Research Article

The convergence rate of the hazard function with functional explanatory variable: case of spacial data with k Nearest Neighbor method

Nadjet Bellatrach [1], Wahiba Bouabsa [2] and Mohammed Kadi Attouch [3]

[1] Department of Mtahematics, Kasdi Merbah University of ouargula, Ouargula 30000, Algeria;

[1,2,3] Lab. of Statistics and Process Stochastic, University Djillali Liabes, BP 89 Sidi Bel Abbes 22000, Algeria,

email: (bellatrachnadjet@gmail.com (N. Bellatrach), wahiba_bouab@yahoo.fr (W. Bouabsa), attou_kadi@yahoo.fr (M. K. Attouch)).

Article History: Do Received: 02 November 2022; Revised: 05 December 2022; Accepted: 22 February 2023; Published online: 8 April 2023.

Abstract: In this paper we introduced a new hazard estimator when the co-variables are functional in nature. This estimator is a mix of both the k Nearest Neighbors shortly (kNN) procedure and spacial functional data. Then the convergence rate are introduced when the considered sample is collected in spatial order with mixing structure. In theory there is an estimation of the risk point then a discussion of application difficulties, such as data driven bandwidth choice. Furthermore, a comparison study based on simulated and real data is also provided to illustrate the performances and the usefulness of the kNN approach and to prove the highly sensitive of the kNN approach to the presence of even a small proportion of outliers in the data.

Keywords: Functional data, spacial data, kNN estimation, kernel estimate, convergence rate.

1. Introduction

The hazard operator is one of the most used models to explore the relationship between two random variables. In this paper, we deal with the problem of the nonparametric estimation of the conditional hazard function when the observations are spatial with the kNN approach.

The estimation of the hazard function plays a very important role in statistics. Indeed, it is used in risk analysis or for the study of survival phenomena in many fields such as (medicine, geophysics, reliability, . . .). The literature on hazard function estimation is very abundant, when the observations are vectorial. Take, for example [65, [59], [44], and [27], for recent references. In all these works, the authors consider independent observations or dependent data from time series. The first results on the nonparametric hazard estimation in functional statistics, were obtained by [31]. They studied the almost complete convergence of a kernel estimator for the hazard function of a real random variable conditional on a functional explanatory variable. [54] has shown that the kernel estimator presented by [28] is strongly consistent and asymptotically normally distributed. A generalization of these results in the spatial data case was obtained by [41]. More specifically, they studied the almost complete convergence of an adapted version of this estimator. The same authors have treated the L^2-convergence rate by giving the exact expression involved in the leading terms of the quadratic error and the asymptotic normality of the construct estimator (see [42]), then we cite for the most recent advances and references, [4], [39], [61], [62], [16], [33].

The importance of this research topic is motivated by the growth in the number of concrete problems for which data are collected in a spatial order. Such problems are encountered in many fields such as epidemiology, econometrics, environmental and earth sciences, agronomy, imaging, etc. The first results considered for spatial dependence have been obtained by [64]. He obtained the asymptotic normality for the density kernel estimator, whereas the nonparametric spatial regression problem has been studied by [50] and [7], who used the Nadaraya-Watson weights to obtain a kernel estimator, establishing the weak convergence and asymptotic distribution. The nonparametric auto-regression model in a prediction context on random fields has been studied by [12]. We refer to [32] for the almost complete uniform convergence on the functional component of this spacial nonparametric model .

Then [48] use the kNN technique to estimate the spatial nonparametric regression. They showed the asymptotic normality of the construct estimate. We return to [26], [58] and [23] for recent advances and references in nonparametric spatial data analysis.

The main purpose of this contribution is to use the kNN procedure. It is an alternative smoothing approach allows to estimate the hazard operator with varied bandwidth parameter strongly depends on the data. Precisely, the bandwidth parameter is priory defined according to the distance between the functional random variable. The kNN algorithm, an interactive method, permits to explore the topological as well as the specter component of the data. Pushed by this sophistic procedure on the bandwidth selection, the functional kNN smoothing approach has received a lot of attention in the last few years. Gyorfi's book [36] is a thorough analysis of kNN estimators in the finite dimensional context. Work in this area was started by [13], and a large number of articles are now available in various estimating contexts, which including regression, discrimination, density and mode estimation, and clustering analysis.

Then the important results in this topic were established by [11]. They gave the convergence rate of the almost complete consistency of the constructed estimator. Using the same techniques Attouch and Bouabsa [3] have established the almost couplet consistency of the conditional mode estimator. We refer to [4] for the conditional hazard function, we make reference to [14], [20], [45], [53], [19], [18], [6],[43], [11], [63], [49], [2], [40] and we cite for the most recent advances and references [39], [1], [9]. However, the difficulty in the kNN smoothing is the fact that the bandwidth parameter is a random variable, unlike the classical regression in which the smoothing parameter is a deterministic scalar. So, the study of the asymptotic properties of our proposed estimator is complicated, and it requires some additional tools and techniques. So the purpose of this paper show us that can be used to further investigate the estimation of functional nonparametric hazard opera in the case of spacial datasets. This is motivated by the fact that the robust hazard estimator has several advantages over the classical kernel regression estimator. The main profit in using a robust regression is that it allows reducing the effect of outlier data.

The exploration and analysis of data in presence of outliers is a great challenge in statistics. In particular, in spatial statistics several robust models resistant to these anomalies have been studied. Concerning the nonparametric modeling case, the first approach is given by [66]. The latter consider a local linear estimate of the regression function based on the least absolute deviation. Then[37] establish the consistency and asymptotic normality of a spatial version of the local linear estimate of the conditional quantiles. The spatial version of the M-estimation of the regression function has introduced by [34]. They obtained the almost complete convergence and the asymptotic normality of this estimate. [21] have paid attention to study nonparametric quantile regression by the L^1 method. They stated the weak consistency and the asymptotic normality of the constructed estimator.

In this paper under some general assumptions, we state the almost complete convergence (with rate). Noting that, this work is the link between the study of [41] with [39] and [1]. For instance, we remember that the complexity of our research comes from the fact that the bandwidth parameter in the kNN method is a random variable. Precisely, the bandwidth parameter is priory defined according to the distance between the functional random variable. Such consideration allows for exploring the topological as well as the specter component of the data.

In NFDA, kNN hazard with spacial data is new. We present the estimator of our spatial model with kNN method in Section 2. In Section 3 we give the assumptions, then we study the almost complete convergence of this estimator. In Section 4 we give all the results and their proofs. As an application we treat, in Subsection 5.1, the estimation of maximum risk conditional on a functional explanatory variable. In Subsection 5.2 we emphasize the UIB results' direct practical influence on data-driven bandwidth choice. Finally, a simulation study is given in Subsection 5.4 proove the good performances of our estimator then Subsection 5.5 is devoted to a real data application to evaluate the performance of this estimate. We conclude the paper by a conclusion Section 6.

2. The model and its estimator with the kNN method

Let M be a natural number in $\mathbb{N}^*$. We consider the random field $W_{\mathbf{i}} = (A_{\mathbf{i}}, B_{\mathbf{i}}), \mathbf{i} \in \mathbb{N}^M$ with values in $\mathcal{G} \times \mathbb{R}$, where $(\mathcal{G}, d)$ is a semi-metric space of possibly infinite dimension. In this context, $(A_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^M}$ can be a functional random variable. It should be noted that, for a good ten years, the statistical community has been preoccupied with the development of models and methods adapted to this context of functional data. While the first studies in this direction mainly focused on linear models (see [8], [56]), recent developments (see [28]) report non-parametric models suitable for this type of data.

Next, a point a in $\mathcal{G}$ (respectively, a compact $Q \in \mathbb{R}$), we assume that the spatial observations $(A_{\mathbf{i}}, B_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^M}$ have the same distribution as $W = (A, B)$ and that the regular version of the conditional probability of B knowing $A = a$ exists and admits a bounded density with respect to the Lesbegue measure on $\mathbb{R}$, denoted f_k^a. With kNN method the functional parameter studied in this article, denoted h_k^a, is defined, for all $b \in \mathbb{R}$ such that $F^a(b) < 1$, by

$$h_k^a(b) = \frac{f_k^a(b)}{1 - F_k^a(b)},$$

where, F_k^a is the conditional distribution function of B knowing $A = a$, with

$$F_k^a(\cdot) = \mathbb{P}(B \leqslant \cdot \,|A = a).$$

Furthermore, it was assumed that the functional random field is observed on the set $I_{\mathbf{n}} = \{\mathbf{i} = (i_1, \dots, i_M) \in \mathbb{N}^{\mathrm{M}}, 1 \leqslant i_\ell \leqslant n_\ell, \ell = 1, \dots, M\}$, $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^{\mathrm{M}}$ and estimated with the kNN method the cdf by

$$\widehat{F}_{\mathrm{k}}^{\mathrm{a}}(b) = \frac{\sum_{\mathbf{i}\in I_{\mathbf{n}}} L\left(P_{\mathrm{k}}^{-1} d(a, A_{\mathbf{i}})\right) J\left(h_{\mathrm{J}}^{-1}(b - B_{\mathbf{i}})\right)}{\sum_{\mathbf{i}\in I_{\mathbf{n}}} L\left(P_{\mathrm{k}}^{-1} d(a, A_{\mathbf{i}})\right)}, \quad \forall b \in \mathbb{R},$$

where L is a kernel and J is a conditional distribution function (cdf), defined by

$$J(\cdot) = \frac{\int_0 L(s)\mathrm{d}s}{\int_0^{1/2} L(s)\mathrm{d}s},$$

and $P_k = P_{k,\mathbf{n}}$ (resp. $h_J = h_{J,\mathbf{n}}$) is a sequence of positive real numbers which belong to ar interval $(x_{\mathbf{n}}, y_{\mathbf{n}})$ (resp. $(v_{\mathbf{n}}, s_{\mathbf{n}})$), with $\lim_{\mathbf{n}\to\infty} v_{\mathbf{n}} = \lim_{\mathbf{n}\to\infty} s_{\mathbf{n}} = 0$.

