, as in the case of the expectation; see [15, 16] for more information. Since it is the only elicitable coherent risk measure, we refer to [17] and its references. Refer to [18] for applying the expectile regression in heteroscedasticity analysis. We refer you to the recent paper by [19, 20] for further justification of the expectile model's application. For an overview of the use of expectile curves in regression analysis, see [21] and the extensive discussions of that paper, especially the research [22] for an evaluation of expectiles and [23] for a critical perspective. Despite their disparities in construction, quantiles and expectiles share similar characteristics. As demonstrated by the research of [24], the primary reason is that expectiles are identical to quantiles if the original distribution is transformed. Quantiles and expectiles, which comprise information about a random variable's complete distribution, are extensions of the median and mean, respectively. Expectiles are superior replacements for quantiles in a variety of pertinent applications. Motivating advantages include the fact that expectiles are more sensitive than quantiles to the magnitude of infrequent catastrophic losses and that they depend on both the tail realizations of the predictor and their probability, whereas quantiles depend only on the frequency of tail realizations. This sensitivity of expectiles to tail behavior enables more prudent and responsive risk management. Observe that the quantiles are not always adequate and can be criticized for being challenging to compute because the corresponding loss function is not continuously differentiable. Of course, these features increase the importance of the expectile in some specific areas, namely, in financial risk analysis. However, the robustness of the quantile is also a good advantage in some alternative areas, namely for the prediction issue. From a statistical point of view, there is a bothersome trade-off between sensitivity and robustness. Thus, it will be interesting to develop a bridge between the two phenomenons. This is the main motivation of the expectile regression as an alternative model to the quantile when the sensitivity is required. The functional expectile with regression was recently introduced by [20, 25]. They demonstrated almost complete consistency with the rate and the limit distribution of NWKE of this functional model. Such results were obtained under the independence structure. The last contribution was extended to the mixing situation by [19]. The authors of [26] stated the BCC of the functional NWKE when the functional time is ergodic. The spatial structure was handled by [27]. It should be noted that this last situation is of great importance in practice. The framework of the present contribution concerns the nonparametric spatial statistic. The pioneer works of this domains are the reaearches [15, 16, 28–31], among other. This subject was investigated for the first time in functional statistics by [32]. They proved the BCC functional NWKE is the regression operator.

The third axe of this contribution concerns the spatio-functional data analysis (SFDA). Such a topic is relatively recent. It is classified by [33] as second-generation functional data analysis. Of course, the main difficulty of this topic in the analysis of spatial data comes from the fact that observations indexed in the multi-dimensional space do not have a linear order, as it happens for classical time series data. For a brief literature review in SFDA, we cite the first pioneer work [34]. They consider the problem of the functional linear regression estimation by the spline method when the observations are spatially correlated. Among the wide range of applications of SFDA we mention the hyperspatial imagery processing developed by [35]. We return to [36] for more advanced application of SFDA in the environmental area. The mathematical development of nonparametric spatio-functional data analysis was started by [37]. They state the asymptotic property of the kernel estimation of the density of a functional random field. We refer to [38] for spatial local linear estimation in functional data.

Recently, the authors of [39] treated spatio-functional quantile regression using different techniques, both parametric and nonparametric. However, despite the great importance of the semi-parametric modeling of spatio-functional data in practice, this problem has not been explored yet. To the best of our knowledge, the present contribution is the first work in this direction. It is worth noting that, the functional data analysis (FDA) is actually in continuous progression. For an overview of the recent developments and trends in this area, we may refer the reader to some journal special issues devoted to this topic, such as [40–45] to cite a few.

Precisely, the main aim of the present work is to estimate the spatial expectile regression when the input and the output variables are linked with the FSIR structure. We construct an estimator of the nonparametric part of this link function using kernel weighting. The consistency with the rate of the constructed estimator is the main asymptotic result of this work. Moreover, the main novelty of this contribution is the treatment of this model under the FSIR structure. From a practical point of view, this kind of semiparametric model is very common in the econometric area. Such popularity is motivated by two important features. The first is its characteristic as an excellent reducer of the data dimension, and the second feature is the easy interpretation of the co-variability between the output variable and the regressor of the functional index included in these models. The estimability of the functional index is discussed using two cross-validation rules. The first is based on the weighted least squared error, and the second is obtained by the maximum likelihood function. A simulation investigation is conducted to compare these criteria. Finally, we highlight the great impact of the present contribution in financial time series data via a real-data application, allowing us to show the superiority of our spatial model over its competitors and to explore the spatial interaction in financial data. Recall that the flexibility of the additive model and the importance of the spatial interaction of the financial activities are the principal motivations to investigate the expectile by regression in spatio-functional SIM. Thus, our study allows us to digitalize the spatial co-movement in financial time series data using new algorithms adapted to high-frequency data observed over a thinner discretization grid. The new method proposed in this contribution constitutes an alternative approach to classical models based on quantile by regression and shortfall regression.

This paper is organized as follows: In the following sections, we introduce the model and the kernel estimator of the conditional expectile function. The main asymptotic results are stated in Section 3. Section 4.2 is devoted to the functional index θ's estimability issue. The applicability of considered rules will be examined in Section 5. The last section also includes a real world data application. Some concluding remarks are given in Section 6. Finally, the proofs of the main results are given in Appendix.

2. Materials and methods

2.1. Spatio-functional framework

Set $N \in \mathbb{N}^*$ and assume $(\mathcal{X}_\mathbf{i}, Y_\mathbf{i})$, $\mathbf{i} \in \mathbb{N}^N$ as a $\mathcal{F} \times \mathbb{R}$-valued strictly stationary spatial process. $(\mathcal{X}_\mathbf{i}, Y_\mathbf{i})$, $\mathbf{i} \in \mathbb{N}^N$ is defined in $(\Omega, \mathcal{A}, Pr)$, a given probability space. Specifically, $\mathcal{F}$ is a separable Hilbert space where $\langle \cdot, \cdot \rangle$ is the inner product and with an orthonormal basis denoted by $\{e_p : p \geq 1\}$. We suppose that the process $(\mathcal{X}_\mathbf{i}, Y_\mathbf{i})$ is observed under a rectangular area

$$I_\mathbf{n} = \{\mathbf{i} = (i_1, \cdots, i_N) \in \mathbb{N}^N : 1 \leq i_{k=1,\cdots,N} \leq n_{k=1,\cdots,N}\}.$$

We will write that the N-uplet in $\mathbb{N}^N$,

$$\mathbf{n} = (n_1, \cdots, n_N) \to \infty,$$

if $\min\{n_k\} \to \infty$ and $|n_j/n_k| < C$ for a constant $0 < C < \infty$, and for all j, k is less than N. Typically, we explore the spatial interaction of

$$Z_{\mathbf{i}} = (\mathcal{X}_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i} \in I_{n \in \mathbb{N}}}$$

by assuming that there exists a non-increasing function $\varphi(t)$ towards 0 as $t \to \infty$, such that, all subsets E' and E of $\mathbb{N}^N$ have finite cardinals

$$\begin{aligned} \alpha\Big(\mathcal{B}(E'), \mathcal{B}(E)\Big) &= \sup_{B \in \mathcal{B}(E'),\, C \in \mathcal{B}(E)} \Big| Pr(B)Pr(C) - Pr(B \cap C) \Big| \\ &\leq \psi\Big(\mathrm{Card}(E'), \mathrm{Card}(E)\Big)\varphi\Big(\mathrm{dist}\,(E', E)\Big), \end{aligned} \tag{2.1}$$

where $\mathrm{Card}(E')$ and $\mathrm{Card}(E)$ are the cardinality of E' and E, $\mathcal{B}(E')$ (resp. $\mathcal{B}(E)$) is the Borel σ-algebra generated from $\{Z_{\mathbf{i}} : \mathbf{i} \in E'\}$ (resp. $\{Z_{\mathbf{i}} : \mathbf{i} \in E\}$). The quantity $\mathrm{dist}\Big(E', E\Big)$ represents the Euclidean metric between E' and E. The function

$$\psi : \mathbb{N}^2 \to \mathbb{R}^+$$

is a positive symmetric function and nondecreasing such that, for all $m', m \in \mathbb{N}$,

$$\psi(m', m) \leq C \min(m', m), \tag{2.2}$$

for some $C > 0$. Additionally, the function φ satisfies

$$\sum_{i=1}^{\infty} i^{\delta} \varphi(i) < \infty \text{ for some } \delta > 0. \tag{2.3}$$

All these assumptions that characterize the spatio-functional framework of this study are standards. They are similar to those used by [16, 46]. It should be pointed out that if $N = 1$, then $(X_{\mathbf{i}}, Y_{\mathbf{i}})$ is called strongly mixing (see [47] for discussion on mixing and examples). Furthermore, the spatial linear process [48] satisfies this mixing assumption (see [49], for more details on the required conditions). We refer to [1, 50–53] for additional information on mixing coefficients for random fields.

2.2. The FSIR structure of the spatial expectile

Now, we assume that the behavior of $Y_{\mathbf{i}}$ is linked to $\mathcal{X}_{\mathbf{i}}$ through the FSIR model with functional θ, a fixed index in $\mathcal{F}$. Therefore, $Y_{\mathbf{i}}$ and $\mathcal{X}_{\mathbf{i}}$ are linked by

$$\mathbb{E}[Y_{\mathbf{i}} \mid \mathcal{X}_{\mathbf{i}}] = \mathbb{E}[Y_{\mathbf{i}} \mid \langle \theta, \mathcal{X}_{\mathbf{i}} \rangle], \quad \theta \text{ in } \mathcal{F}. \tag{2.4}$$

The model identifiability has been examined by [10]. A sufficient condition of the FSIR-identifiability is:

(i) The link function $r(x) = \mathbb{E}[Y_{\mathbf{i}} \mid \mathcal{X}_{\mathbf{i}} = x]$ is differentiable.

(ii) The functional index θ is such that $\langle \theta, e_1 \rangle = 1$.

The goal of this paper is to estimate the q^{th} expectile with regression of Y given $X = \mathfrak{z}$ defined, for $0 < q < 1$, by

$$\xi_q(\theta, \mathfrak{z}) = \arg\min_{t\in R} \Gamma_q(Y, \theta, \mathfrak{z}, t),$$

where

$$\begin{aligned}\Gamma_q(Y, \theta, \mathfrak{z}, t) &= \mathbb{E}\left[q(Y-t)^2 \mathbb{1}_{(Y-t)>0} \mid \langle \mathfrak{z}, \theta\rangle\right] \\ &+ \mathbb{E}\left[(1-q)(Y-t)^2 \mathbb{1}_{(Y-t)\le 0} \mid \langle \mathfrak{z}, \theta\rangle\right]. \end{aligned} \tag{2.5}$$

It is worth noticing that (2.5) generalizes the conditional *expec*tation of Y given $\langle \mathfrak{z}, \theta\rangle$, which coincides with $\xi_q(\theta, \mathfrak{z})$, specifically when $q = 1/2$ and such as the conditional quantile generalizes the conditional median. On the other hand, (2.5) is similar to the conditional q-quan*tile* of Y given $X = x$, which can be obtained by replacing $(Y - t)^2$ by $|Y - t|$ in (2.5). Hence, the name conditional q-*expectile*. By straightforward calculus, we demonstrate that $\xi_q(\theta, \mathfrak{z})$ is a zero of

$$\frac{q}{1-q} = \frac{G_1(\theta, \mathfrak{z}, t)}{G_2(\theta, \mathfrak{z}, t)},$$

where

$$\begin{aligned} G_1(\theta, \mathfrak{z}, t) &= -\mathbb{E}\left[(Y-t)\mathbb{1}_{(Y-t)\le 0} \mid \langle \mathfrak{z}, \theta\rangle\right], \\ G_2(\theta, \mathfrak{z}, t) &= \mathbb{E}\left[(Y-t)\mathbb{1}_{(Y-t)>0} \mid \langle \mathfrak{z}, \theta\rangle\right]. \end{aligned}$$

Since the function

$$G(\theta, \mathfrak{z}, t) := \frac{G_1(\theta, \mathfrak{z}, t)}{G_2(\theta, \mathfrak{z}, t)}$$

is a non-decreasing function, see for example [20], the expectile $\xi_q(\theta, \mathfrak{z})$ of order q is expressed as

$$\xi_q(\theta, \mathfrak{z}) = \inf\left\{t \in R: \ G(\theta, \mathfrak{z}, t) \ge \frac{q}{1-q}\right\}. \tag{2.6}$$

Finally, the spatial estimator of the q^{th} expectile with the regression of Y given $\mathcal{X}$ is

$$\widehat{\xi_q}(\theta, \mathfrak{z}) = \inf\left\{t \in R/ \ \widehat{G}(\theta, \mathfrak{z}, t) \ge \frac{q}{1-q}\right\}, \tag{2.7}$$

where

$$\widehat{G}(\theta, \mathfrak{z}, t) = \frac{\widehat{G_1}(\theta, \mathfrak{z}, t)}{\widehat{G_2}(\theta, \mathfrak{z}, t)}$$

with

$$\widehat{G_1}(\theta, \mathfrak{z}, t) = \frac{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(h_\mathbf{n}^{-1} IN_\theta(x - \mathcal{X}_\mathbf{i})\right) \mathbb{1}_{(Y_\mathbf{i}-t)\le 0}(Y_\mathbf{i} - t)}{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(h_\mathbf{n}^{-1} IN_\theta(x - \mathcal{X}_\mathbf{i})\right)}$$

and

$$\widehat{G_2}(\theta, \mathfrak{z}, t) = \frac{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(h_\mathbf{n}^{-1} N_\theta(x - \mathcal{X}_\mathbf{i})\right) \mathbb{1}_{(Y_\mathbf{i}-t)>0}(Y_\mathbf{i} - t)}{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(h_\mathbf{n}^{-1} N_\theta(x - \mathcal{X}_\mathbf{i})\right)},$$

where

$$N_\theta(z) = \langle z, \theta \rangle,$$

$K(\cdot)$ is a kernel function, and $h_{\mathbf{n}}$ is a sequence of positive real numbers tending to zero as $\mathbf{n}$ tends to infinity.

3. Results

Recall that our aim is the establishment of the BCC of the estimator $\widehat{\xi_q}(\theta, \mathfrak{z})$ to $\xi_q(\theta, \mathfrak{z})$ under the strong mixing structure (2.1). To accomplish this goal, we define

$$B(\theta, \mathfrak{z}, r) := \{x' \in \mathcal{F} \ /|N_\theta(x' - \mathfrak{z})| \leq r\} \quad \text{for} \quad r > 0.$$

Next, we introduce the following conditions:

(H1) $\forall r > 0, Pr(X \in B(\theta, \mathfrak{z}, r)) =: \phi(\theta, \mathfrak{z}, r) > 0$. Additionally, $\phi(\theta, \mathfrak{z}, r) \longrightarrow 0$ as $r \longrightarrow 0$.

(H2) The random field $(\mathcal{X}_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}}$ satisfies

$$\begin{cases} \mathbb{E}\Big[\big|Y_{\mathbf{i}}\big|^p \Big| N_\theta(\mathcal{X}_{\mathbf{i}})\Big] < C < \infty, \text{ for some } p > 4, \\ \text{For all } \mathbf{i} \neq \mathbf{j}, \ \mathbb{E}\Big[\big|Y_{\mathbf{j}} Y_{\mathbf{i}}\big| \Big| N_\theta(\mathcal{X}_{\mathbf{i}}), N_\theta(\mathcal{X}_{\mathbf{j}})\Big] \leq C < \infty, \\ \text{and} \\ 0 < \sup_{\mathbf{i} \neq \mathbf{j}} Pr\Big[(\mathcal{X}_{\mathbf{i}}, \mathcal{X}_{\mathbf{j}}) \in B(\theta, \mathfrak{z}, h) \times B(\theta, \mathfrak{z}, h)\Big] \leq C(\phi(\theta, \mathfrak{z}, h_{\mathbf{n}}))^{(a+1)/a}, \\ \text{for some } 0 < a < \delta N^{-1}. \end{cases}$$

(H3) The kernel function $K(\cdot)$ is supported in $[0, 1]$ and there exist C and $C' > 0$, such that

$$C \, \mathbb{1}_{(0,1)}(\cdot) \leq K(\cdot) \leq C' \, \mathbb{1}_{(0,1)}(\cdot).$$

(H4) The functions $G_l(\mathfrak{z}, \cdot, \cdot)$, for $l = 1, 2$, continuously-differentiable in $\mathbb{R}$, for all $(t_1, t_2) \in \mathbb{R}$ and $\mathcal{X}_1, \mathcal{X}_2 \in \mathcal{F}$, we have

$$\Big|G_l(\theta, \mathcal{X}_1, t_1) - G_l(\theta, \mathcal{X}_2, t_2)\Big| \leq C(\|\mathcal{X}_1 - \mathcal{X}_2\|^{k_l} + |t_1 - t_2|^{\varsigma_l}),$$

for some ς_l, $k_l > 0$.

(H5) There exists $(\gamma_{\mathbf{n}})$, a sequence of nonnegative real numbers, such that

$$\begin{cases} \gamma_{\mathbf{n}}^{-1} \phi^{(1-\beta)/\beta}(\theta, \mathfrak{z}, h_{\mathbf{n}}) \to 0 \text{ with } \beta = \dfrac{p-2}{p}, \\ \sum_{\mathbf{n}} \widehat{\mathbf{n}}^{((1+2\varsigma_1)/2\varsigma_1)} \gamma_{\mathbf{n}}^{-p} < \infty, \\ \sum_{\mathbf{n}} \widehat{\mathbf{n}}^{((1+2\varsigma_1)/2\varsigma_1) - \delta/2N} \gamma_{\mathbf{n}}^{\delta/N} \phi^{-\delta/2N}(\theta, \mathfrak{z}, h_{\mathbf{n}}) \log^{\delta/2N} \widehat{\mathbf{n}} < \infty, \end{cases}$$

where

$$\widehat{\mathbf{n}} = \prod_{k}^{N} n_k.$$

3.1. Comments on the assumptions

The assumptions are standard. Specifically, hypothesis (H1) is stated for several functional processes. For the same special case, see [54] or [55]. These works evaluated (H1) over some Gaussian processes. Condition (H2) is added to get the same convergence rate as in the i.i.d. setting. The assumption (H3) it also standard in this context of functional statistics. In particular, its technical assumption is verified by numerous standard kernels, such as Epanechnikov kernel, beta kernel, triangular kernel, and others (see for the same example [56]). (H4) is a moderate regularity postulate imposed to state the bias term. The requirements in (H5) are reasonable technical conditions to simplify the proofs.

The BCC of $\widehat{\xi}_q(\mathfrak{z})$, when $(\mathcal{X}_{\mathbf{i}}, Y_{\mathbf{i}})$ satisfies (2.2) and (2.3), is given in the following theorem.

Theorem 1. *Under hypotheses (H1)–(H5) and in addition, if we have*

$$\frac{\partial G(\theta, \mathfrak{z}, \xi_q(\mathfrak{z}))}{\partial t} > 0,$$

then,

$$\widehat{\xi}_q(\mathfrak{z}) - \xi_q(\mathfrak{z}) = O\left(h_{\mathbf{n}}^{k_1}\right) + O\left(h_{\mathbf{n}}^{k_2}\right) + O\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta, \mathfrak{z}, h_{\mathbf{n}})}\right)^{1/2}$$

in BCC-consistency-mode, as $\mathbf{n} \to \infty$.

We point out that the structure of the obtained convergence rate keeps its usual form of the functional kernel smoothing approach in the sense that it is decomposed into its principal terms. The first one is the bias term expressed with respect to the degree of the smoothing assumption in (H4). The second term is the stochastic term, which is expressed with respect to the functional structure through the function $\phi(\cdot)$ of the assumption (H1).

Proposition 3.1. *Under conditions (H1)–(H5), we have*

$$\sup_{t\in[\xi_q(\mathfrak{z})-\epsilon_0,\, \xi_q(\mathfrak{z})+\epsilon_0]} \left|\widehat{G}(\theta, \mathfrak{z}, t) - G(\theta, \mathfrak{z}, t)\right| = O\left(h_{\mathbf{n}}^{k_1}\right) + O\left(h_{\mathbf{n}}^{k_2}\right) + O\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta, \mathfrak{z}, h_{\mathbf{n}})}\right)^{1/2}$$

in BCC-consistency-mode, as $\mathbf{n} \to \infty$.

4. Discussion

4.1. The impact of the spatial dependency in practice

It is clear that the convergence rate of the spatial estimator $\widehat{\xi}_q$ is comparable to the i.i.d case. This statement follows from the observation that, in our theoretical study, we seek to determine the appropriate conditions to get a good asymptotic property of the estimator, namely those reduce the convergence rate of the estimator. However, in practice, such an optimal situation is not usually available. Thus, it is very interesting to examine the effect of the spatial correlation of the data on the computationablity of the constructed estimator. Indeed, in the nonfunctional case, the spatial correlation is evaluated through the covariogram or the variogram function; see [57]. Alternatively, in functional statistics, we use the trace-variogram function (see [58]) to examine the spatial correlation.

Specifically, we adopt the ideas of [32] by adding the spatial controller to the definition of the estimator. Indeed, we compute $\widehat{\xi_q}(\theta, X_{\mathbf{k}})$, for a new observation $X_{\mathbf{k}}$ in new site $\mathbf{k} \notin \mathbf{I_n}$ by replacing $\widehat{G_1}$ and $\widehat{G_2}$ in (2.7) with

$$\widehat{G_1}(\theta, \mathfrak{z}, t) = \frac{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} IN_\theta(x - \mathcal{X}_{\mathbf{i}})\right) \mathbb{1}_{\mathbf{W_k}}(\mathbf{i}) \mathbb{1}_{(Y_{\mathbf{i}}-t)\leq 0}(Y_{\mathbf{i}} - t)}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} IN_\theta(x - \mathcal{X}_{\mathbf{i}})\right) \mathbb{1}_{\mathbf{W_k}}(\mathbf{i})}$$

and

$$\widehat{G_2}(\theta, \mathfrak{z}, t) = \frac{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} N_\theta(x - \mathcal{X}_{\mathbf{i}})\right) \mathbb{1}_{\mathbf{W_k}}(\mathbf{i}) \mathbb{1}_{(Y_{\mathbf{i}}-t)> 0}(Y_{\mathbf{i}} - t)}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} N_\theta(x - \mathcal{X}_{\mathbf{i}})\right) \mathbb{1}_{\mathbf{W_k}}(\mathbf{i})},$$

where $\mathbf{W_k}$ is a vicinity set of the fixed site $\mathbf{k}$ defined by

$$\mathbf{W_k} = \{\mathbf{i}, \text{ such that } \gamma(\mathbf{i}, \mathbf{k}) \leq \iota_{\mathbf{n}}\}, \tag{4.1}$$

where γ is the trace-variogram function and $\iota_{\mathbf{n}}$ is a appropriate sequence of positive real numbers. We point out that the trace-variogram function γ is estimated empirically by

$$\widehat{\gamma}(\mathbf{l}, \mathbf{k}) = \frac{1}{2\#N_{\mathbf{l},\mathbf{k}}} \sum_{\mathbf{i},\mathbf{j}\in N_{\mathbf{l},\mathbf{k}}} d(X_{\mathbf{i}}, X_{\mathbf{j}})$$

with

$$N_{\mathbf{l},\mathbf{k}} = \{\mathbf{i}, \mathbf{j} \in I_{\mathbf{n}} \text{ such that } ||\mathbf{i} - \mathbf{j}|| = ||\mathbf{l} - \mathbf{k}||\},$$

and $\#N_{\mathbf{l},\mathbf{k}}$ is the cardinal of $N_{\mathbf{l},\mathbf{k}}$. Observe that the use of the trace-variogram function also allows for the integration of the functional nature of the data. Furthermore, in the isotropic case where the dependence is related only to the distance between the locations, we can proceed with the vicinity set

$$\mathbf{V_k} = \{\mathbf{i}, \text{ such that } \mathit{distance}(\mathbf{i}, \mathbf{k}) \leq \nu_{\mathbf{n}}\}, \tag{4.2}$$

where $\nu_{\mathbf{n}}$ is an appropriate sequence of positive real numbers. Of course, the distance here is the locating function between the different sites defined by the user.

4.2. Spatio-functional index estimation

This section is devoted to showing how we implement $\widehat{\xi_q}(\mathfrak{z})$ in practice. Naturally, the applicability of $\widehat{\xi_q}(\mathfrak{z})$ depends on the precision of the parameters utilized in the estimator. In this paragraph, we focus on the principal one, which is the single index θ. It is worth noting that the FSIR estimation has been developed by multiple authors, for instance, see [59,60]. However, in this study we will introduce and control the spatial structure of the data. Precisely, we will control this aspect over the three usual selector procedures mentioned below.

4.2.1. Spatial cross-validation by least squares error: SCVLSE-rule

The single index model is widely employed in econometrics. It is usually used to reduce the number of factors in econometric data analysis. In mathematical statistics, this kind of model belongs to the family of additive models. Theoretically, these techniques are used to improve the convergence rate of the nonparametric approach. We integrate this vicinity subset of the previous paragraph in the least squares rule to select the best index as

$$\hat{\theta} = \arg\min_{\theta\in\Theta\subset\mathcal{F}} \sum_{\mathbf{k}\in\mathcal{I}_\mathbf{n}} (Y_\mathbf{k} - \widehat{R}_\theta^{-\mathbf{k}}(X_\mathbf{k}))^2, \tag{4.3}$$

where $\widehat{R}_\theta^{-\mathbf{k}}$ is the leave-one-out estimator of the conditional expectation, which is defined by

$$\widehat{R}_\theta^{-\mathbf{k}}(X_\mathbf{k}) = \frac{\sum_{\mathbf{j}\in\mathcal{I}_\mathbf{n}-\{\mathbf{k}\}} K\left(h_\mathbf{n}^{-1} N_\theta((X_\mathbf{k}) - \mathcal{X}_\mathbf{j})\right) Y_\mathbf{j} \mathbb{1}_{\mathbf{W}_\mathbf{k}}(\mathbf{j})}{\sum_{\mathbf{j}\in\mathcal{I}_\mathbf{n}-\{\mathbf{k}\}} K\left(h_\mathbf{n}^{-1} N_\theta((X_\mathbf{k}) - \mathcal{X}_\mathbf{j})\right) \mathbb{1}_{\mathbf{W}_\mathbf{k}}(\mathbf{j})}.$$

As discussed in the second section, this rule is justified by the fact that conditional expectation can be viewed as a particular case of the q-expectile with $q = 0.5$. However, the rule (4.3) can be generalized for various orders q by taking

$$\hat{\theta} = \arg\min_{\theta\in\Theta\subset\mathcal{F}} \sum_{\mathbf{k}\in\mathcal{I}_\mathbf{n}} \rho_q(Y_\mathbf{k} - \widehat{\xi}_q^{-\mathbf{k}}(X_\mathbf{k})), \tag{4.4}$$

where

$$\rho_q(s) = |q - \mathbb{1}_{\{s<0\}}| s^2,$$

and $\widehat{\xi}_q^{-\mathbf{k}}$ is the leave one-out estimator of ξ_q constructed in the same manner as $\widehat{R}_\theta^{-\mathbf{k}}$.