Although, recently, this kind of kernel estimation is widely studied in the literature, these studies exclusively concern fixed bandwidths. In our paper we derive UIB consistency with the corresponding rates and emphasise applications for the construction of automatic data-driven smoothing parameters in Section.

From $\widehat{F}_k^a$, we deduce an estimator of the conditional density, denoted $\hat{f}_k^a$, where

$$f_k^a(\cdot) = (F_k^a)'(\cdot).$$

So

$$\hat{f}_{\mathrm{k}}^{\mathrm{a}}(b) = \frac{h_J^{-1} \sum_{\mathbf{i}\in I_{\mathbf{n}}} L\left(P_{\mathrm{k}}^{-1} d(a, A_{\mathbf{i}})\right) J'\left(h_{\mathrm{J}}^{-1}(b - B_{\mathbf{i}})\right)}{\sum_{\mathbf{i}\in I_{\mathbf{n}}} L\left(P_{\mathrm{k}}^{-1} d(a, A_{\mathbf{i}})\right)}, \quad \forall b \in \mathbb{R},$$

where J' is the derivative of J. The kNN estimator of the conditional hazard function is noted by

$$\hat{h}_{\mathrm{k}}^{\mathrm{a}}(b) = \frac{\hat{f}_k^a(b)}{1 - \hat{F}_k^a(b)}, \quad \forall b \in \mathbb{R},$$

The aim of this work is to study the almost complete convergence of the estimator $\hat{h}_k^a$ to h_k^a, when the functional random field $(W_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^M}$ satisfies the following mixing condition.

$$\begin{cases}\text{There is a function } \Phi(s) \downarrow 0 \text{ when } s \to \infty, \text{such that} \\ \forall Z, Z' \text{subsetof } \mathbb{N}^M \text{ finite cardinal} \\ \alpha\big(\mathcal{B}(Z), \mathcal{B}(Z')\big) = \sup_{Y \in \mathcal{B}(Z)} \sup_{E \in \mathcal{B}(Z')} |\mathbb{P}(Y \cap E) - \mathbb{P}(Y)\mathbb{P}(E)| \\ \leqslant \Psi\big(Card(Z), Card(Z')\big)\Phi\big(dist(Z, Z')\big),\end{cases}$$

where, $\mathcal{B}(Z)$ (resp. $\mathcal{B}(Z')$) the Borelian tribe generated by $(W_{\mathbf{i}}, \mathbf{i} \in Z)$ (resp. $(W_{\mathbf{i}}, \mathbf{i} \in Z')$), $Card(E)$ (resp. Card (Z')) is the cardinal of Z (resp. Z'), $dist(Z, Z')$ designates the Euclidean distance between Z and Z' and Ψ a symmetric function: $\mathbb{N}^2 \to \mathbb{R}^+$, decreasing with respect to the two variables separately and satisfying one of the following conditions

$$\Psi(u, v) \leqslant C\min(u, v), \quad u, v \in \mathbb{N},, \qquad (1)$$

or

$$\Psi(u, v) \leqslant C(u + v + 1)^{\tilde{\vartheta}}, \quad u, v \in \mathbb{N}, \qquad (2) \qquad (2)$$

for some $\tilde{\vartheta} \geqslant 1$ and $C > 0$, note that these conditions were used by [59] and they are verified by many spatial models (see [34]).

Recall that when equation (1) holds with $\Psi \equiv 1$ or $N = 1$, the random field $W_{\mathbf{i}} = (A_{\mathbf{i}}, B_{\mathbf{i}})$ is said to as highly mixing .

In addition, we suppose that the process W meets the following mixing condition:

$$\sum_{i=1}^{\infty} i^{\ell}\Phi(i) < \infty, \ell > 0. \qquad (3) \qquad (3)$$

We note that the conditions (2) and (3) are identical to mixing conditions utilized by [10] and [63].

3. Asymptotic properties

In the following, the author denotes by C and/or C' any strictly positive constants. Recall that in this spatial context, $\mathbf{n} \to \infty$ means that $\min\{\mathbf{n}_\ell\} \to \infty$ and that for each $1 \leqslant j, \ell \leqslant M$ we have $\infty > C > \left|\frac{\mathbf{n}_j}{\mathbf{n}_\ell}\right|$. Let us introduce the following hypotheses

- (K1)
  - (K1-i) $\mathbb{P}(A \in B(a, t)) = \varphi_a(t) > 0$ where $B(a, t)$ is the closed ball, centered at a and of radius t.
  - (K1-ii) for all $q \in (0,1), \lim_{t \to 0} \frac{\varphi_a(qt)}{\varphi_x(t)} = \varsigma_a(q) < \infty$.
- (K2) The function Φ checks $\sum_{i=1}^{\infty} i^{\gamma}\Phi(i) < \infty, \gamma > 5M$.
- (K3)

$0 < \sup_{\mathbf{i} \neq \mathbf{j}} \mathbb{P}\big[(A_{\mathbf{i}}, A_{\mathbf{j}}) \in B(a, P_k) \times B(a, P_k)\big] \leqslant C(\varphi_a(P_k))^{(x+1)/x}$, with $1 < x < \gamma M^{-1}$.

- (K4) Recall that a is a fixed functional element and N_a is a fixed neighbourhood of a. The nonparametric model for the conditional distribution and the conditional density is constructed by assuming that $C > 0$ such that for all $(b_1, b_2) \in \mathbb{R}^2$ and all $(a_1, a_2) \in N_a \times N_a$, one obtains
  - (K4-i)

$$\left|F_{\mathrm{k}}^{a_1}(b_1) - F_{\mathrm{k}}^{a_2}(b_2)\right| \leqslant C\big(d(a_1, a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4}\big),$$

$$\beta_3 > 0, \beta_4 > 0.$$

- (K4-ii)

$$\left|F_{\mathrm{k}}^{a_1}(b_1) - F_{\mathrm{k}}^{a_2}(b_2)\right| \leqslant C\big(d(a_1, a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4}\big),$$

$$\beta_3 > 0, \beta_4 > 0.$$

• (K5) kernel L has a continuous first derivative on $(0,1/2)$ and is supported inside $(0,1/2)$. Furthermore, there are some constants $0 < C < C' < \infty$, such as

$$0 < C\mathrm{I}_{\left(\frac{0,1}{2}\right)}(\cdot) \leqslant L(\cdot) \leqslant C'\mathrm{I}_{\left(\frac{0,1}{2}\right)},$$

and

$$L(1/2) - \int_0^{1/2} L'(s)\varsigma_a(s)\mathrm{d}s > 0.$$

• (K6) J is a function of class C^2 and a support compact.

• (K7) There exist $0 < \alpha < (\delta - 5M)/3M$ and $\zeta_0 > 0$, such as

$$\lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}^{\alpha} h_J = \infty \quad \text{and} \quad C\widehat{\mathbf{n}}^{\frac{(5+3\alpha)M-\delta}{\delta}} + \zeta_0 \leqslant h_J \varphi_a(P_k),$$

where $\widehat{\mathbf{n}} = n_1 \dots n_M$.

• (K8)

- (K8-i)

$$\mathcal{G}_0 = \left\{\cdot \mapsto \frac{1}{\int_0^1 L(s)\mathrm{d}s}\int_0^{\left(\gamma^{-1}(y-\cdot)\right)} L(s)\mathrm{d}s, \text{ for } \delta > 0\right\},$$

is a pointwise measurable class.

- (K8-ii)

$$\mathcal{L}_0 \text{is such that} \sup_U \int_0^1 \sqrt{1 + \log \mathcal{N}\left(\varepsilon \|K_0\|_{U,2}, \mathcal{L}_0, d_U\right)} \mathrm{d}\varepsilon < \infty,$$

where K_0 is the envelope function of the set $\mathcal{L}_0$.

• (K9)

- (K9-i)

$$\frac{\log \mathbf{n}}{\mathbf{n}\min\left(x_{\mathbf{n}}, \varphi_a(x_{\mathbf{n}})\right)} \to 0.$$

- (K9-ii)

$$\frac{\log \mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\min\left(x_{\mathbf{n}}, \varphi_a(x_{\mathbf{n}})\right)} \to 0.$$

Remarks on the hypotheses

Our research is the link between the work of [38], and [40], so several assumptions are the same considered in all this researches.

4. Result and proof

Theorem 4.1 *Under the conditions (1) and (2) and hypotheses (K1)-(K7), and (K9), we have that*

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}} \leqslant P_k \leqslant y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leqslant h_J \leqslant s_{\mathbf{n}}} \left|\widehat{h}_k^a(b) - h_k^a(b)\right| = O\left(y_{\mathbf{n}}^{\beta_3}\right) + O\left(s_{\mathbf{n}}^{\beta_4}\right) + O_{a.co.}\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right).$$

Proof

The proof of this theorem is based on the fact that

$$\left|\hat{h}_k^a(b)-h_k^a(b)\right| \leqslant \frac{1}{\left|1-\hat{F}_k^a(b)\right|}\left[\left|\hat{f}_k^a(b)-f^a(b)\right| \frac{|f_k^a(b)|}{|1-F_k^a(b)|}\left|\hat{F}_k^a(b)-F_k^a(b)\right|\right].$$

Theorem 4.1 is obtained from Theorems 4.2 and Theorem 4.3 and from Corollary 4.1 below.

Corollary 4.1 *Under the conditions of Theorem 4.1, there exists a* $\delta>0$ *such that*

$$\sum_{\mathbf{n}\in\mathbb{N}^M} \mathbb{P}\left\{\sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}\left|1-\hat{F}_k^a(b)\right|<\delta\right\}<\infty.$$

Proof

Remember that any $P_k\in(x_{\mathbf{n}},y_{\mathbf{n}})$ and $h_J\in(v_{\mathbf{n}},s_{\mathbf{n}})$, we have

$$\left|1-\hat{F}_k^a(b)\right| \leqslant \frac{(1-F_k^a(b))}{2} \;\; implies \;\; \left|\hat{F}_k^a(b)-F_k^a(b)\right| \geqslant \frac{(1-F_k^a(b))}{2}.$$

Therefore, we have that

$$\sum_{\mathbf{n}\in\mathbb{N}^M} \mathbb{P}\left(\inf_{P_k\in(x_{\mathbf{n}},y_{\mathbf{n}}),h_J\in(v_{\mathbf{n}},s_{\mathbf{n}})}\left|1-\hat{F}_k^a(b)\right| \leqslant \frac{1-F_k^a(b)}{2}\right)$$
$$\leqslant \sum_{\mathbf{n}\in\mathbb{N}^M} \mathbb{P}\left(\sup_{P_k\in(x_{\mathbf{n}},y_{\mathbf{n}}),h_J\in(v_{\mathbf{n}},s_{\mathbf{n}})}\left|\hat{F}_k^a(b)-F_k^a(b)\right| \geqslant \frac{1-F_k^a(b)}{2}\right)<\infty.$$

Applying Theorem 4.2 now gives the desired result.

Theorem 4.2 *Under hypotheses (K1), and (K7), (K8-ii) and (K9-i) we have that*

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \sup_{k}\left|\hat{F}_k^a(b)-F_k^a(b)\right| = O\left(y_{\mathbf{n}}^{\beta_3}\right)+O\left(s_{\mathbf{n}}^{\beta_4}\right)+O_{a.co.}\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\varphi_a(x_{\mathbf{n}})}}\right).$$

Proof

The demonstrations are based respectively on the following decompositions

$$\hat{F}_k^a(b)-F_k^a(b)=\frac{1}{\hat{F}_D^a}\left\{\left(\hat{F}_N^a(b)-\mathbb{E}\hat{F}_N^a(b)\right)-\left(F_k^a(b)-\mathbb{E}\hat{F}_N^a(b)\right)\right\}+$$
$$+\frac{F_k^a(b)}{\hat{F}_D^a}\left(\mathbb{E}\hat{F}_D^a-\hat{F}_D^a\right). \tag{4}$$

Define

$$\hat{F}_D^a=\frac{1}{\hat{\mathbf{n}}\mathbb{E}\left[L\left(P_k^{-1}d(a,A_1)\right)\right]}\sum_{\mathbf{i}\in I_{\mathbf{n}}} L\left(P_k^{-1}d(a,A_{\mathbf{i}})\right),$$
$$\hat{F}_N^a(b)=\frac{1}{\hat{\mathbf{n}}\mathbb{E}\left[L\left(P_k^{-1}d(a,A_1)\right)\right]}\sum_{\mathbf{i}\in I_{\mathbf{n}}} L\left(P_k^{-1}d(a,A_{\mathrm{i}})\right)J\left(h_J^{-1}(b-B_{\mathrm{i}})\right),$$

where **1** is the spatial index of the fixed components 1.

The proof follows from the Lemmas bellows.

Lemma 4.1. *Under hypotheses (1) and (2) and hypotheses (K1), (K3), (K5) and (K7)-(K9-i) we have that*

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\left|\hat{F}_D^a-\mathbb{E}\hat{F}_D^a\right| = O\left(\left(\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}\varphi(x_{\mathbf{n}})}\right)^{\frac{1}{2}}\right) \quad a.co.$$

Proof

In the following note for all $\mathbf{i} \in I_{\mathbf{n}}$

$$L_{\mathrm{i}}(a) = L\left(P_k^{-1} d(a, A_{\mathrm{i}})\right), \quad J_{\mathrm{i}}(b) = J\left(h_J^{-1}(b - B_{\mathrm{i}})\right).$$

One has to show that there exists $\zeta_0 > 0$ such that

$$\sum_n \mathbb{P}\left\{\sup_{x_{\mathbf{n}} \leqslant P_k \leqslant y_0} \sqrt{\frac{n\varphi_a(x_{\mathbf{n}})}{\log n}} \left|\hat{F}_D^a - \mathbb{E}\left[\hat{F}_D^a\right]\right| \geqslant \zeta_0\right\} < \infty, for\ some\ y_0 > 0.$$

Following Bernstein's inequality in ([24], p. 321) the proof is based on Bernstein's inequality for empirical processes, by defining

$$h_{L,j} = 2^j x_{\mathbf{n}} \text{ and } K(n) = max\{j: h_{L,j} \leqslant 2y_0\},$$

hence

$$\sup_{x_{\mathbf{n}} \leqslant P_k \leqslant y_0} \sqrt{\frac{n\varphi_a(x_{\mathbf{n}})}{\log n}} \left|\hat{F}_D^a - \mathbb{E}\left[\hat{F}_D^a\right]\right| \leqslant \max_{1 \leqslant j \leqslant K(\mathbf{n})} \sup_{h_{L,j-1} \leqslant P_k \leqslant h_{L,j}} \sqrt{\frac{\mathbf{n}\varphi_a(P_k)}{\log n}} \left|\hat{F}_D^a - \mathbb{E}\left[\hat{F}_D^a\right]\right|.$$

To write the difference, the demonstration is based on the concepts similar to those used by [12], Thus

$$\hat{F}_D{}^a(a) - \mathbb{E}[\hat{F}_D{}^a(a))] = \frac{1}{\hat{\mathbf{n}}\mathbb{E}[L_{\mathbf{1}}(a)]} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \Theta_{\mathbf{i}}(a),$$

where $\Theta_{\mathbf{i}}(a) = (1/\sqrt{\mathbf{n}}) \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} (L_i - \mathbb{E}[L_i])$ corresponds to the empirical process based on variables $A_1, A_2, \dots, A_{\mathbf{n}}$. Then, consider the following class of functions

$$\mathcal{L}_{L,j} = \{z \longmapsto L(\gamma^{-1} d(a,z)) \text{ where } h_{L,j-1} \leqslant \gamma \leqslant P_{k,j}\}.$$

Therefore,

$$\leqslant \sum_{\mathbf{i}=\mathbf{1}}^{K(\mathbf{n})} \mathbb{P}\left\{\frac{1}{\sqrt{\mathbf{n}\varphi_a\left(\frac{P_{k,j}}{2}\right)\log\mathbf{n}}} \left\|\sqrt{\mathbf{n}}\Theta_{\mathbf{i}}(a)\right\|\mathcal{L}_{L,j} \geqslant \zeta_0\right\},$$

$$\leqslant K(\mathbf{n}) \max_{j=1,\dots,K(\mathbf{n})} \mathbb{P}\left\{\max_{1 \leqslant k \leqslant \mathbf{n}} \left\|\sqrt{k}\Theta_k(a)\right\|_{\mathcal{L}_{L,j}} \geqslant \zeta_0 \sqrt{\mathbf{n}\varphi_a\left(\frac{P_{k,j}}{2}\right)\log\mathbf{n}}\right\}. \quad (5)$$