4.2.2. Spatial cross-validation maximum likelihood: SCVML-rule

An alternative approach based on the maximum likelihood method is used to select the best optimal index. Indeed, with this rule, the best index model is realized by maximizing the conditional likelihood function

$$\theta = \arg\max_{\theta\in\Theta\subset\mathcal{F}} f(y \mid N_\theta(\mathcal{X})),$$

where $f(\cdot \mid \cdot)$ is the density of Y conditioning on $N_\theta(X)$. So, the practical determination of the single index is found on the nonparametric estimation of $f(\cdot \mid \cdot)$. Once again, to explore the spatial correlation, we integrate the same vicinity subset in the conditional density function

$$\widehat{f}(y|X_\mathbf{k}) = \frac{h_\mathbf{n}^{-1} \sum_{\mathbf{i}\in\mathbf{I_n}} K\left(h_\mathbf{n}^{-1} N_\theta(x - \mathcal{X}_\mathbf{i})\right) K\left(h_\mathbf{n}^{-1}(y - Y_\mathbf{i})\right) \mathbb{1}_{\mathbf{W}_\mathbf{k}}(\mathbf{i})}{\sum_{\mathbf{i}\in\mathbf{I_n}} K\left(h_\mathbf{n}^{-1} N_\theta(x - \mathcal{X}_\mathbf{i})\right) \mathbb{1}_{\mathbf{W}_\mathbf{k}}(\mathbf{i})}.$$

Therefore, we have

$$\hat{\theta} = \arg\max_{\theta\in\Theta\subset\mathcal{F}} \frac{1}{n} \sum_{i=1}^{n} \log \hat{f}(Y_i \mid \widehat{r}_{\theta,n}^{-i}(\mathcal{X}_i)). \tag{4.5}$$

This criterion was used by [61] in the nonfunctional case. The use of conditional density outperforms the first criterion (4.3) based on the conditional expectation. Since the conditional density is more informative than the conditional expectation, (4.5) seems to be more adequate than (4.3).

5. Computational study

5.1. A simulation study

The aim of this subsection is the evaluation of the effectiveness of the proposed estimator using a finite sample size. The goal is to show how the spatial interaction impacts the choice the functional-index θ as well as the smoothing parameter $h_{\mathbf{n}}$. Particularly, we will check the effect of spatial dependency over the two cross-validation rules of the previous sections. For this purpose, we simulate spatio-functional data using the SFIM as follows:

$$Y_{\mathbf{i}} = r_\theta(X_{\mathbf{i}}) + \epsilon_{\mathbf{i}}$$

with

$$r_\theta(\cdot) = r(N_\theta(\cdot)).$$

The function $r(\cdot)$ denotes a nonparametric regression link and $\epsilon_{\mathbf{i}}$ is a white noise spatial process that is supposed to be a Gaussian isotropic random field. The covariance function of this spatial process is

$$C(u) = e^{\left(\frac{-u^2}{2}\right)}.$$

For this experimental analysis, we generate the response variable by taking

$$r(\mathfrak{z}) = \int_0^1 \frac{1}{x^2(t) + 1}$$

and $\theta = e_1$ is the first element of the basis function of Karhunen-Loève decomposition. Thereby, θ is the eigenfunction that corresponds to the greatest eigenvalue of the covariance of the process $(X_{\mathbf{i}})_{\mathbf{i}}$. The latter is drawn from the formula

$$X_{\mathbf{i}} = \sin(\pi W_{\mathbf{i}} t) + W_{\mathbf{i}} t \cos(W_{\mathbf{i}} t), \quad t \in [0, 1],$$

where $W_{\mathbf{i}}$ is a random Weibull field with covariance that has an exponential function

$$C(u) = e^{(-\psi u)} \text{ for } u \geq 0.$$

The simulation result is displayed in Figure 1.

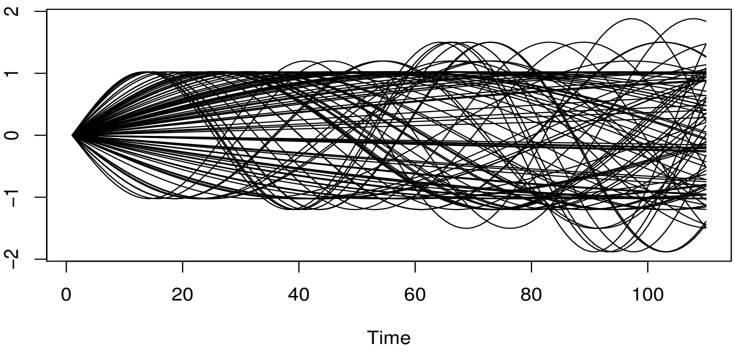

Figure 1. Functional covariates.

Recall that θ is unknown in practice. To estimate θ, we compare the two rules of the previous sections. Based on the two rules, we select the best functional index from Θ the finite subset defined by

$$\Theta = \Theta_{\mathbf{n}} = \left\{ \theta \in \mathcal{F}, \theta = \sum_{i=1}^{k_0} c_i e_i, \; \|\theta\| = 1, \text{ and } \exists j \in 1, \ldots, k \text{ such that } N_\theta(e_j) > 0 \right\},$$

where $(c_i)_i$ are some calibrated real constants that ensure model identifiability. Usually, we choose the $(c_i)_i$ from $\{-1, 0, 1\}$ with calibration. Next, for this computational study, we assume that the functional subset Θ belongs to the Hilbert subspace spanned by the finite basis functions of $(e_i)_{i=1,\cdots,k_0}$. This basis function constitutes the k_0-eigenfunction associated with the k_0 largest eigenvalue. For the sake of brevity, we have fixed $k_0 = 5$. Thereafter, we use the same cross-validation rules to select the smoothing parameter $h_{\mathbf{n}}$. Specifically, we have

$$h_{\mathbf{n}} = \arg\min \arg\min_{h \in H_{\mathbf{n}}} \sum_{\mathbf{k} \in \mathcal{I}_{\mathbf{n}}} \rho_q(Y_{\mathbf{k}} - \widehat{\xi}_q^{-\mathbf{k}}(X_{\mathbf{k}})),$$

where

$$H_{\mathbf{n}} = \left\{ a \geq 0 : \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \mathbb{1}_{B(3,a)}(X_{\mathbf{i}}) = k \right\},$$

$k \in \{5, 15, 25, \cdots, 0.5\widehat{\mathbf{n}}\}$, and the ball is defined with respect to the L_2-distance between the functional regressors. We simulate with quadratic kernel-defined as

$$K(t) = \frac{3}{2}(1 - t^2)\mathbb{1}_{[0,1)}.$$

The effectiveness of the estimator $\widehat{\xi}_q$ is evaluated by computing the mean square errors (MSE)

$$\texttt{MSE}(p) = \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{j}} \left(\xi_q(X_{\mathbf{j}}) - \widehat{\xi}_q(X_{\mathbf{j}}) \right)^2, \tag{5.1}$$

where the theoretical expectile regression ξ_q is obtained by using the routine code `qenorm` in the `R`-package `VGAM`.

The results are recorded in Table 1. It compares the SCVLSE-rule and SCVML-rule with an arbitrary selector method. The latter of these rules is obtained by dividing the optimal parameters of the SCVLSE-rule by 2. We emphasize our choice of utilizing the neighborhood set $\mathbf{W_k}$, which enables the integration of the spatial component into the functional component through the trace-variogram function. The sequence $\iota_{\mathbf{n}}$ in the vicinity set $\mathbf{W_k}$ is selected among the quantile of order q of the estimator vector trace-variogram $\gamma(\mathbf{k}, \mathbf{i})$ given by the code *trace.variog* in the *geofd* package. The following table contains the MSE(q) for $q = 0.01$, $q = 0.05$, $q = 0.1$, $\widehat{\mathbf{n}} = 200$ and $\psi = 0.5, 2$, which describes the covariance function of the random Weibull field in the regressor.

Table 1. *MSE*-results.

rule	value of ψ	q=0.01	q=0.05	q=0.1
SCVLSE-rule	0.5	0.34	0.37	0.31
	2	0.19	0.15	0.22
SCVML-rule	0.5	0.28	0.32	0.34
	2	0.11	0.09	0.08
Arbitrary -rule	0.5	0.77	0.73	0.84
	2	0.61	0.68	0.62

Unsurprisingly, the efficiency of $\widehat{\xi_q}$ is heavily affected by the spatial correlation as well as the selection rule to choose the parameter involved in the computation of $\widehat{\xi_q}$, such as the functional index θ and the smoothing parameters $h_{\mathbf{n}}$. It is clear that the arbitrary way significantly destroys the estimation quality. On the other hand, the effect of the spatial correlation of the data also impacts the estimation quality. Indeed, it is clear that the estimation quality decreases with large values of ψ. To better illustrate this observation, we plot in Figure 2 theMSE value with respect to the values of ψ.

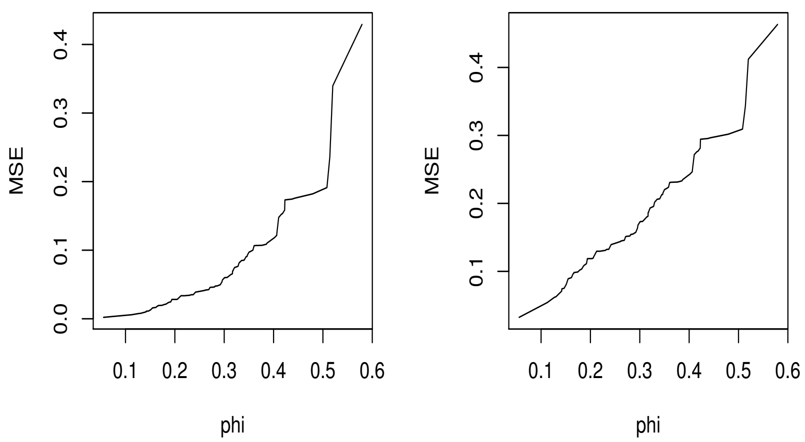

Figure 2. Left plot: MSE with respect to the SCVLSE-rule. Right plot: MSE with respect to the SCVML-rule.

In the second illustration, we compare our approach to its competitive such as the parametric and the nonparametric methods. To conduct a fair comparison between the three algorithms, we regenerate the response variable using three different situations. Indeed, in addition to the initial data of the first illustration, we define

$$\begin{aligned} Y_{\mathbf{i}} &= \int \theta(t)(X_{\mathbf{i}})(t)dt + \epsilon_{\mathbf{i}}, \text{ linear case,} \\ Y_{\mathbf{i}} &= r(X_{\mathbf{i}}) + \epsilon_{\mathbf{i}}, \text{purely nonparametric case,} \end{aligned}$$

where θ, r, $\epsilon_{\mathbf{i}}$, and $X_{\mathbf{i}}$ are defined in the first setting. We were inspired by [39] to define the spatial functional linear expectile regression as

$$\bar{\xi}_q(\mathcal{X}_{\mathbf{k}}) = \langle \ddot{\xi}_q, \mathcal{X}_{\mathbf{k}} \rangle \quad \text{with} \quad \ddot{\xi}_q = \sum_{j=1}^{M} \widehat{s_{j,q}} \widehat{\upsilon_j},$$

where the vector

$$\begin{pmatrix} \widehat{s_{1,q}} \\ \widehat{s_{2,q}} \\ \vdots \\ \widehat{s_{M,q}} \end{pmatrix} = \arg\min_{\zeta \in \mathbb{R}^M} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \rho_q \left(Y_{\mathbf{i}} - \sum_{j=1}^{m} \zeta_j \langle \widehat{\upsilon_j}, X_{\mathbf{i}} \rangle \right),$$

and where the $(\zeta_j)_j$ are the components of ζ in the basis function, and $(\widehat{\upsilon_j})_{j=1,\cdots,M}$, for the M eigenfunctions associated with the M greatest eigenvalues of the spatial empirical version of the covariance operator

$$\Gamma_{\mathbf{n}}(u) = \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \langle X_{\mathbf{i}}, u \rangle X_{\mathbf{i}}.$$

For the spatial nonparametric functional expectile regression we adopt the estimator of [27], which is defined by estimating G_1 and G_2 as

$$\widetilde{G_1}(\theta, \mathcal{X}_{\mathbf{k}}, t) = \frac{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} \|\mathcal{X}_{\mathbf{k}} - \mathcal{X}_{\mathbf{i}}\|\right) \mathbb{1}_{\mathbf{W}_{\mathbf{k}}}(\mathbf{i}) \mathbb{1}_{(Y_{\mathbf{i}}-t) \le 0}(Y_{\mathbf{i}} - t)}{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} \|\mathcal{X}_{\mathbf{k}} - \mathcal{X}_{\mathbf{i}}\|\right) \mathbb{1}_{\mathbf{W}_{\mathbf{k}}}(\mathbf{i})}$$

and

$$\widetilde{G_2}(\theta, \mathcal{X}_{\mathbf{k}}, t) = \frac{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} \|\mathcal{X}_{\mathbf{k}} - \mathcal{X}_{\mathbf{i}}\|\right) \mathbb{1}_{\mathbf{W}_{\mathbf{k}}}(\mathbf{i}) \mathbb{1}_{(Y_{\mathbf{i}}-t) > 0}(Y_{\mathbf{i}} - t)}{\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K\left(h_{\mathbf{n}}^{-1} \|\mathcal{X}_{\mathbf{k}} - \mathcal{X}_{\mathbf{i}}\|\right) \mathbb{1}_{\mathbf{W}_{\mathbf{k}}}(\mathbf{i})}.$$

Using the same procedure as in the first setting, we select the sequence of the vicinity set $(\iota_{\mathbf{n}})$, the smoothing parameter, the same metric, and the same kernel. We examine the performance of the three approaches $\bar{\xi}_q$, $\widetilde{\xi}_q$ and $\widehat{\xi}_q$ using the MSE of (5.1) presented in Figures 3–5. Unsurprisingly, the semi-parametric estimator $\widehat{\xi}_q(\mathcal{X}_{\mathbf{k}})$ is more stable for the three situations. In the sense that its MSE has slow variability with respect the different situations. It is of order 0.33 in the linear case, 0.36 in the semi-parametric case and 0.38 of the purely nonparametric case. The MSEs of $\bar{\xi}_q$ are (0.27, 0.45, 0.88) versus (0.97, 0.67, 0.31) for $\widetilde{\xi}_q$. We observe also that the estimators $\bar{\xi}_q$ and $\widetilde{\xi}_q$ outperform only for the parametric and nonparametric situations, respectively.

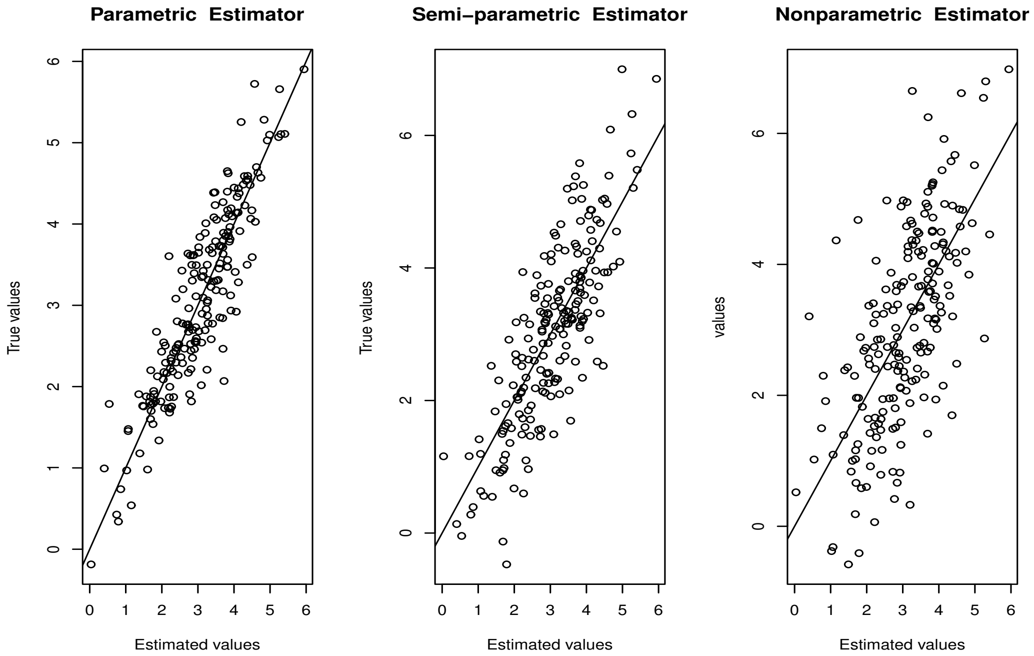

Figure 3. Case 1: The data are generated according to the semiparametric model.

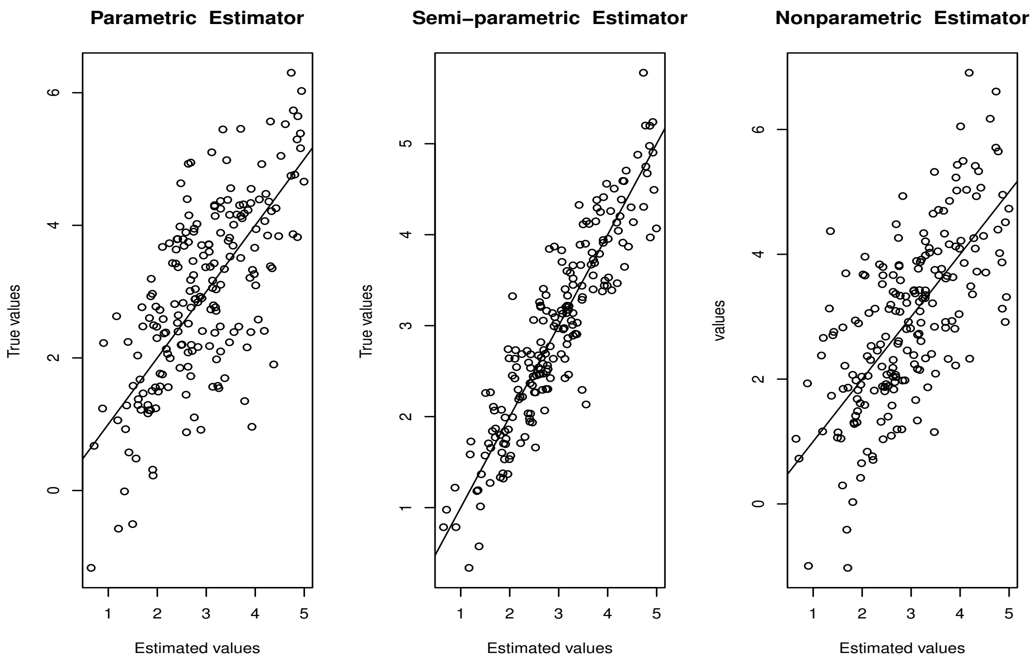

Figure 4. Case 2: The data are generated according to the semiparametric model.

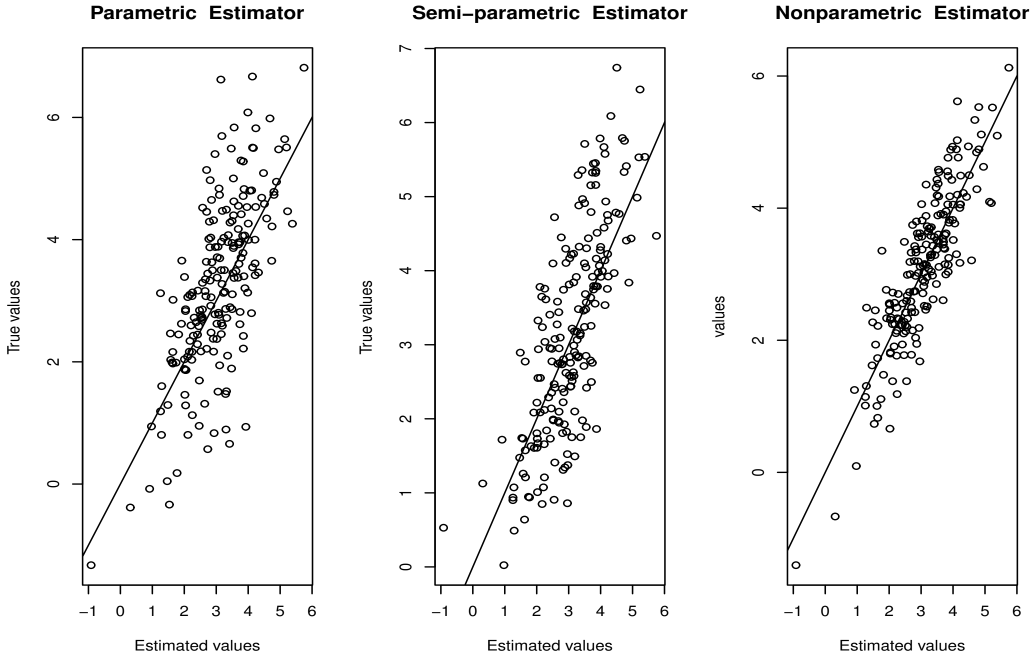

Figure 5. Case 3: The data are generated according to the nonparametric model.

5.2. Real data application

Commonly, the financial area is the natural field of expectile regression. In this area, expectile regression is used as an alternative model of risk to the expected-shortfall and the value-at-risk (VaR). It is more informative than the mentioned risk tools, such as the VaR function. This statement results from the fact that the expectile model is the tail expectation, whereas the VaR function is the quantile model of the tail probabilities. The tail expectation function covers the frequencies as well as the values, whereas the tail probability is based only on the frequency. The novelty of the spatial expectile is the possibility to fit the financial risk of the co-movements of various investments in different sectors or stock markets. In a sense, the spatio-functional correlation of the financial data regroups the time conventional correlation as well as the pairwise relationships between the stock markets through known financial metrics. It is worth noting the fact that the spatial linkage in financial data is unrelated to the geographic localization of the stock markets. Such spatial financial distances are usually deafened by some spatial matrix weighting. Thus, we seek in this computational part to inspect the behavior of the spatio-functional conditional expectile concerning some common spatial matrix weighting. For this aim, we consider the Euro Stoxx-50 index data. Such spatio-functional observations are available at `https://fred.stlouisfed.org/series` (accessed on 14 March 2023). We proceed with the difference logarithmic of this data for the closed prices $r(\cdot)$ of the period between 22 February 2022 to 23 February 2023. Specifically, we build a functional variable from

$$Z(t) = -100 \log\left(\frac{r(t)}{r(t-1)}\right)$$

as a continuous process. The real interest variable Y is Z (of the-last-day of month), and the functional insert variable $X(\cdot)$ illustrates the values of Z (for one month). In Figure 6, we plot the observed functional regressors.

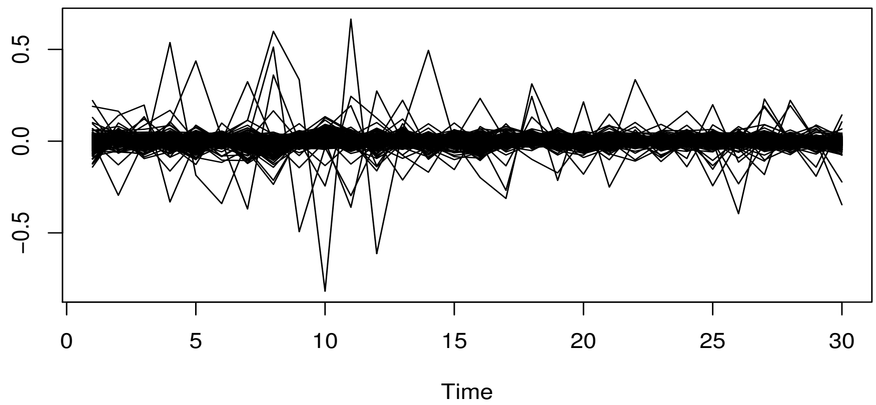

Figure 6. The spatio-functional regressor.