We consider the spatial decomposition of [60] on on $\Theta_{\mathbf{i}}(a)$ variables, defined, for the fixed integer $\varrho_{\mathbf{n}}$, as follows

$$V(1, \mathbf{n}, a, \mathbf{j}) = \sum_{\substack{i\ell = 2j\kappa \\ \kappa = 1,\dots,\cdots \\ \kappa,\dots,1}}^{2j\kappa\varrho_{\mathbf{n}} + \varrho_{\mathbf{n}}} \Theta_{\mathrm{i}}(a),$$

$$V(2, \mathbf{n}, a, \mathbf{j}) = \sum_{\substack{i_\kappa = 2j_\kappa\varrho_{\mathbf{n}}+1 \\ \kappa = 1,\dots,M}}^{2j_\kappa\varrho_{\mathbf{n}} + \varrho_{\mathbf{n}}} \sum_{\substack{ij \\ M}}^{2(j\varrho_{\mathbf{n}} + \varrho_{\mathbf{n}} + 1} \Theta_{\mathrm{i}}(a),$$

$$V(3, \mathbf{n}, x, \mathbf{j}) = \sum_{\substack{i_k = 2j_k\varrho_{\mathbf{n}} + 1 i_{N-1} = 2j_{N-1}\varrho_{\mathbf{n}} + \varrho_{\mathbf{n}} + 1 \\ k = 1,\dots,N-2}}^{2j_k\varrho_{\mathbf{n}} + \varrho_n} \sum_{\substack{n \\ j_N}}^{2(j\rho_{\mathbf{n}} + 1} \sum_{\mathbf{i}}^{2j_N\varrho_{\mathbf{n}} + \varrho_{\mathbf{n}}} \Theta_{\mathrm{i}}(a),$$

$$V(4,\mathbf{n},x,\mathbf{j})=\sum_{\substack{i_h=2j_h\varrho_\mathbf{n}+1\\k=1,\ldots,N-2}}^{2(j_N-1+1)\varrho_\mathbf{n}}\ \sum_{i_{N-1}=2jN-1}^{2(j_N+1)\varrho_\mathbf{n}+\varrho_\mathbf{n}+1}\ \sum_{i_N=2jN}\Theta_{\varrho_\mathbf{n}+\varrho_\mathbf{n}+1}\Theta_\mathbf{i}(a),$$

and so on. The last two terms are as follows

$$V(2^{M-1},\mathbf{n},a,\mathbf{j})=\sum_{\substack{i_\kappa=2j_\kappa\varrho_\mathbf{n}+\varrho_\mathbf{n}+1\\\kappa=1,\ldots,M-1}}^{2(j_\kappa+1)\varrho_\mathbf{n}}\ \sum_{i_M=2j_M\varrho_\mathbf{n}+1}^{2j_M\varrho_\mathbf{n}+\varrho_\mathbf{n}+1}\Theta_\mathrm{i}(a),$$

$$V(2^M,\mathbf{n},a,\mathbf{j})=\sum_{\substack{i_\kappa=2(j_\kappa+1)\varrho_\mathbf{n}\\\kappa=1,\ldots,M}}^{2(j_\kappa+1)\varrho_\mathbf{n}}\Theta_\mathrm{i}(a).$$

For $\tau_i=2^{-1}n_i{\varrho_\mathbf{n}}^{-1}, i=1,\ldots,M$ and $\mathcal{H}=\{0,\ldots,\tau_1-1\}\times\ldots\times\{0,\ldots,\tau_M-1\}$, we pose

$$\Delta(\mathbf{n},a,i)=\sum_{\mathbf{j}\in\mathcal{H}}V(i,\mathbf{n},a,\mathbf{j}).$$

Without loss of generality, one can write

$$\left|{\hat{F}_D}^a(a)-\mathbb{E}\left[{\hat{F}_D}^a(a)\right]\right|=\frac{1}{\hat{\mathbf{n}}\mathbb{E}[L_\mathbf{1}(a)]}\sum_{i=1}^{2^M}\Delta(\mathbf{n},a,i).\qquad(6)$$

Even if $n_\mathbf{i}$ is not exactly equal to $2\tau_\mathbf{i}\varrho_\mathbf{n}$, one can group the remaining variables in a block $\Delta(\mathbf{n},a,2^M+1)$ (this will not change the proof see [7]).).

Now, under the last equation (6), for all $\zeta>0$, one obtains

$$\mathbb{P}\left(\left|{\hat{F}_D}^a a(a)-\mathbb{E}\left[{\hat{F}_D}^a(a)\right]\right|\geqslant\zeta\right)\leqslant2^M\max_{i=1,\ldots}\mathbb{P}(\Delta(\mathbf{n},a,i)\geqslant\zeta\hat{\mathbf{n}}\mathbb{E}[L_\mathbf{1}(a)]).$$

Hence, it suffices to calculate

$$\mathbb{P}(\Delta(\mathbf{n},a,i)\geqslant\zeta\hat{\mathbf{n}}\mathbb{E}[L_\mathbf{1}(a)])\quad\text{for all}\quad i=1,\ldots,2^M.$$

This only deals with the case $i=1$. For this, we number the variables $(V(1,\mathbf{n},a,\mathbf{j});\mathbf{j}\in\mathcal{H})$ and apply (Lemma 4.1 of [10]) on the re-numbered variables. with the new numbering are noted $D_1,\ldots,D_N$,

where

$N=\prod_{\kappa=1}^M\tau_\kappa=2^M\hat{\mathbf{n}}{\varrho_\mathbf{n}}^{-M}\leqslant\hat{\mathbf{n}}{\varrho_\mathbf{n}}^{-M}$. Note that for all D_j there is a certain **j** in $\mathcal{H}$ such as

$$D_j=\sum_{\mathbf{i}\in I(1,\mathbf{n},a,\mathbf{j})}\Theta_\mathbf{i}(a),$$

where $I(1,\mathbf{n},a,\mathbf{j})=\{\mathbf{i}:2j_\kappa\varrho_\mathbf{n}+1\leqslant i_\kappa\leqslant2j_\kappa\varrho_\mathbf{n}+\varrho_\mathbf{n};\kappa=1,\ldots,M\}$. The distance between these sets is greater than ${\varrho_\mathbf{n}}^M$ and each set contains ${\varrho_\mathbf{n}}^M$ elements.

The definition (Lemma 4.1 of [10]) allows to approximate $D_1,D_2,\ldots,D_N$ by independent random variables $D_1^*,\ldots,D_N^*$ of the same law as $D_{j=1,\ldots,N}$ and such that

$$\sum_{j=1}^N\mathrm{E}\left|D_j-D_j^*\right|\leqslant2CN(\varrho_\mathbf{n}^M\psi(N-1)\varrho_\mathbf{n}^M,\varrho_\mathbf{n}^M)\varphi(\varrho_\mathbf{n}).$$

Then, by the Bernstein and Markov inequalities

$$\mathbb{P}(\Delta(\mathbf{n},a,i)\geqslant\zeta\hat{\mathbf{n}}\mathbb{E}[L_\mathbf{1}(a)])\leqslant\beta_1+\beta_2,$$

$$\beta_1 = \mathbb{P}\left(\left|\sum_{j=1}^{N} D_j^*\right| \geqslant \frac{N\zeta\hat{\mathbf{n}}\mathbb{E}[L_1(a)]}{2N}\right)$$
$$\leqslant 2\exp\left(-\frac{\left(\zeta\hat{\mathbf{n}}\mathbb{E}[L_1(a)]\right)^2}{N\mathrm{Var}[D_1^*]+C\varrho_{\mathbf{n}}^M\zeta\hat{\mathbf{n}}\mathbb{E}[L_1(a)]}\right),$$

where

$$\beta_2 = \mathbb{P}\left(\sum_{j=1}^{N}\left|D_j - D_j^*\right| \geqslant \frac{\zeta\hat{\mathbf{n}}\mathbb{E}[L_1(a)]}{2}\right) \leqslant \frac{C}{\zeta\hat{\mathbf{n}}\mathbb{E}[L_1(a)]}\sum_{j=1}^{N}\mathbb{E}\left|D_j - D_j^*\right|$$
$$\leqslant CN\varrho_{\mathbf{n}}^M\left(\zeta\hat{\mathbf{n}}\mathbb{E}[L_1(a)]\right)^{-1}\psi((N-1)\varrho_{\mathbf{n}}^M,\varrho_{\mathbf{n}}^M)\varphi(\varrho_{\mathbf{n}}).$$

As $\hat{\mathbf{n}} = 2^M N\varrho_{\mathbf{n}}^M$ et $\psi((N-1)\varrho_{\mathbf{n}}^M,\varrho_{\mathbf{n}}^M) \leqslant \varrho_{\mathbf{n}}^M$ so for $\zeta = \zeta_0\sqrt{\frac{\log\hat{\mathbf{n}}}{\hat{\mathbf{n}}\varphi_a(P_{k,j}/2)}}$

$$\beta_2 \leqslant \hat{\mathbf{n}}\varrho_{\mathbf{n}}^M(\log\hat{\mathbf{n}})^{-1/2}\left(\hat{\mathbf{n}}\varphi_a(P_k)\right)^{-1/2}\varphi(\varrho_{\mathbf{n}}),$$

Take $\varrho_{\mathbf{n}}C\left(\frac{\hat{\mathbf{n}}\varphi_a(P_{k,j}/2)}{\log\hat{\mathbf{n}}}\right)^{1/2M}$, then

$$\beta_2 \leqslant (\log\hat{\mathbf{n}})^{-1}\hat{\mathbf{n}}\varphi(\varrho_{\mathbf{n}}), \tag{7}$$

according to hypotheses (K7), to show that $\sum_{\mathbf{n}\in I_{\mathbf{n}}}\hat{\mathbf{n}}\varphi(\varrho_{\mathbf{n}}) < \infty$.

Now one can deal with β_1. For this, it suffices to evaluate $Var[D_1^*]$. In effect

$$Var[D_1^*] = Var\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},a,\mathbf{1})}\Theta_{\mathbf{i}}(a)\right] = \sum_{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},a,\mathbf{1})}\left|Cov\left(\Theta_{\mathbf{i}}(a),\Theta_{\mathbf{j}}(a)\right)\right|.$$

Let us pose $Z_{\mathbf{n}} = \sum_{\mathbf{i}\in I(1,\mathbf{n},a,\mathbf{1})} Var[\Theta_{\mathbf{i}}(a)]$ and $T_{\mathbf{n}} = \sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},a,\mathbf{1})}\left|Cov\left(\Theta_{\mathbf{i}}(a),\Theta_{\mathbf{j}}(a)\right)\right|$. By virtue of (K1), one has $Var[\Theta_{\mathbf{i}}(a)] \leqslant C\left(\varphi_a\left(P_{k,j}/2\right) + \left(\varphi_a\left(P_{k,j}/2\right)\right)^2\right)$, so $Z_{\mathbf{n}} = O\left(\varrho_{\mathbf{n}}^M\varphi_a\left(P_{k,j}/2\right)\right)$. As it concerns $T_{\mathbf{n}}$, the techniques of [52] are used considering the following sets

$$Q_1 = \{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},a,\mathbf{1}): 0 < \|\mathbf{i}-\mathbf{j}\| \leqslant \omega_{\mathbf{n}}\}$$
$$Q_2 = \{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},a,\mathbf{1}): \|\mathbf{i}-\mathbf{j}\| > \omega_{\mathbf{n}}\},$$

where $\omega_{\mathbf{n}}$ is a real sequence tending to $+\infty$. Thereby,

$$T_{\mathbf{n}} = \sum_{(\mathbf{i},\mathbf{j})\in Q_1}\left|Cov\left(\Theta_{\mathbf{i}}(a),\Theta_{\mathbf{j}}(a)\right)\right| + \sum_{(\mathbf{i},\mathbf{j})\in Q_2}\left|Cov\left(\Theta_{\mathbf{i}}(a),\Theta_{\mathbf{j}}(a)\right)\right| = T_{\mathbf{n}}^1 + T_{\mathbf{n}}^2.$$

On the one hand, we have

$$T_{\mathbf{n}}^1 = \sum_{(\mathbf{i},\mathbf{j})\in Q_1}\left|\mathbb{E}\left[L_{\mathbf{i}}(a)L_{\mathbf{j}}(a)\right] - \mathbb{E}[L_{\mathbf{i}}(a)]\mathbb{E}\left[L_{\mathbf{j}}(a)\right]\right|$$
$$\leqslant C\varrho_{\mathbf{n}}^M\omega_{\mathbf{n}}^M\varphi_a(P_k)\left(\left(\varphi_a(P_k)\right)^{\frac{1}{a}} + \varphi_a(P_k)\right) \leqslant C\varrho_{\mathbf{n}}^M\Omega_{\mathbf{n}}^M\varphi_a(P_k)^{(a+1)/a}.$$

While on the other

$$T_{\mathbf{n}}^2 = \sum_{(\mathbf{i},\mathbf{j})\in Q_2}\left|Cov\left(\Theta_{\mathbf{i}}(a),\Theta_{\mathbf{j}}(a)\right)\right|.$$

Since kernel L is bounded, from ([60], Lemma 2.1(ii)), one obtains

$$\left|Cov\left(\Delta_{\mathbf{i}}(a),\Delta_{\mathbf{j}}(a)\right)\right| \leqslant C\varphi(\|\mathbf{i}-\mathbf{j}\|).$$

Thus

$$T_{\mathbf{n}}^2 \leqslant C\sum_{(\mathbf{i},\mathbf{j})\in Q_2}\varphi(\|\mathbf{i}-\mathbf{j}\|) \leqslant C\varrho_{\mathbf{n}}^M\sum_{\mathbf{i}:\|\mathbf{i}\|\geqslant\omega_{\mathbf{n}}}\varphi(\|\mathbf{i}\|)$$
$$\leqslant C\varrho_{\mathbf{n}}^N\omega_{\mathbf{n}}^{-Na}\sum_{\mathbf{i}:\|\mathbf{i}\|\geqslant\omega_{\mathbf{n}}}\|\mathbf{i}\|^{Na}\varphi(\|\mathbf{i}\|).$$

Take $\omega_{\mathbf{n}} = \left(\phi_x(h_K)\right)^{-\frac{1}{Na}}$,

so

$$T_{\mathbf{n}}^2 \leqslant C\varrho_{\mathbf{n}}^M \omega_{\mathbf{n}}^{-Ma} \sum_{\mathbf{i}:\|\mathbf{i}\|\geqslant\omega_{\mathbf{n}}} \| \mathbf{i} \|^{Ma} \varphi(\| \mathbf{i} \|)$$
$$\leqslant C\varrho_{\mathbf{n}}^N \phi_x(h_K) \sum_{\mathbf{i}:\|\mathbf{i}\|\geqslant\omega_{\mathbf{n}}} \| \mathbf{i} \|^{Ma} \varphi(\| \mathbf{i} \|).$$

From (K2) one can write $T_{\mathbf{n}}^2 \leqslant C\varrho_{\mathbf{n}}^M \varphi_a(P_k)$. Moreover, with this choice of $\omega_{\mathbf{n}}$, one obtains

$$T_{\mathbf{n}}^1 \leqslant C\varrho_{\mathbf{n}}^M \varphi_a\left(P_{k,j}/2\right),$$

where

$$Var[D_1^*] = O\left(\varrho_{\mathbf{n}}^M \varphi_a\left(P_{k,j}/2\right)\right).$$

Using this last result, with the definitions of $\varrho_{\mathbf{n}}, N$ and ζ, it is shown that

$$\beta_1 \leqslant \exp(-C\zeta_0 \log\widehat{\mathbf{n}}).$$

(K9-i) and moreover, since $L(n) \leqslant 2\log n$, from equation (5)

$$K(n) \max_{j=1,\ldots,K(n)} \mathbb{P}\left\{\max_{1\leqslant k\leqslant n} \left\|\sqrt{\mathbf{n}}\Theta_n(a)\right\|_{\mathcal{L}_{L,j}} > \zeta_0 \sqrt{\mathbf{n}\varphi_a\left(\frac{P_{k,j}}{2}\right)\log \mathbf{n}}\right\}$$
$$\leqslant C(\log\mathbf{n})\mathbf{n}^{-C'_0\zeta^2}.$$

Now, by choosing ζ_0 such that $C'_0\zeta^2 > 1$, one obtains:

$$\sup_{x_n\leqslant P_k\leqslant y_n} \left|\widehat{F}_D{}^a - \mathbb{E}\left[\widehat{F}_D{}^a\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}\varphi(x_{\mathbf{n}})}}\right).$$

Lemma 4.2. *Under hypotheses (K1), (K4-i), (K5) and (K7)*

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \left|F_k^a(b) - \mathbb{E}\hat{F}_N^a(b)\right| = O\left(y_{\mathbf{n}}^{\beta_3}\right) + O\left(s_{\mathbf{n}}^{\beta_4}\right).$$

Proof.

Given the fact that all random variables are distributed in the same way,

$$\forall b \in \mathcal{S}, \sup_{x_{\mathbf{n}}\leqslant P_K\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant v_{\mathbf{n}}} \left|F_N^a(b) - \mathbb{E}\hat{F}_N^a(b)\right| = \mathbb{E}\left[L_1(a)\mathbb{I}_{B(a,P_k)}(A)[\mathbb{E}[J_1(b)/A] - F^a(b)]\right].$$

Integrating by parties, one can see that

$\mathbb{E}(J_1(b)/A) = \int_{\mathbb{R}} J\left(h_J^{-1}(b-z)\right) f_k^A(z)\mathrm{d}z = h_J^{-1} \int_{\mathbb{R}} J'\left(h_J^{-1}(b-z)\right) F_k^A(z)\mathrm{d}z.$

Taking into account the change in a common variable $t = \frac{b-z}{h_J}$, one obtains

$$\mathbb{E}(J_1(b)/A) = \int_{\mathbb{R}} J'(b) F_k^A\left(b - h_J t\right)\mathrm{d}t,$$

thus

$$|\mathbb{E}(J_1(b)/A) - F_k^a(b)| \leqslant \int_{\mathbb{R}} J'(b)\left|F_k^A\left(b - h_J t\right) - F_k^a(b)\right|\mathrm{d}t.$$

Under (K4)

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} (A)\left|\mathbb{E}\left(\frac{J_1(b)}{A}\right) - F_k^a(b)\right|$$

$$\leqslant \int_{\mathbb{R}} J'(b)\left(P_k^{\beta_1} + |t|^{\beta_2} h_J^{\beta_2}\right)\mathrm{d}t.$$

Knowing that J' is a probability density, hypothesis (K6) thus completes the demonstration of Lemma 4.2, leading finally to

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \left|F_k^a(b) - \mathbb{E}\hat{F}_N^a(b)\right| = \mathrm{C}(y_{\mathbf{n}}^{\beta_3} + s_{\mathbf{n}}^{\beta_4}).$$

Lemma 4.3. *Under the same conditions of Theorem 4.2*

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \left|\hat{F}_N^a(b) - \mathbb{E}\hat{F}_N^a(b)\right| = O\left(\left(\frac{\log \hat{\mathbf{n}}}{\hat{\mathbf{n}}\varphi_x(x_{\mathbf{n}})}\right)^{\frac{1}{2}}\right), \qquad a.co.$$