Of course, the first step of spatial modeling is spatial detrending, which is necessary to use the stationarity assumption. To do that, we generate a stationary process $(\widetilde{Y}_{\mathbf{l}}, \widetilde{X}_{\mathbf{l}})_{\mathbf{l}}$ from the initial spatial observation $(Y_{\mathbf{l}}, X_{\mathbf{i}})_{\mathbf{l}}$ as

$$\begin{cases} \tilde{X}_{\mathbf{l}} = X_{\mathbf{l}} - m_1(\mathbf{l}), \\ \widetilde{Y}_{\mathbf{l}} = Y_{\mathbf{l}} - m_2(\mathbf{l}). \end{cases}$$

Next, we compute the conditional expectile estimator $\widehat{\xi_q}$ from the statistics $(\widehat{X}_{\mathbf{i}}, \widehat{Y}_{\mathbf{i}})_{\mathbf{i}}$ instead of from the initial observations $(X_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}}$. Thus, the used observations are obtained by estimating the real functions $m_1(\cdot)$ and $m_2(\cdot)$ by

$$\widehat{m}_1(\mathbf{i}_0) = \sum_{\mathbf{i} \in \mathbf{I_n}} W_{\mathbf{i_0 i}} X_{\mathbf{i}}$$

and

$$\widehat{m}_2(\mathbf{j}_0) = \sum_{\mathbf{j} \in \mathbf{I_n}} W'_{\mathbf{j_0 j}} Y_{\mathbf{j}},$$

where $W_{\mathbf{ij}}$ and $W'_{\mathbf{ij}}$ are given spatial weighting matrices. As mentioned above, we evaluate the impact of this step in the spatio-financial risk by comparing three common weighting matrices:

- The first one is

$$W^1_{\mathbf{ij}} = \begin{cases} 1, & \text{if } \mathbf{i} \text{ and } \mathbf{j} \text{ are in the same sector,} \\ 0, & \text{if not.} \end{cases}$$

- The second one

$$W^2_{\mathbf{ij}} = \begin{cases} 1, & \text{if } \mathbf{i} \text{ and } \mathbf{j} \text{ are in the same sector and the same country,} \\ 0.5, & \text{if } \mathbf{i} \text{ and } \mathbf{j} \text{ are in the same sector,} \\ 0, & \text{if not.} \end{cases}$$

- The third one

$$W^3_{\mathbf{ij}} = \begin{cases} 1, & \text{if } \mathbf{i} = \mathbf{j}, \\ 0, & \text{if not.} \end{cases}$$

The last matrix allows us to examine also the behavior of the spatio-functional expetile regression without the detrending part. Now, to run our spatio-functional model, we retrain the same schemes as those exercised in the artificial example to designate the parameters of $\widehat{\xi_q}$. Specifically, we use the quadratic function supported on $(0, 1)$ as the kernel and we select the single index and the smoothing parameter by the rule SCVLSE as

$$(\theta_o pt, h_{\mathbf{n}}) = \arg\min \arg \min_{h\in H_{\mathbf{n}},\theta in\Theta_{\mathbf{n}}} \sum_{\mathbf{k}\in\mathcal{I}_{\mathbf{n}}} \rho_q(Y_{\mathbf{k}} - \widehat{\xi_q}^{-\mathbf{k}}(X_{\mathbf{k}})),$$

where Θ_n and $H_{\mathbf{n}}$ are defined in the same manner as in the previous section. However, we use the metric of principal component analysis to specify the ball in $H_{\mathbf{n}}$. For both subsets Θ_n and $H_{\mathbf{n}}$, the principal component analysis is performed over 5-eigenfunctions associated to the 5 largest eigenvalues. The feasibility of the proposed spatial risk analysis is checked for $q = 0.95$ and $q = 0.05$ by dividing the data several times (exactly 55 times). The observations are divided at random into two parts: the training sample (220 observations) and the testing sample (150 observations). Finally, we assess the feasibility of the detrending steps by measuring

$$Eror = \frac{1}{150}\sum_{\mathbf{i}=1}^{150} \rho_{0.01}(Y_{\mathbf{i}} - \widehat{\xi}_{0.01}(X_{\mathbf{i}})),$$

where

$$\rho_q(s) = |q - \mathbb{1}_{\{s<0\}}|s^2.$$

The values $Eror$ for the 55 random splitting operations of the sample are plotted in Figure 7.

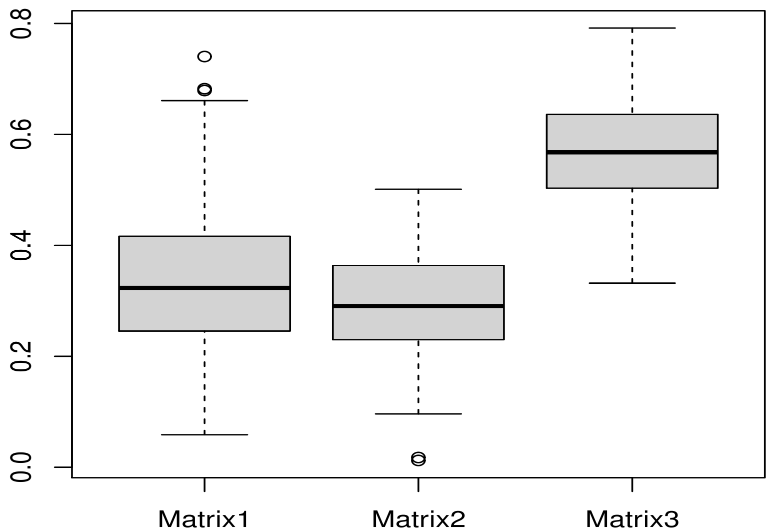

Figure 7. Comparison of the *Eror* values between the three matrices.

It is obvious that there is a remarkable difference between the stationary case and the nonstationary situation. In a sense, when we execute with the initial data without detrending (matrix 3), the estimation quality is poor compared to the detrending step (using matrices 1 and 2). In particular, the detrending step permits us to increase the effectiveness of the estimator $\widehat{\xi_q}$ by reducing the risk measure $Eror(\cdot)$. The scatter plot in Figure 7 demonstrates that the median of $Eror$-values remarkably varies between the three matrices. It is around 0.3 for the matrix $W^1_{\mathbf{ij}}$, around 0.2 for $W^2_{\mathbf{ij}}$ and 0.6 in the nonstationary case associated with the matrix $W^3_{\mathbf{ij}}$.

6. Conclusions

In this paper, we have demonstrated the BCC consistency of spatio-functional expectile regression under the FSIR structure. Such a result is established over some general postulates, allowing us to explore the nonparametric nature of the expectile operator, the functional path of the financial time series data, and the spatial correlation of the observed data. On the other hand, since the degree of correlation greatly impacts the convergence speed of the estimator, we have modeled this feature using various rules. First, we have evaluated the effect of spatial dependency on the choice of the single index model. For this purpose, we have employed vicinity-set techniques and the spatial weighting matrix. We observed that the two approaches fit the spatial dependency correctly and are a good tool for controlling the spatial covariation of the financial data. Indeed, since financial transactions are performed via the internet, the spatial dependency between them is not based only on the location of the financial institutions. Thus, the vicinity-set techniques (defined by the trace-variogram function) and the spatial weighting matrix algorithm allows for the integration of all the different elements affecting the spatial correlation. Moreover, we show that the insertion of the expectile operator in financial risk management is carried out via two principal steps: detrending and determination of the estimator. Such a strategy increases the efficiency of our algorithm in practice. Additionally, to this theoretical and practical development, the present contribution opens very interesting tracks for future research. For example, it will be a priority in the future to study the asymptotic property of the parametric estimation of the spatio-functional expectile operator. Such a prospect is motivated by the expectile operator's ability to behave as linear, nonparametric, or semiparametric forms [62–64]. Thus, the ideas of [30] can be extended here. Second, the asymptotic normality of the estimators is also crucial in mathematical statistics. It allows us to determine the confidence interval with a given confidence level. Extending our results to other functional time series cases (ergodic, long memory, associated process) would be interesting. However, it would require nontrivial mathematics that is well beyond the scope of this paper.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors would like to thank the associate-editor and the four anonymous reviewer for their valuable comments and suggestions which improved substantially the quality of an earlier version of this paper. The authors thank and extend their appreciation to the funders of this project: 1) Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R358), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia; 2) The Deanship of Scientific Research at King Khalid University through the Research Groups Program under grant number R.G.P. 1/366/44.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. S. Bouzebda, I. Soukarieh, Non-parametric conditional U-processes for locally stationary functional random fields under stochastic sampling design, *Mathematics*, **11** (2023), 16. https://doi.org/10.3390/math11010016
2. N. A. Cressie, *Statistics for spatial data*, John Wiley & Sons, Inc., 2015. https://doi.org/10.1002/9781119115151
3. X. Guyon, *Random fields on a network: modeling, statistics, and applications*, Springer-Verlag, 1995.
4. B. D. Ripley, Spatial statistics: developments, 1980–1983, *Int. Stat. Rev.*, **52** (1984), 141–150. https://doi.org/10.2307/1403097
5. M. Rosenblatt, *Stationary sequences and random fields*, Springer-Verlag, 1985. https://doi.org/10.1007/978-1-4612-5156-9
6. S. Bouzebda, A. Laksaci, M. Mohammedi, Single index regression model for functional quasi-associated time series data, *REVSTAT*, **20** (2022), 605–631. https://doi.org/10.57805/revstat.v20i5.391
7. S. Bouzebda, A. Laksaci, M. Mohammedi, The k-nearest neighbors method in single index regression model for functional quasi-associated time series data, *Rev. Mat. Complutense*, **36** (2023), 361–391. https://doi.org/10.1007/s13163-022-00436-z
8. W. Härdle, P. Hall, H. Ichimura, Optimal smoothing in single-index models, *Ann. Stat.*, **21** (1993), 157–178. https://doi.org/10.1214/aos/1176349020
9. M. Hristache, A. Juditsky, V. Spokoiny, Direct estimation of the index coefficient in a single-index model, *Ann. Stat.*, **29** (2001), 595–623. https://doi.org/10.1214/aos/1009210682
10. F. Ferraty, A. Peuch, P. Vieu, Modèle à indice fonctionnel simple, *C. R. Math.*, **336** (2003), 1025–1028. https://doi.org/10.1016/S1631-073X(03)00239-5
11. D. Chen, P. Hall, H. G. Müller, Single and multiple index functional regression models with nonparametric link, *Ann. Stat.*, **39** (2011), 1720–1747. https://doi.org/10.1214/11-AOS882
12. H. Ding, Y. Liu, W. Xu, R. Zhang, A class of functional partially linear single-index models, *J. Multivar. Anal.*, **161** (2017), 68–82. https://doi.org/10.1016/j.jmva.2017.07.004

13. M. Mohammedi, S. Bouzebda, A. Laksaci, O. Bouanani, Asymptotic normality of the k-NN single index regression estimator for functional weak dependence data, *Commun. Stat.*, 2022. https://doi.org/10.1080/03610926.2022.2150823

14. W. K. Newey, J. L. Powell, Asymmetric least squares estimation and testing, *Econometrica*, **55** (1987), 819–847. https://doi.org/10.2307/1911031

15. Z. Lu, X. Chen, Spatial kernel regression estimation: weak consistency, *Stat. Probab. Lett.*, **68** (2004), 125–136. https://doi.org/10.1016/j.spl.2003.08.014

16. L. T. Tran, Kernel density estimation on random fields, *J. Multivar. Anal.*, **34** (1990), 37–53. https://doi.org/10.1016/0047-259X(90)90059-Q

17. F. Bellini, V. Bignozzi, G. Puccetti, Conditional expectiles, time consistency and mixture convexity properties, *Insurance*, **82** (2018), 117–123. https://doi.org/10.1016/j.insmatheco.2018.07.001

18. Y. Gu, H. Zou, High-dimensional generalizations of asymmetric least squares regression and their applications, *Ann. Stat.*, **44** (2016), 2661–2694. https://doi.org/10.1214/15-AOS1431

19. I. M. Almanjahie, S. Bouzebda, Z. Kaid, A. Laksaci, Nonparametric estimation of expectile regression in functional dependent data, *J. Nonparametr. Stat.*, **34** (2022), 250–281. https://doi.org/10.1080/10485252.2022.2027412

20. M. Mohammedi, S. Bouzebda, A. Laksaci, The consistency and asymptotic normality of the kernel type expectile regression estimator for functional data, *J. Multivar. Anal.*, **181** (2021), 104673. https://doi.org/10.1016/j.jmva.2020.104673

21. T. Kneib, Beyond mean regression, *Stat. Modell.*, **13** (2013), 275–303. https://doi.org/10.1177/1471082X13494159

22. P. H. Eilers, Discussion: the beauty of expectiles, *Stat. Modell.*, **13** (2013), 317–322. https://doi.org/10.1177/1471082X13494313

23. R. Koenker, Discussion: living beyond our means, *Stat. Modell.*, **13** (2013), 323–333. https://doi.org/10.1177/1471082X13494314

24. M. C. Jones, Expectiles and M-quantiles are quantiles, *Stat. Probab. Lett.*, **20** (1994), 149–153. https://doi.org/10.1016/0167-7152(94)90031-0

25. I. M. Almanjahie, S. Bouzebda, Z. C. Elmezouar, A. Laksaci, The functional *k*NN estimator of the conditional expectile: uniform consistency in number of neighbors, *Stat. Risk Modell.*, **38** (2022), 47–63. https://doi.org/10.1515/strm-2019-0029

26. F. Alshahrani, I. M. Almanjahie, Z. C. Elmezouar, Z. Kaid, A. Laksaci, M. Rachdi, Functional ergodic time series analysis using expectile regression, *Mathematics*, **10** (2022), 3919. https://doi.org/10.3390/math10203919

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

28. G. Biau, B. Cadre, Nonparametric spatial prediction, *Stat. Infer. Stochastic Process.*, **7** (2004), 327–349. https://doi.org/10.1023/B:SISP.0000049116.23705.88

29. M. Hallin, Z. Lu, L. T. Tran, Local linear spatial regression, *Ann. Stat.*, **32** (2004), 2469–2500. https://doi.org/10.1214/009053604000000850

30. J. Li, L. T. Tran, Nonparametric estimation of conditional expectation, *J. Stat. Plann. Infer.*, **139** (2009), 164–175. https://doi.org/10.1016/j.jspi.2008.04.023

31. R. Xu, J. Wang, L_1-estimation for spatial nonparametric regression, *J. Nonparametr. Stat.*, **20** (2008), 523–537. https://doi.org/10.1080/10485250801976717

32. S. D. Niang, M. Rachdi, A. F. Yao, Kernel regression estimation for spatial functional random variables, *Far East J. Theor. Stat.*, **37** (2011), 77–113.

33. S. Koner, A. M. Staicu, Second-generation functional data, *Annu. Rev. Stat. Appl.*, **10** (2023), 547–572. https://doi.org/10.1146/annurev-statistics-032921-033726

34. J. O. Ramsay, T. Ramsay, L. M. Sangalli, *Spatial functional data analysis*, Springer-Verlag, 2011. https://doi.org/10.1007/978-3-7908-2736-1_42

35. M. Lv, J. E. Fowler, L. Jing, Spatial functional data analysis for the spatial–spectral classification of hyperspectral imagery, *IEEE Geosci. Remote Sens. Lett.*, **16** (2019), 942–946. https://doi.org/10.1109/LGRS.2018.2884077

36. J. Mateu, E. Romano, Advances in spatial functional statistics, *Stochastic Environ. Res. Risk Assess.*, **31** (2017), 1–6. https://doi.org/10.1007/s00477-016-1346-z

37. S. D. Niang, A. F. Yao, Kernel spatial density estimation in infinite dimension space, *Metrika*, **76** (2013), 19–52. https://doi.org/10.1007/s00184-011-0374-4

38. A. Chouaf, A. Laksaci, On the functional local linear estimate for spatial regression, *Stat. Risk Modell.*, **29** (2012), 189–214. https://doi.org/10.1524/strm.2012.1114

39. M. Rachdi, A. Laksaci, F. A. A. Awadhi, Parametric and nonparametric conditional quantile regression modeling for dependent spatial functional data, *Spat. Stat.*, **43** (2021), 100498. https://doi.org/10.1016/j.spasta.2021.100498

40. G. Aneiros, S. Novo, P. Vieu, Variable selection in functional regression models: a review, *J. Multivar. Anal.*, **188** (2022), 104871. https://doi.org/10.1016/j.jmva.2021.104871

41. S. Bouzebda, B. Nemouchi. Central limit theorems for conditional empirical and conditional U-processes of stationary mixing sequences, *Math. Methods Stat.*, **28** (2019), 169–207. https://doi.org/10.3103/S1066530719030013

42. S. Bouzebda, M. Chaouch, Uniform limit theorems for a class of conditional Z-estimators when covariates are functions, *J. Multivar. Anal.*, **189** (2022), 104872. https://doi.org/10.1016/j.jmva.2021.104872

43. S. Bouzebda, B. Nemouchi, Weak-convergence of empirical conditional processes and conditional U-processes involving functional mixing data, *Stat. Infer. Stochastic Process.*, **26** (2023), 33–88. https://doi.org/10.1007/s11203-022-09276-6

44. J. Hristov, Special issue: trends in fractional modelling in science and innovative technologies, *Symmetry*, **15** (2023), 884. https://doi.org/10.3390/sym15040884

45. H. G. Müller, Special issue on "functional and object data analysis": guest editor's introduction, *Canad. J. Stat.*, **50** (2022), 8–19. https://doi.org/10.1002/cjs.11690

46. M. Carbon, M. Hallin, L. T. Tran, Kernel density estimation for random fields: the L_1 theory, *J. Nonparametr. Stat.*, **6** (1996), 157–170. https://doi.org/10.1080/10485259608832669

47. P. Doukhan, *Mixing*, Springer-Verlag, 1994. https://doi.org/10.1007/978-1-4612-2642-0
48. D. Tjøstheim, Statistical spatial series modelling, *Adv. Appl. Probab.*, **10** (1978), 130–154. https://doi.org/10.2307/1426722
49. X. Guyon, Estimation d'un champ par pseudo-vraisemblance conditionnelle: étude asymptotique et application au cas markovien, *Proceedings of the Sixth Franco-Belgian Meeting of Statisticians*, 1987.
50. R. C. Bradley, Some examples of mixing random fields, *Rocky Mountain J. Math.*, **23** (1993), 495–519. https://doi.org/10.1216/rmjm/1181072573
51. J. Dedecker, P. Doukhan, G. Lang, L. R. J. Rafael, S. Louhichi, C. Prieur, *Weak dependence: with examples and applications*, Springer-Verlag, 2007. https://doi.org/10.1007/978-0-387-69952-3
52. D. Kurisu, Nonparametric regression for locally stationary random fields under stochastic sampling design, *Bernoulli*, **28** (2022), 1250–1275. https://doi.org/10.3150/21-bej1385
53. I. Soukarieh, S. Bouzebda, Weak convergence of the conditional U-statistics for locally stationary functional time series, *Stat. Infer. Stochastic Process.*, 2023. https://doi.org/10.1007/s11203-023-09305-y
54. V. I. Bogachev, *Gaussian measures*, American Mathematical Society, 1998.
55. W. V. Li, Q. M. Shao, Gaussian processes: inequalities, small ball probabilities and applications, *Handb. Stat.*, **19** (2001), 533–597. https://doi.org/10.1016/S0169-7161(01)19019-X
56. F. Ferraty, P. Vieu, *Nonparametric functional data analysis*, Springer-Verlag, 2006. https://doi.org/10.1007/0-387-36620-2
57. N. A. Cressie, *Spatial prediction in a multivariate setting*, Elsevier, 1993.
58. J. Mateu, R. Giraldo, *Geostatistical functional data analysis*, John Wiley & Sons, Ltd., 2021. https://doi.org/10.1002/9781119387916
59. A. Ait-Saïdi, F. Ferraty, R. Kassa, P. Vieu, Cross-validated estimations in the single-functional index model, *Statistics*, **42** (2008), 475–494. https://doi.org/10.1080/02331880801980377
60. A. Toma, C. Fulga, Robust estimation for the single index model using pseudodistances, *Entropy*, **20** (2018), 374. https://doi.org/10.3390/e20050374
61. M. Bonneu, X. Milhaud, A modified Akaike criterion for model choice in generalized linear models, *Statistics*, **25** (1994), 225–238. https://doi.org/10.1080/02331889408802447
62. S. Bouzebda, M. Cherfi, General bootstrap for dual ϕ-divergence estimates, *J. Probab. Stat.*, **2012** (2012), 834107. https://doi.org/10.1155/2012/834107
63. S. Bouzebda, A. Keziou. A new test procedure of independence in copula models via χ^2-divergence, *Commun. Stat.*, **39** (2009), 1–20. https://doi.org/10.1080/03610920802645379
64. S. Bouzebda, A. Keziou, New estimates and tests of independence in semiparametric copula models, *Kybernetika*, **46** (2010), 178–201.

Appendix

This section is devoted to the proof of our main result. The previously presented notation continues to be used in the following.

Proof of Theorem 1. The function $\widehat{G}$ is non-decreasing, and therefore its derivative is non-negative at $\xi_q(\theta, \mathfrak{z})$. So, for any $\epsilon > 0$, we have

$$\begin{aligned}\sum_{\mathbf{n}} Pr\left(\left|\widehat{\xi_q}(\theta, \mathfrak{z}) - \xi_q(\theta, \mathfrak{z})\right| > \epsilon\right) \\ \leq \sum_{\mathbf{n}} Pr\left(\left|\widehat{G}(\theta, \mathfrak{z}, \xi_q(\theta, \mathfrak{z}) - \epsilon) - G(\theta, \mathfrak{z}, \xi_q(\theta, \mathfrak{z}) - \epsilon)\right| \geq C\epsilon\right) \\ + \sum_{\mathbf{n}} Pr\left(\left|\widehat{G}(\theta, \mathfrak{z}, \xi_q(\theta, \mathfrak{z}) + \epsilon) - G(\theta, \mathfrak{z}, \xi_q(\theta, \mathfrak{z}) + \epsilon)\right| \geq C\epsilon\right).\end{aligned}$$

Next, it suffices to use the Proposition 3.1, and for

$$t = \xi_q(\theta, \mathfrak{z}) \pm \epsilon,$$

we obtain

$$\sum_{\mathbf{n}} Pr\left(\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|\widehat{G}(\theta, \mathfrak{z}, t) - G(\theta, \mathfrak{z}, t)\right| > C\epsilon\right) < \infty. \tag{A.1}$$

So, in order to show the consistency of $\widehat{\xi_q}(\theta, \mathfrak{z})$, we have to prove the uniform BCC of $\widehat{G}_{1,2}(\theta, \mathfrak{z}, t)$, as $\mathbf{n} \to \infty$. This requirement is a consequence of Proposition 3.1. Hence, it suffices to prove this proposition first. □

Proof of Proposition 3.1. For this proposition, we decompose $\widehat{G}(\theta, \mathfrak{z}, t) - G(\theta, \mathfrak{z}, t)$ as follows:

$$\begin{aligned}\widehat{G}(\theta, \mathfrak{z}, t) - G(\theta, \mathfrak{z}, t) &= \frac{\widehat{G}_1(\theta, \mathfrak{z}, t)}{\widehat{G}_2(\theta, \mathfrak{z}, t)} - \frac{G_1(\theta, \mathfrak{z}, t)}{G_2(\theta, \mathfrak{z}, t)} \\ &= \frac{1}{\widehat{G}_2(\theta, \mathfrak{z}, t)}\left[\widehat{G}_1(\theta, \mathfrak{z}, t) - G_1(\theta, \mathfrak{z}, t)\right] \\ &\quad + \frac{G(\theta, \mathfrak{z}, t)}{\widehat{G}_2(\theta, \mathfrak{z}, t)}\left[G_2(\theta, \mathfrak{z}, t) - \widehat{G}_2(\theta, \mathfrak{z}, t)\right].\end{aligned} \tag{A.2}$$

Thus, the proposition, as well as the theorem, are consequences of the following intermediate results.

Lemma 1. *Under the conditions of Proposition 3.1, we have*

$$\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|\widehat{G}_1(\theta, \mathfrak{z}, t) - \mathbb{E}\left[\widehat{G}_1(\theta, \mathfrak{z}, t)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta, \mathfrak{z}, h_{\mathbf{n}})}}\right) \quad as \quad \mathbf{n} \to \infty \tag{A.3}$$

and

$$\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|\widehat{G}_2(\theta, \mathfrak{z}, t) - \mathbb{E}\left[\widehat{G}_2(\theta, \mathfrak{z}, t)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta, \mathfrak{z}, h_{\mathbf{n}})}}\right) \quad as \quad \mathbf{n} \to \infty.$$

□

Proof of Lemma 1. Since the proof of the two terms is the same, we focus on the first term only. To evaluate the maximum of this dispersion term on a compact interval, we recover $[\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]$, by finite compact intervals $\left[t_j-\ell_{\mathbf{n}}, t_j+\ell_{\mathbf{n}}\right]$, with

$$\ell_{\mathbf{n}} = \widehat{\mathbf{n}}^{-1/2\varsigma_l}$$

and

$$d_{\mathbf{n}} = O\left(\widehat{\mathbf{n}}^{1/2\varsigma_1}\right).$$

Let

$$\mathcal{G}_{\mathbf{n}} = \left\{t_j-\ell_{\mathbf{n}}, t_j+\ell_{\mathbf{n}}, 1\leq j\leq d_{\mathbf{n}}\right\}, \tag{A.4}$$

the subset of the covering interval's extremities. Now, from the monotonicity of $\widehat{G}_1(\theta,\mathfrak{z},\cdot)$ and $\mathbb{E}[\widehat{G}_1(\theta,\mathfrak{z},\cdot)]$, we write, for $1\leq j\leq d_{\mathbf{n}}$,

$$\begin{aligned}
\widehat{G}_1(\theta,\mathfrak{z},t_j-\ell_{\mathbf{n}}) &\leq \sup_{t\in(t_j-\ell_{\mathbf{n}},t_j+\ell_{\mathbf{n}})} \widehat{G}_1(\theta,\mathfrak{z},t) \leq \widehat{G}_1(\theta,\mathfrak{z},t_j+\ell_{\mathbf{n}}),\\
\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t_j-\ell_{\mathbf{n}})\right] &\leq \sup_{t\in(t_j-\ell_{\mathbf{n}},t_j+\ell_{\mathbf{n}})} \mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right] \leq \mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t_j+\ell_{\mathbf{n}})\right].
\end{aligned}$$

It follows that

$$\begin{aligned}
&\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]-\widehat{G}_1(\theta,\mathfrak{z},t)\right|\\
&\quad\leq \max_{1\leq j\leq d_{\mathbf{n}}}\max_{z\in\{t_j-\ell_{\mathbf{n}},t_j+\ell_{\mathbf{n}}\}} \left|\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]-\widehat{G}_1(\theta,\mathfrak{z},z)\right| + 2^{\varsigma_1}C_2\ell_{\mathbf{n}}^{\varsigma_1}.
\end{aligned}$$

Since

$$\ell_{\mathbf{n}}^{\varsigma_1} = o\left(\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right)^{1/2}\right),$$

we prove only the fact that

$$\max_{1\leq j\leq d_{\mathbf{n}}}\max_{z\in\{t_j-\ell_{\mathbf{n}},t_j+\ell_{\mathbf{n}}\}} \left|\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]-\widehat{G}_1(\theta,\mathfrak{z},z)\right| = O\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right), \quad a.co. \tag{A.5}$$