Proof

By the compactness of S,

$$S \subset \bigcup_{j=1}^{\sigma_{\mathbf{n}}} \left(\Gamma_j - \wp_{\mathbf{n}}, \Gamma_j + \wp_{\mathbf{n}}\right),$$

with $\wp_{\mathbf{n}} = \mathbf{n}^{-\alpha-1/2}$ and $\sigma_{\mathbf{n}} = O\left(\mathbf{n}^{\alpha+1/2}\right)$. Then the monotony of $\mathbb{E}\left[\hat{f}_N^a(b)\right]$ and $\hat{f}_N^a(b)$ gives, for $1 \leqslant j \leqslant \sigma_{\mathbf{n}}$, that

$$\mathbb{E}\left[\hat{F}_N^a\left(\Gamma_j - \wp_{\mathbf{n}}\right)\right] \leqslant \mathbb{E}\left[\hat{F}_N^a\left(\Gamma_j + \wp_{\mathbf{n}}\right)\right]$$
$$\hat{F}_N^a\left(\Gamma_j - \wp_{\mathbf{n}}\right) \leqslant \hat{F}_N^a\left(\Gamma_j + \wp_{\mathbf{n}}\right).$$

Now, from condition (K4-i), one obtains, for any $b_1, b_2 \in S$, that

$$\left|\mathbb{E}\left[\hat{F}_N^a(b_1)\right] - \mathbb{E}\left[\hat{F}_N^a(b_2)\right]\right| \leqslant C\left(d(a_1,a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4}\right).$$

It follows that

$$\sup_{t\in S}\left|\hat{F}_N^a(b) - \mathbb{E}\left[\hat{F}_N^a(b)\right]\right|$$
$$\leqslant \max_{1\leqslant j\leqslant \sigma_{\mathbf{n}},\ z\in\{\Gamma_j-\wp_{\mathbf{n}},\Gamma_j+\wp_{\mathbf{n}}\}} \left|\hat{f}_N^a(z) - \mathbb{E}\left[\hat{F}_N^a(z)\right]\right| + 2C\frac{\wp_{\mathbf{n}}}{h_J}\hat{F}_D^a.$$

Then,

$$\wp_{\mathbf{n}} = o\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right).$$

Thus, all it remains to prove is that

$$\mathbb{P}\left(\sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \max_{1\leqslant j\leqslant \sigma_n,\ z\in\{\Gamma_j-\wp_{\mathbf{n}},\Gamma_j+\wp_{\mathbf{n}}\}} \left|\hat{F}_N^a(z) - \mathbb{E}\left[\hat{F}_N^a(z)\right]\right| > \zeta\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right),$$
$$\leqslant 2\sigma_{\mathbf{n}} \max_{1\leqslant j\leqslant \sigma_{\mathbf{n}}} \max_{z\in\{\Gamma_j-\wp_{\mathbf{n}},\Gamma_j+\wp_{\mathbf{n}}\}} \max_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \max_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \left|\hat{F}_N^a(z) - \mathbb{E}\left[\hat{F}_N^a(z)\right]\right| > \zeta\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_n)}}\right).$$

Now, look at the quantity

$$\mathbb{P}\left(\sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \left|\hat{F}_N^a(z) - \mathbb{E}\left[\hat{F}_N^a(z)\right]\right| > \zeta\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_n)}}\right),$$

for all $z = \Gamma_j \mp \wp_{\mathbf{n}}, 1 \leqslant j \leqslant \sigma_{\mathbf{n}}$. The proof of the above inequality is based on Bernstein's inequality for empirical processes, i.e.

$$\Theta_{i_1}(a) = \frac{1}{\sqrt{\mathbf{n}}}\sum_{i=1}^{\mathbf{n}} \left(L_i J_i - \mathbb{E}[L_i(a)J_i(a)]\right).$$

Then, one obtains for all $z = \Gamma_j \mp \wp_{\mathbf{n}},\ 1 \leqslant j \leqslant \sigma_{\mathbf{n}}$, that

$$\mathbb{P}\left\{\sup_{x_{\mathbf{n}} \leqslant P_k \leqslant y_0} \sup_{v_{\mathbf{n}} \leqslant h_j \leqslant s_0} \left|\hat{F}_N^a(z) - \mathbb{E}\left[\hat{F}_N^a(z)\right]\right| \geqslant \zeta_0{}'\right\} \leqslant \sigma_{\mathbf{n}}(\beta_1 + \beta_2).$$

The definition of $\sigma_{\mathbf{n}}$, equation(7) and hypotheses (K7) allow to write

$$\sum_{\mathbf{n} \in I_{\mathbf{n}}} \hat{\mathbf{n}}^{\alpha+\frac{1}{2}}\ \beta_2 < \infty.$$

A suitable choice of ζ_0 gives

$$\sum_{\mathbf{n} \in I_{\mathbf{n}}} \hat{\mathbf{n}}^{\alpha+\frac{1}{2}}\ \beta_1 < \infty.$$

With the same technic of demonstration like Lemma 4.1, the only difference is that $\Theta_{\mathbf{i_1}}(a)$ is used instead of $\Theta_{\mathbf{i}}(a)$. Hence this leads finally to

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leqslant P_k \leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}} \leqslant h_j \leqslant s_{\mathbf{n}}} |F_N^a(b) - \mathbb{E}\,\hat{F}_N^a(b)| = O_{a.co.}\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\varphi_a(x_{\mathbf{n}})}}\right).$$

Corollary 4.1. *Under the hypotheses of Lemma 4.1*

$$\sum_{\mathbf{n} \in \mathbb{N}^N} \mathbb{P}\left(\inf_{x_{\mathbf{n}} \leqslant P_k \leqslant y_{\mathbf{n}}} \hat{F}_D^a < C\right) < \infty.$$

Proof

By simple analytical arguments, one obtains from equations (K1) and

(K5)-(K6) $\mathbb{E}\left[\frac{1}{\mathbf{n}\varphi_a(P_k)}\sum_{i=1}^{\mathbf{n}} L\left(P_k^{-1}d(a,A_i)\right)\right] \to L(1/2) - \int_0^{1/2} L'(q)\varsigma_a(s)\mathrm{d}s > 0.$

Then, for $\mathbf{n}$ large enough there exists a constant $C' > 0$, such that

$$\mathbb{E}\left[\hat{F}_D^a\right] \geqslant C' \text{forall} \quad P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}}).$$

As a result of selecting $C = C'/2$, one obtains

$$\mathbb{P}\left(\inf_{P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}})} \hat{F}_D^a \leqslant C\right) \leqslant \mathbb{P}\left(\sup_{P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}})} \left|\mathbb{E}\left[\hat{F}_D^a\right] - \hat{F}_D^a\right| \geqslant C\right),$$

and Lemma 4.1 leads to the desired result.

Theorem 4.3 *Under the hypotheses (K1),(K7) and,(K8-ii), (K9-ii) we have that*

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leqslant P_k \leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}} \leqslant h_j \leqslant s_{\mathbf{n}}} \left|\hat{f}_k^a(b) - f_k^a(b)\right| = O\left(y_{\mathbf{n}}^{\beta_3}\right) + O\left(s_{\mathbf{n}}^{\beta_4}\right) + O_{a.co.}\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n} v_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}}\right).$$

Proof

The demonstrations are based, respectively, on the following decompositions

$$\hat{f}_k^a(b) - f_k^a(b) = \frac{1}{\hat{F}_D^a}\left\{\left(\hat{f}_N^a(b) - \mathbb{E}\hat{f}_N^a(b)\right) - \left(f_k^a(b) - \mathbb{E}\hat{f}_N^a(b)\right)\right\} + \\ + \frac{f_k^a(b)}{\hat{F}_D^a}\left(\mathbb{E}\hat{F}_D^a - \hat{F}_D^a\right). \quad (8)$$

Define

$$\hat{f}_N^a(b) = \frac{1}{\hat{\mathbf{n}} h_J \mathbb{E}\left[L\left(P_k^{-1} d(a, A_{\mathbf{1}})\right)\right]} \sum_{\mathbf{i}\in I_{\mathbf{n}}} L\left(P_k^{-1} d(a, A_{\mathbf{i}})\right) J'\left(h_J^{-1}(b - B_{\mathbf{i}})\right).$$

So, the following two Lemmas are required to prove Theorem 4.3.

Lemma 4.4 *Under the hypotheses (K1),(K4-ii), (K5)-(K7) and (K9-ii) we have*

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \left|f_k^a(b) - \mathbb{E}\hat{f}_N^a(b)\right| = O\left(y_{\mathbf{n}}^{\beta_3}\right) + O\left(s_{\mathbf{n}}^{\beta_4}\right).$$

Proof

Using the stationarity of data, the conditional on the explicative variable, and the change in the usual variable $t = \frac{b-z}{h_J}$, we obtain

$$h_J^{-1}\mathbb{E}(J_1'(b)/A) = h_J^{-1}\int_{\mathbb{R}} J'(b) f_k^A\left(b - h_J t\right)\mathrm{d}t,$$

and we deduce that

$$\left|h_J^{-1}\mathbb{E}(J_1'(b)/A) - f_k^a(b)\right| \leqslant \int_{\mathbb{R}} J'(b)\left|f^A\left(b - h_J t\right) - f_k^a(b)\right|\mathrm{d}t.$$

Under the condition (K4-ii)

$$\forall b \in \mathcal{S},\quad 1 = \ \mathrm{I}\ = \mathrm{I}_{B(a,P_k)}(A)\left|h_J^{-1}\mathbb{E}(J_1'(b)/A) - f_k^a(b)\right| \leqslant \int_{\mathbb{R}} J'(b)\left(h_k^{\beta_3} + |t|^{\beta_4} h_J^{\beta_4}\right)\mathrm{d}t.$$

Knowing that J' is a probability density, the hypothesis (K6) thus completes the demonstration of Lemma 4.4, so this leading finally to

$$\forall b \in \mathcal{S}, \sup_{x_{\mathbf{n}}\leqslant P_K\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant v_{\mathbf{n}}} \left|\ f_k^a(b) - \mathbb{E}\hat{f}_N^a(b)\right| \leqslant C\left(y_{\mathbf{n}}^{\beta_3} + s_{\mathbf{n}}^{\beta_4}\right).$$

Lemma 4.5 *Under the same conditions of Theorem 4.3 , we have*

$$\sup_{b\in\mathcal{S}} \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}} \sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}} \left|\hat{f}_N^a(b) - \mathbb{E}\hat{f}_N^a(b)\right| = O_{a.co.}\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}\quad v_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right).$$

Proof

The demonstration is very similar to that of Lemma 4.3. Indeed, we consider the covering

$$S \subset \bigcup_{j=1}^{\sigma_{\mathbf{n}}} \left(\Gamma_j - \wp_{\mathbf{n}}, \Gamma_j + \wp_{\mathbf{n}}\right),$$

with $\wp_{\mathbf{n}} = \mathbf{n}^{-1/2+\frac{-3}{2}\alpha}$ and $\sigma_{\mathbf{n}} = O\left(\mathbf{n}^{1/2+\frac{3}{2}\alpha}\right)$. Then the monotony of $\mathbb{E}\left[\hat{f}_N^a(b)\right]$ and $\hat{f}_N^a(b)$ gives, for $1 \leqslant j \leqslant \sigma_{\mathbf{n}}$, that

$$\mathbb{E}\left[\hat{f}_N^a\left(\Gamma_j - \wp_{\mathbf{n}}\right)\right] \leqslant \mathbb{E}\left[\hat{f}_N^a\left(\Gamma_j + \wp_{\mathbf{n}}\right)\right]$$
$$\hat{f}_N^a\left(\Gamma_j - \wp_{\mathbf{n}}\right) \leqslant \hat{f}_N^a\left(\Gamma_j + \wp_{\mathbf{n}}\right).$$

Now, from the condition (K4-ii), we obtain, for any $b_1, b_2 \in S$, that

$$\left|\mathbb{E}\left[\hat{f}_N^a(b_1)\right] - \mathbb{E}\left[\hat{f}_N^a(b_2)\right]\right| \leqslant C\left(d(a_1, a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4}\right).$$

It follows that

$$\sup_{z\in S}\left|\hat{f}_N^a(b) - \mathbb{E}\left[\hat{f}_N^a(b)\right]\right|$$
$$\leqslant \max_{1\leqslant j\leqslant\sigma_{\mathbf{n}}} \max_{z\in\{\Gamma_j-\wp_{\mathbf{n}},\Gamma_j+\wp_{\mathbf{n}}\}} \left|\hat{f}_N^a(z) - \mathbb{E}\left[\hat{f}_N^a(z)\right]\right| + 2C\frac{\wp_{\mathbf{n}}}{h_J^2}\hat{F}_D^a.$$

So,

$$\wp_{\mathbf{n}} = o\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n} v_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right).$$

Therefore, all it remains to prove is that

$$\mathbb{P}\left(\sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}\max_{1\leqslant j\leqslant\sigma_{\mathbf{n}}}\max_{z\in\{\Gamma_j-\wp_{\mathbf{n}},\Gamma_j+\wp_{\mathbf{n}}\}}\left|\hat{f}_N^a(z)-\mathbb{E}\left[\hat{f}_N^a(z)\right]\right|>\zeta\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right)$$
$$\leqslant 2\sigma_{\mathbf{n}}\max_{1\leqslant j\leqslant\sigma_{\mathbf{n}}}\max_{z\in\{\Gamma_j-\wp_{\mathbf{n}},\Gamma_j+\wp_{\mathbf{n}}\}}\max_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\max_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}\left|\hat{f}_N^a(z)-\mathbb{E}\left[\hat{f}_N^a(z)\right]\right|>\zeta\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\Bigg).$$

Now, we look at the quantity

$$\mathbb{P}\left(\sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}\left|\hat{f}_N^a(z)-\mathbb{E}\left[\hat{f}_N^a(z)\right]\right|>\zeta\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_n)}}\right),$$

for all $z=\Gamma_j\mp\wp_{\mathbf{n}}, 1\leqslant j\leqslant\sigma_{\mathbf{n}}$. The proof of the above inequality is based on the Bernstein's inequality for empirical processes,

$$\Theta_{\mathbf{i_2}}(a)=\frac{1}{\sqrt{\mathbf{n}}}\sum_{i=1}^{\mathbf{n}}\left(L_iJ_i'-\mathbb{E}[L_iJ_i']\right),$$

Then, we get, for all $z=\Gamma_j\mp\wp_{\mathbf{n}}, 1\leqslant j\leqslant\sigma_{\mathbf{n}}$, that

$$\mathbb{P}\left\{\sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{0}}}\sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{0}}}\left|\hat{f}_N^a(z)-\mathbb{E}\left[\hat{f}_N^a(z)\right]\right|\geqslant\zeta_0'\right\}\leqslant\sigma_{\mathbf{n}}(\beta_1'+\beta_2'),$$

where, $\beta_1'\leqslant\exp(-C(\zeta_0)\log\hat{\mathbf{n}})$ and

$$B_2'\leqslant\hat{\mathbf{n}}\varrho_{\mathbf{n}}^M(\log\hat{\mathbf{n}})^{-1/2}\left(\hat{\mathbf{n}}h_J\varphi_a(P_k)\right)^{-1/2}\varphi(\varrho_{\mathbf{n}}).$$

We take, $\varrho_{\mathbf{n}}C\left(\frac{\hat{\mathbf{n}}h_J\varphi_a(P_k)}{\log\hat{\mathbf{n}}}\right)^{1/2M}$, the definition of $z_{\mathbf{n}}$ under hypothesis (K7), a good choice of ζ_0 allow us to write

$$\sum_{\mathbf{n}\in I_{\mathbf{n}}}\hat{\mathbf{n}}^{\frac{1}{2}}(\beta_2'+\beta_1')<\infty.$$

We use the same technic of demonstration like Lemma 4.1, the only difference is that we utilise $\Theta_{\mathbf{i_2}}(a)$ instead of $\Theta_{\mathbf{i}}(a)$ So this leading finally to

$$\sup_{b\in\mathcal{S}}\sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\sup_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}\left|f_N^a(b)-\mathbb{E}\hat{f}_N^a(b)\right|=O_{a.co.}\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}v_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right).$$

5 Application

5.1 Estimation of the risk point

In this part, we suggest estimating the high-risk point in $\mathcal{S}$, noted $\eta(a)$, defined as

$$h_k^a\big(\eta(x)\big)=\max_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\max_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}\max_{b\in\mathcal{S}}h_k^a(b).\qquad(9)\qquad(9)$$

This model has a strong background in statistics, particularly in risk analysis (see [54]). In our functional setting, we assume that there is a single point $\eta(x)$ in $\mathcal{S}$ with $\sup_{P_k\in(x_{\mathbf{n}},y_{\mathbf{n}})}$, $\sup_{h_J\in(v_{\mathbf{n}},s_{\mathbf{n}})}$ and verified (9). The natural estimator of $\eta(x)$, denoted by $\hat{\eta}(x)$, is as follows

$$h_k^a\big(\eta(a)\big)=\max_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\max_{v_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}\max_{b\in\mathcal{S}}h_k^a(b).\qquad(10)\qquad(10)$$

In general, this estimate is not unique. As a result, throughout the rest of this article, $\hat{\eta}(a)$ will be used to denote any random variable verified (10).

To analyze the almost complete convergence rate of the estimator $\hat{\eta}(a)$, we keep the same hypotheses from the previous section and assume that the function h_k^a is of class C^2 in respect to b, as shown below.

$$h_k^{a'}\big(\eta(a)\big)=0\quad\text{and}\quad h_k^{a''}\big(\eta(a)\big)>0.\qquad(11)\qquad(11)$$

Finally, Theorem 4.1 allows us to deduce the following corollary

Corollary 5.1 *Under the conditions (1), (2), (K1)-(K7), and (K9), and (11) we have*

$$\sup_{b\in\mathcal{S}}\ \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\ \sup_{\nu_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}|\hat{\eta}(a)-\eta(a)| = O\left(y_{\mathbf{n}}^{\beta_3/2}\right)+O\left(s_{\mathbf{n}}^{\beta_4/2}\right)+O_{a.co.}\left(\left(\frac{\log\mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}\right)^{1/4}\right).$$