To do that we write

$$\begin{aligned}
&Pr\left(\max_{1\leq j\leq d_{\mathbf{n}}}\max_{z\in\{t_j-\ell_{\mathbf{n}},t_j+\ell_{\mathbf{n}}\}} \left|\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]-\widehat{G}_1(\theta,\mathfrak{z},z)\right| > \eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right)\\
&\quad\leq 2d_{\mathbf{n}}\max_{1\leq j\leq d_{\mathbf{n}}}\max_{z\in\{t_j-\ell_{\mathbf{n}},t_j+\ell_{\mathbf{n}}\}} Pr\left(\left|\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]-\widehat{G}_1(\theta,\mathfrak{z},z)\right| > \eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right).
\end{aligned}$$

So, all that remains is to evaluate for all $z = t_j\mp\ell_{\mathbf{n}}$, $1\leq j\leq d_{\mathbf{n}}$,

$$Pr\left(\left|\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]-\widehat{G}_1(\theta,\mathfrak{z},z)\right| > \eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right). \tag{A.6}$$

To simplify the notation, we let

$$Y_{\mathbf{i}}^{-} = \mathbb{1}_{(Y_{\mathbf{i}}-t)\leq 0}(Y_{\mathbf{i}} - t),$$

which is not necessarily bounded. Thus, to evaluate (A.6), we use the truncation method. Indeed, we consider

$$\widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t) = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}(\theta, \mathfrak{z}) Y_{\mathbf{i}}^{*},$$

where

$$Y_{\mathbf{i}}^{*} = Y_{\mathbf{i}}^{-} \mathbb{1}_{(|Y_{\mathbf{i}}|<\gamma_{\mathbf{n}})}$$

with

$$K_{\mathbf{i}}(\theta, \mathfrak{z}) = K(h_{\mathbf{n}}^{-1} N_{\theta}(\mathfrak{z} - \mathcal{X}_{\mathbf{i}})).$$

So, (A.3) is a consequence of the following results:

$$\left|\mathbb{E}[\widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t)] - \mathbb{E}[\widehat{G}_{1}(\theta, \mathfrak{z}, t)]\right| = O\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta, \mathfrak{z}, h_{\mathbf{n}})}}\right), \tag{A.7}$$

$$\left|\widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t) - \widehat{G}_{1}(\theta, \mathfrak{z}, t)\right| = O_{a.co.}\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta, \mathfrak{z}, h_{\mathbf{n}})}}\right) \tag{A.8}$$

and

$$\left|\widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t) - \mathbb{E}[\widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t)]\right| = O_{a.co.}\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta, \mathfrak{z}, h_{\mathbf{n}})}}\right). \tag{A.9}$$

For statement (A.7): For this equation, we use Holder's inequality (for $\alpha = \frac{p}{2}$ and β such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$), which allows us to write that

$$\begin{aligned}
\left|\mathbb{E}[|\widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t)] - \mathbb{E}[|\widehat{G}_{1}(\theta, \mathfrak{z}, t)]\right| &\leq C\frac{1}{\mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]}\mathbb{E}\left[\left|Y^{-}\right| \mathbb{1}_{\{Y\geq\gamma_{\mathbf{n}}\}} K_{\mathbf{1}}(\theta, \mathfrak{z})\right] \\
&\leq \frac{\gamma_{\mathbf{n}}^{-\alpha}}{\mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]}\mathbb{E}^{1/\alpha}\left[\left|Y^{2\alpha}\right|\right]\mathbb{E}^{1/\beta}\left[K_{\mathbf{1}}^{\beta}(\theta, \mathfrak{z})\right] \\
&\leq \frac{\gamma_{\mathbf{n}}^{-\alpha}}{\mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]}\mathbb{E}^{1/\alpha}\left[|Y^{p}|\right]\mathbb{E}^{1/\beta}\left[K_{\mathbf{1}}^{\beta}(\theta, \mathfrak{z})\right].
\end{aligned}$$

It follows that

$$\left|\mathbb{E}[|\widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t)] - \mathbb{E}[|\widehat{G}_{1}(\theta, \mathfrak{z}, t)]\right| \leq C\gamma_{\mathbf{n}}^{-\alpha}\phi_{x}^{(1-\beta)/\beta}(\theta, \mathfrak{z}, h_{\mathbf{n}}).$$

Finally, (H5) allows us to conclude the statement (A.7).

For statement (A.8): The demonstration of this statement is based on the Markov inequality. For all $\epsilon > 0$, we have

$$\begin{aligned}
&Pr\left(\left|\widehat{G}_{1}(\theta, \mathfrak{z}, t) - \widehat{G^{*}}_{1}(\theta, \mathfrak{z}, t)\right| > \epsilon\right) \\
&\quad\leq \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} Pr\left(Y_{\mathbf{i}} > \gamma_{\mathbf{n}}\right) \leq \widehat{\mathbf{n}} Pr\left(Y > \gamma_{\mathbf{n}}\right) \leq \widehat{\mathbf{n}}\gamma_{\mathbf{n}}^{-p}\mathbb{E}\left[Y^{p}\right].
\end{aligned}$$

Choose

$$\epsilon = \epsilon_0\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right)$$

to deduce that

$$\sum_{\mathbf{n}} Pr\left(\left|\widehat{G}_1(\theta,\mathfrak{z},t)-\widehat{G^*}_1(\theta,\mathfrak{z},t)\right| > \epsilon_0\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right)\right) \leq \sum_{\mathbf{n}} \widehat{\mathbf{n}}\gamma_{\mathbf{n}}^{-p} < \infty.$$

For statement (A.9): This is based on the blocks of spatial decomposition insights [30]. Specifically, we write

$$\Lambda_{\mathbf{i}} = K_{\mathbf{i}}(\theta,\mathfrak{z})Y_{\mathbf{i}}^* - \mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})Y_{\mathbf{i}}^*\right],$$

which allows us to write

$$\widehat{G^*}_1(\theta,\mathfrak{z},t) - \mathbb{E}[\widehat{G^*}_1(\theta,\mathfrak{z},t)] = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right]}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}).$$

The latter can be decomposed as

$$\widehat{G^*}_1(\theta,\mathfrak{z},t) - \mathbb{E}[\widehat{G^*}_1(\theta,\mathfrak{z},t)] = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right]}\sum_{i=1}^{2^N}\mathfrak{T}(\mathbf{n},i), \tag{A.10}$$

with, for all $i \in (1,N)$ and $r_i = 2n_i p_{\mathbf{n}}^{-1}$,

$$\mathfrak{T}(\mathbf{n},i) = \sum_{\mathbf{l}\in\mathcal{J}}\mathfrak{M}(i,\mathbf{n},\mathbf{j}),$$

where

$$\mathcal{J} = \{0,\cdots,r_1-1\}\times\cdots\times\{0,\cdots,r_N-1\},$$

with

$$\begin{aligned}
\mathfrak{M}(1,\mathbf{n},\mathbf{l}) &= \sum_{\substack{i_k=2l_kp_{\mathbf{n}}+1\\ k=1,\cdots,N}}^{2l_kp_{\mathbf{n}}+p_{\mathbf{n}}}\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\\
\mathfrak{M}(2,\mathbf{n},\mathbf{l}) &= \sum_{\substack{i_k=2l_kp_{\mathbf{n}}+1\\ k=1,\cdots,N-1}}^{2l_kp_{\mathbf{n}}+p_{\mathbf{n}}}\ \sum_{i_N=2l_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{(l_N+1)p_{\mathbf{n}}}\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\\
\mathfrak{M}(3,\mathbf{n},\mathbf{l}) &= \sum_{\substack{i_k=2l_kp_{\mathbf{n}}+1\\ k=1,\cdots,N-2}}^{2l_kp_{\mathbf{n}}+p_{\mathbf{n}}}\ \sum_{i_{N-1}=2l_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(l_{N-1}+1)p_{\mathbf{n}}}\ \sum_{i_N=2l_Np_{\mathbf{n}}+1}^{2l_Np_{\mathbf{n}}+p_{\mathbf{n}}}\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\\
\mathfrak{M}(4,\mathbf{n},\mathbf{l}) &= \sum_{\substack{i_k=2l_kp_{\mathbf{n}}+1\\ k=1,\cdots,N-2}}^{2l_kp_{\mathbf{n}}}\ \sum_{i_{N-1}=2l_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(l_{N-1}+1)p_{\mathbf{n}}}\ \sum_{i_N=2l_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(l_N+1)p_{\mathbf{n}}}\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\\
&\cdots
\end{aligned}$$

where the last one is

$$\mathfrak{M}(2^N, \mathbf{n}, \mathbf{l}) = \sum_{\substack{i_k=2l_k p_{\mathbf{n}}+p_{\mathbf{n}}+1 \\ k=1,\cdots,N}}^{2(l_k+1)p_{\mathbf{n}}} \Lambda_{\mathbf{i}}(\theta, \mathfrak{z}).$$

We note that, if the $n_{\mathbf{i}}$ is not exactly equal $2r_{\mathbf{i}}p_{\mathbf{n}}$, we regroup the remaining terms in $\mathfrak{T}(\mathbf{n}, 2^N + 1)$. Clearly, the first term $\mathfrak{T}(\mathbf{n}, 1)$ is the leading one and the quantity $\widehat{G^*}_1(\theta, \mathfrak{z}, t)$ is a finite sum. Thus, from (A.10), for $\eta > 0$, we get

$$Pr\left(|\mathbb{E}[\widehat{G^*}_1(\theta, \mathfrak{z}, t)] - \widehat{G^*}_1(\theta, \mathfrak{z}, t)| \geq \eta\right) \leq 2^N \max_{i=1,\cdots 2^N} Pr\left(\mathfrak{T}(\mathbf{n}, i) \geq \eta \widehat{\mathbf{n}} \mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]\right).$$

Therefore, (A.8) is a consequence of

$$Pr\left(\mathfrak{T}(\mathbf{n}, i) \geq \eta \widehat{\mathbf{n}} \mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]\right)$$

for all $i = 1, \cdots, 2^N$.

We now treat the leading term $\mathfrak{T}(\mathbf{n}, 1)$. Of course, the other cases are proved by the same treatment. The proof of this last case is based on the application of Lemma 2 in [46]. Indeed, we recount the

$$M = \prod_{k=1}^N r_k = 2^{-N} \widehat{\mathbf{n}} p_{\mathbf{n}}^{-N} \leq \widehat{\mathbf{n}} p_{\mathbf{n}}^{-N}$$

variables in another arbitrary way, that is $Z_1(\theta, \mathfrak{z}), \cdots, Z_M(\theta, \mathfrak{z})$. In a sense, for each $Z_{\mathbf{j}}(\theta, \mathfrak{z})$ there exists $\mathbf{j}$ in $\mathcal{J}$ such that

$$Z_{\mathbf{j}}(\theta, \mathfrak{z}) = \sum_{\mathbf{i} \in I(1, \mathbf{n}, \mathbf{j})} \Lambda_{\mathbf{i}}(\theta, \mathfrak{z}),$$

where

$$I(1, \mathbf{n}, \mathbf{l}) = \{\mathbf{i} : 2l_k p_{\mathbf{n}} + 1 \leq i_k \leq 2l_k p_{\mathbf{n}} + p_{\mathbf{n}},\ k = 1, \cdots, N\}.$$

Clearly, the sets $I(1, \mathbf{n}, \mathbf{l})$ are distanced by $p_{\mathbf{n}}$ and they contain $p_{\mathbf{n}}^N$ sites. In addition, we have

$$K(h_{\mathbf{n}}^{-1} N_\theta(\mathfrak{z} - \mathcal{X}_{\mathbf{i}})) Y_{\mathbf{i}}^* \leq C\gamma_{\mathbf{n}}.$$

Then, from Lemma 2 in [46], we extract independent random variables $Z_1^*(\theta, \mathfrak{z}), \cdots, Z_M^*(\theta, \mathfrak{z})$ having the same distribution as

$$Z_{l=1,\cdots,M}(\theta, \mathfrak{z}),$$

such that

$$\sum_{j=1}^r \mathbb{E}|Z_{\mathbf{j}}(\theta, \mathfrak{z}) - Z_{\mathbf{j}}^*(\theta, \mathfrak{z})| \leq 2C\gamma_{\mathbf{n}} M p_{\mathbf{n}}^N \psi(p_{\mathbf{n}}^N(M-1), p_{\mathbf{n}}^N) \varphi(p_{\mathbf{n}}). \tag{A.11}$$

So, we write

$$Pr\left(\mathfrak{T}(\mathbf{n}, i) \geq \eta \widehat{\mathbf{n}} \mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]\right) \leq \mathfrak{R}_1(\mathbf{n}) + \mathfrak{R}_2(\mathbf{n}),$$

where

$$\mathfrak{R}_1(\mathbf{n}) = Pr\left(\left|\sum_{j=1}^M Z_{\mathbf{j}}^*\right| \geq \frac{M\eta \widehat{\mathbf{n}} \mathbb{E}\left[K_{\mathbf{1}}(\theta, \mathfrak{z})\right]}{2M}\right),$$

$$\mathfrak{N}_2(\mathbf{n}) = Pr\left(\sum_{j=1}^{M} |Z_{\mathbf{j}} - Z_{\mathbf{j}}^*| \geq \frac{\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right]}{2}\right).$$

We start by evaluating the term $\mathfrak{N}_1$. As an independent array, we use the Bernstein's inequality. The latter is based on the variance quantity

$$\operatorname{Var}\left[Z_1^*(\theta,\mathfrak{z})\right] = \operatorname{Var}\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},\mathbf{1})} \Lambda_{\mathbf{i}}(\theta,\mathfrak{z})\right].$$

For this purpose, we use the fact that

$$\mathbb{E}\left[Y_{\mathbf{i}}^p|\mathcal{X}_{\mathbf{i}}\right] < \infty,$$

for $p > 2$, to prove that

$$\begin{aligned}
\operatorname{Var}\left[\Lambda_{\mathbf{i}}(\theta,\mathfrak{z})\right] &\leq C\mathbb{E}\left[K_{\mathbf{i}}^2 Y_{\mathbf{i}}^{*2}\right] \leq C\mathbb{E}\left[K_{\mathbf{i}}^2 Y_{\mathbf{i}}^2\right] \\
&\leq C\mathbb{E}\left[K_{\mathbf{i}}^2\mathbb{E}\left[Y_{\mathbf{i}}^2|\mathcal{X}_{\mathbf{i}}\right]\right] \\
&\leq C\mathbb{E}\left[K_{\mathbf{i}}^2\right] \leq C\phi(\theta,\mathfrak{z},h).
\end{aligned}$$

Therefore,

$$\sum_{\mathbf{i}\in I(1,\mathbf{n},\mathbf{1})} \operatorname{Var}\left[\Lambda_{\mathbf{i}}(\theta,\mathfrak{z})\right] = O\left(p_{\mathbf{n}}^N\phi(\theta,\mathfrak{z},h)\right).$$

Next, we use the second part of (H2) to prove that, for all $\mathbf{i} \neq \mathbf{j}$,

$$\begin{aligned}
\operatorname{Cov}(\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\Lambda_{\mathbf{j}}(\theta,\mathfrak{z})) &\leq C\mathbb{E}\left[K_{\mathbf{i}}K_{\mathbf{j}}|Y_{\mathbf{i}}Y_{\mathbf{j}}|\right] \leq C\mathbb{E}\left[K_{\mathbf{i}}K_{\mathbf{j}}\mathbb{E}\left[|Y_{\mathbf{i}}Y_{\mathbf{j}}||\mathcal{X}_{\mathbf{i}}\mathcal{X}_{\mathbf{j}}\right]\right] \\
&\leq C\mathbb{E}\left[K_{\mathbf{i}}K_{\mathbf{j}}\right] \leq C\phi_x^{(a+1)/a}(\theta,\mathfrak{z},h).
\end{aligned}$$

On the other hand, since

$$\mathbb{E}\left[Y_{\mathbf{i}}^p|\mathcal{X}_{\mathbf{i}}\right] < \infty,$$

then, for all $\mathbf{i} \neq \mathbf{j}$, Hölder's inequality allows us to write

$$\begin{aligned}
\operatorname{Cov}(\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\Lambda_{\mathbf{j}}(\theta,\mathfrak{z})) &\leq \|\Lambda_{\mathbf{i}}(\theta,\mathfrak{z})\|_p^2\varphi^{1-2/p}(\|j-i\|) \\
&\leq C\phi_x^{2/p}(\theta,\mathfrak{z},h)\varphi^{1-2/p}(\|i-j\|)).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})} \left|\operatorname{Cov}(\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\Lambda_{\mathbf{j}}(\theta,\mathfrak{z}))\right| \\
&\leq \sum_{\{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})\,\|\mathbf{i}-\mathbf{j}\|\leq u_{\mathbf{n}}\}} \left|\operatorname{Cov}(\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\Lambda_{\mathbf{j}}(\theta,\mathfrak{z}))\right| \\
&\quad + \sum_{\{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})\,\|\mathbf{i}-\mathbf{j}\|>u_{\mathbf{n}}\}} \left|\operatorname{Cov}(\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}),\Lambda_{\mathbf{j}}(\theta,\mathfrak{z}))\right| \\
&\leq Cp_{\mathbf{n}}^N\phi(\theta,\mathfrak{z},h)\left(u_{\mathbf{n}}^N\phi(\theta,\mathfrak{z},h)^{1/a} + u_{\mathbf{n}}^{-Na}\phi_x^{2/p-1}(\theta,\mathfrak{z},h)\sum_{\mathbf{i}:\|\mathbf{i}\|\geq u_{\mathbf{n}}} \|\mathbf{i}\|^{Na}\,\varphi^{1-2/p}\left(\|\mathbf{i}\|\right)\right).
\end{aligned}$$

It suffices to choose

$$u_{\mathbf{n}} = \phi(\theta, \mathfrak{z}, h)^{2/Np(a+1)-1/Na},$$

and we obtain

$$\sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})} \left|\mathrm{Cov}(\Lambda_{\mathbf{i}}(\theta,\mathfrak{z}), \Lambda_{\mathbf{j}}(\theta,\mathfrak{z}))\right| \leq C p_{\mathbf{n}}^{N} \phi(\theta, \mathfrak{z}, h).$$

So, we infer that

$$\mathrm{Var}\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},\mathbf{1})} \Lambda_{\mathbf{i}}(\theta,\mathfrak{z})\right] = O\left(p_{\mathbf{n}}^{N}\phi(\theta,\mathfrak{z},h)\right).$$

Now, we replace in the inequality of $\mathfrak{R}_1$, that is,

$$\mathfrak{R}_1(\mathbf{n}) \leq 2\exp\left(-\frac{(\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right])^2}{MVar\left[Z_1^*\right] + C\eta\gamma_{\mathbf{n}} p_{\mathbf{n}}^{N}\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right]}\right). \tag{A.12}$$

This gives

$$\mathfrak{R}_1(\mathbf{n}) \leq \exp\left(-C\eta_0 \log\widehat{\mathbf{n}}\right).$$

Finally, a good choice of η_0 allows us to write that

$$\sum_{\mathbf{n}} \mathfrak{R}_1(\mathbf{n}) < \infty.$$

For the term $\mathfrak{R}_2(\mathbf{n})$, we use the Markov inequality and (A.11) to get that

$$\mathfrak{R}_2(\mathbf{n}) \leq 2M\gamma_{\mathbf{n}} p_{\mathbf{n}}^{N} (\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right])^{-1} \psi(p_{\mathbf{n}}^{N}(M-1), p_{\mathbf{n}}^{N})\varphi(p_{\mathbf{n}}).$$

Now, since

$$\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right] \leq C\phi(\theta,\mathfrak{z},h_{\mathbf{n}}), \quad \widehat{\mathbf{n}} = 2^N M p_{\mathbf{n}}^{N}$$

and

$$\psi(p_{\mathbf{n}}^{N}(M-1), p_{\mathbf{n}}^{N}) \leq p_{\mathbf{n}}^{N},$$

by choosing

$$\eta = \eta_0 \sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\theta,\mathfrak{z},h)}},$$

we readily obtain

$$\mathfrak{R}_2(\mathbf{n}) \leq \widehat{\mathbf{n}}\gamma_{\mathbf{n}} p_{\mathbf{n}}^{N} \left(\log\widehat{\mathbf{n}}\right)^{-1/2} \left(\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h)\right)^{-1/2} \varphi(p_{\mathbf{n}}).$$

It suffices to choose

$$p_{\mathbf{n}} = C\left(\frac{\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h)}{\log\widehat{\mathbf{n}}\gamma_{\mathbf{n}}^2}\right)^{1/2N}$$

to get

$$\mathfrak{R}_2(\mathbf{n}) \leq \widehat{\mathbf{n}}\,\varphi(p_{\mathbf{n}}).$$

From (H5), we conclude that

$$\sum_{\mathbf{n}} \mathfrak{R}_2(\mathbf{n}) < \infty.$$

In a similar way, as $\mathbf{n} \to \infty$, we get

$$\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|\widehat{G}_2(\theta,\mathfrak{z},t) - \mathbb{E}\left[\widehat{G}_2(\theta,\mathfrak{z},t)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\phi(\theta,\mathfrak{z},h_{\mathbf{n}})}}\right),$$

which completes the demonstration of this lemma. □

Lemma 2. *Assume that the conditions (H1), (H3), and (H4) are fulfilled. We have, as* $\mathbf{n} \to \infty$,

$$\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|G_1(\theta,\mathfrak{z},t) - \mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]\right| = O\left(h_{\mathbf{n}}^{k_1}\right) \tag{A.13}$$

and

$$\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|G_2(\theta,\mathfrak{z},t) - \mathbb{E}\left[\widehat{G}_2(\theta,\mathfrak{z},t)\right]\right| = O\left(h_{\mathbf{n}}^{k_2}\right).$$

Proof of Lemma 2. Similar to the previous lemma, as the proof of two terms is the same, we focus on the term $\widehat{G}_1$. For this goal, we use the stationarity of the observations and write

$$\begin{aligned}&\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right] - G_1(\theta,\mathfrak{z},t)\\ =\ & \frac{1}{\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right]}\left\{\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\left(G_1(\theta,\mathcal{X}_{\mathbf{1}},t) - G_1(\theta,\mathfrak{z},t)\right)\right]\right\}.\end{aligned}$$

Then, from (H4), we get

$$\begin{aligned}&\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\left(G^1(\theta,;\mathcal{X}_{\mathbf{1}},t) - G_1(\theta,\mathfrak{z},t)\right)\right]\\ =\ & \mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\left(G_1(\theta,\mathcal{X}_{\mathbf{1}},t) - G_1(\theta,\mathfrak{z},t)\right)\mathbb{1}_{B(\mathfrak{z},h_n)}(\mathcal{X}_{\mathbf{1}})\right].\end{aligned}$$

Therefore, we infer

$$\begin{aligned}&\left|\mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right] - G_1(\theta,\mathfrak{z},t)\right|\\ =\ & \frac{1}{\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\right]}\left|\mathbb{E}\left[K_{\mathbf{1}}(\theta,\mathfrak{z})\left(G^1(\theta,\mathcal{X}_{\mathbf{1}},t) - G_1(\theta,\mathfrak{z},t)\right)\mathbb{1}_{B(\mathfrak{z},h_n)}(\mathcal{X}_{\mathbf{1}})\right]\right|.\end{aligned}$$

It follows that

$$\left|G_1(\theta,\mathfrak{z},t) - \mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]\right| \leq C h_{\mathbf{n}}^{k_1}.$$

The last inequality is uniform in t, then write

$$\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|G_1(\theta,\mathfrak{z},t) - \mathbb{E}\left[\widehat{G}_1(\theta,\mathfrak{z},t)\right]\right| = O\left(h_{\mathbf{n}}^{k_1}\right).$$

By the same analytical arguments, we obtain

$$\sup_{t\in[\xi_q(\theta,\mathfrak{z})-\delta,\xi_q(\theta,\mathfrak{z})+\delta]} \left|G_2(\theta,\mathfrak{z},t) - \mathbb{E}\left[\widehat{G}_2(\theta,\mathfrak{z},t)\right]\right| = O\left(h_{\mathbf{n}}^{k_2}\right).$$

Hence, the proof is complete. □

Lemma 3. *Under the conditions of Proposition 3.1, as* $\mathbf{n} \to \infty$*, we have*

$$\sum_{\mathbf{n}} Pr\left(\inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} \left| \widehat{G}_2(\theta, \mathfrak{z}, t) \right| \leq \epsilon \right) < \infty. \tag{A.14}$$

Proof of Lemma 3. Note that

$$\inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} \left| \widehat{G}_2(\theta, \mathfrak{z}, t) \right| \leq \frac{1}{2} \inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} G_2(\theta, \mathfrak{z}, t)$$
$$\Longrightarrow \sup_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} \left| \widehat{G}_2(\theta, \mathfrak{z}, t) - G_2(\theta, \mathfrak{z}, t) \right| > \frac{1}{2} \inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} G_2(\theta, \mathfrak{z}, t).$$

This statement means

$$Pr\left(\inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} \left| \widehat{G}_2(\theta, \mathfrak{z}, t) \right| \leq \frac{1}{2} G_2(\theta, \mathfrak{z}, t) \right)$$
$$\leq Pr\left(\sup_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} \left| \widehat{G}_2(\theta, \mathfrak{z}, t) - G_2(\theta, \mathfrak{z}, t) \right| > \frac{1}{2} \inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} G_2(\theta, \mathfrak{z}, t) \right).$$

Finally, we combine the Lemmas 1 and 2 and choose

$$\epsilon = \frac{1}{2} \inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} G_2(t; x) > 0$$

to conclude that

$$\sum_{\mathbf{n}} Pr\left(\inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta,\, \xi_q(\theta,\mathfrak{z})+\delta]} \left| \widehat{G}_2(\theta, \mathfrak{z}, t) \right| \leq \frac{1}{2} \inf_{t \in [\xi_q(\theta,\mathfrak{z})-\delta, \xi_q(\theta,\mathfrak{z})+\delta]} G_2(\theta, \mathfrak{z}, t \right) \;<\; \infty. \tag{A.15}$$

Hence, the proof is complete. □

AIMS Press