Proof

Under the hypothesis (K5), Taylor's development of the function h_k^a in the neighborhood of $\eta(a)$, particularly for the point $\hat{\eta}(a)$, is

$$h_k^a(\hat{\eta}(a)) = h_k^a(\eta(a)) + (\hat{\eta}(a)-\eta(a))^2\frac{h_k^{a''}(\eta^*)}{j!},$$

where, η^* is between $\hat{\eta}(a)$ and $\eta(a)$, so it is necessarily in the compact $\mathcal{S}$. Thus

$$|\hat{\eta}(a)-\eta(a)|^2 \leqslant \frac{2!}{\min_{b\in\mathcal{S},P_k\in(x_{\mathbf{n}},y_{\mathbf{n}}),h_J\in(\nu_{\mathbf{n}},s_{\mathbf{n}})} h_k^{a''}(b)}|h_k^a(\hat{\eta}(a))-h_k^a(\eta(a))|,$$

and with the use of analytic arguments, it is demonstrated that

$$|h_k^a(\hat{\eta}(a))-h_k^a(\eta(a))| \leqslant 2\sup_{b\in\mathcal{S}}\ \sup_{P_k\in(x_{\mathbf{n}},y_{\mathbf{n}})}\ \sup_{h_J\in(\nu_{\mathbf{n}},s_{\mathbf{n}})}\left|\hat{h}_k^a(b)-h_k^a(b)\right|.$$

Using the result of Theorem 4.1, we show that

$$\sup_{b\in\mathcal{S}}\ \sup_{x_{\mathbf{n}}\leqslant P_k\leqslant y_{\mathbf{n}}}\ \sup_{\nu_{\mathbf{n}}\leqslant h_J\leqslant s_{\mathbf{n}}}|\hat{\eta}(a)-\eta(a)| = O\left(y_{\mathbf{n}}^{\beta_3}\right)+O\left(s_{\mathbf{n}}^{\beta_4}\right)+O_{a.co.}\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right).$$

5.2 Implementation of uniforme in bandwidth for the choice of bandwidth results

When applying a nonparametric estimator in practice, one is challenged with the bandwidth choice difficulty. Such a decision should be based on data, which means that the bandwidths that will be utilized in practice are random variables that will vary depending on the entire statistical sample. In statistical models, the asymptotic theory for data-driven bandwidth has always been connected to some kind of uniform in bandwidth studies shortly (UIB). (notice, for example ([38], "Lemma 1, p. 1473"), even if the word UIB did not arise in the literature until much later.

Throughout this section, we'll look at how the general UIB results reported earlier in this work can be applied to a wide range of data-driven bandwidth applications in of functional spatial with mixing structure data. This will be described in Section 5.3 the estimation problem (cdf, density and hazard functions).

5.3 Selecting a Bandwidth in hazard Distribution Function

The same kind of ideas can be developed for dealing with random bandwidths in other setting than regression. To avoid tedious repetitions, the situations investigated in Section 2 will be presented all together. Let $\tilde{P}_k$ and $\tilde{h}_J$ be random variables taking values respectively in the intervals $(x_{\mathbf{n}}, y_{\mathbf{n}})$ and $(\nu_{\mathbf{n}}, s_{\mathbf{n}})$. Denote by $\tilde{F}_k, \tilde{f}_k$ and $\tilde{h}_k$ the estimators obtained by plugging in the random bandwidths $\tilde{P}_k$ and $\tilde{h}_J$ into $\hat{F}_k, \hat{f}_k$ and $\hat{h}_k$. Then, as direct applications of Theorems 4.2, 4.3 and Theorem 4.1, one gets the following results

$$\tilde{F}_k^a(b)-F_k^a(b) = O\left(y_{\mathbf{n}}^{\beta_3}\right)+O\left(s_{\mathbf{n}}^{\beta_4}\right)+O\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}\varphi_a(x_{\mathbf{n}})}}\right),\ \text{a.co.} \qquad (12) \qquad (12)$$

$$\tilde{f}_k^a(b)-f^a(b) = O\left(y_{\mathbf{n}}^{\beta_3}\right)+O\left(s_{\mathbf{n}}^{\beta_4}\right)+O\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right),\text{a.co.} \qquad (13) \qquad (13)$$

and

$$\tilde{h}_k^a(b)-h_k^a(b) = O\left(y_{\mathbf{n}}^{\beta_3}\right)+O\left(s_{\mathbf{n}}^{\beta_4}\right)+O\left(\sqrt{\frac{\log\mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}\right),\text{a.co.} \qquad (14)$$

As shown in the examples below, these generic solutions for random bandwidths have obvious applicability in data-driven bandwidth selection problems.

Example1-Conditional distribution function

When focusing on c.d.f. estimation, a statistical prediction can be performed by means of the conditional median, where $\hat{B}_i^{(\prime)}$ is the leave-one-out prediction of B_i is defined to be the solution of the following equation:

$$\hat{F}_k^{A_i,-i}(v) = \frac{1}{2},$$

where,

$$\hat{F}_k^{a,-i} = \frac{\sum_{\mathbf{i}\in I_\mathbf{n}} \sum_{j\neq i} L\left(P_k^{-1} d\left(a, A_j\right)\right) J\left(h_J^{-1}(b - B_\mathbf{i})\right)}{\sum_{\mathbf{i}\in I_\mathbf{n}} \sum_{j\neq i}^{\mathbf{n}} L\left(P_k^{-1} d\left(a, A_j\right)\right)}.$$

In practice, based on the statistical sample, one utilizes a data-driven bandwidth, defined as

$$\tilde{P}_k = \tilde{P}_k(A_1, B_1, \ldots, A_M, B_M).$$

As a result of this, consistency results can be achieved for any type of automatic data-driven bandwidth selection as a consequence of (12).

One of the most ahead of the scheduled is cross-validation, which involves minimizing the following least square prediction criterion

$$\left(\tilde{p}_{k,CV}^{(\prime)}, \tilde{h}_{J,CV}^{(\prime)}\right) = \arg\min_{P_k\in(x_\mathbf{n}, y_\mathbf{n}), h_J\in(v_\mathbf{n}, s_\mathbf{n})} \sum_{i=1}^{\mathbf{n}} \left(B_i - \hat{B}_i^{(\prime)}\right)^2.$$

Example2-Density function

The conditional mode can be used to make predictions from the conditional density estimator. B_i could well be predicted precisely as the answer to the following maximisation problem

$$\hat{B}_i^{(\prime\prime)} = \operatorname{argmax} \hat{f}_k^{A_i,-i}(b),$$

where the leave-one-out approach defines $\hat{f}^{a,-i}$ is

$$\hat{f}_k^{a,-i} = \frac{h_J^{-1} \sum_{\mathbf{i}\in I_\mathbf{n}} \sum_{j\neq i} L\left(P_k^{-1} d\left(a, A_j\right)\right) J'\left(h_J^{-1}(b - B_\mathbf{i})\right)}{\sum_{\mathbf{i}\in I_\mathbf{n}} \sum_{j\neq i}^{\mathbf{n}} L\left(P_k^{-1} d\left(a, A_j\right)\right)}.$$

Then, there are some more obvious choices for data-driven bandwidths.

$$\left(\tilde{P}_{k,CV}^{(\prime\prime)}, \tilde{h}_{J,CV}^{(\prime\prime)}\right) = \arg\min_{P_k\in(x_\mathbf{n}, y_\mathbf{n}), h_J\in(v_\mathbf{n}, s_\mathbf{n})} \sum_{i=1}^{\mathbf{n}} \left(B_i - \hat{B}_i^{(\prime\prime)}\right)^2.$$

The following results is direct consequences of Theorems 4.2, 4.3 and Theorem 4.1.

Let $j = 1$ or $j = 2$, and denote by $\hat{F}_k^{(j),a}, \hat{f}_k^{(j),a}$ and $\hat{h}_k^{(j),a}$, respectively, the estimators $\hat{F}_k, \hat{f}_k$ and $\hat{S}$ constructed with the cross-validated bandwidths $\tilde{P}_{k,CV}^{(j)}$ and $\tilde{h}_{J,CV}^{(j)}$. Then, the following three results hold:

- Under the conditions of Theorem 4.2, we have that:

$$\left|\hat{F}_k^{(j),a}(b) - F_k^a(b)\right| = O\left(y_\mathbf{n}^{\beta_3}\right) + O\left(s_\mathbf{n}^{\beta_4}\right) + O\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\varphi_a(x_\mathbf{n})}}\right), \text{a.co.}$$

- Under the hypotheses of Theorem 4.3, we obtain that

$$\left|\hat{f}_k^{(j),a}(b) - f_k^a(b)\right| = O\left(y_\mathbf{n}^{\beta_3}\right) + O\left(s_\mathbf{n}^{\beta_4}\right) + O\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n} v_\mathbf{n} \varphi_a(x_\mathbf{n})}}\right), a.co.$$

- Under the conditions of Theorem 4.1, we have that

$$\left|\hat{h}_k^{(j),a}(b) - h_k^a(b)\right| = O\left(y_\mathbf{n}^{\beta_3}\right) + O\left(s_\mathbf{n}^{\beta_2}\right) + O\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n} v_\mathbf{n} \varphi_a(x_\mathbf{n})}}\right), \text{a.co.}$$

These are of particular interest in practice since they fully automate the practical concerns that arise when employing the estimators $\hat{F}_k, \hat{f}_k$ and $\hat{h}_k$, our paper appears to be the first in this area for hazard, conditional distribution, conditional density settings.

5.4 Simulated data application

In this section we compare the finite-sample performance of kNN hazard function in the context of functional prediction via classical hazard function via a short Monte Carlo study. More precisely, we compare the finite-sample efficiency of both regression functions as spatial prediction tools. In order to highlight the main feature of our procedure, we compare their sensitivity to the presence of outliers. For this purpose, we consider the following model

$$B_{\mathbf{i}} = 5\log((4 - A_{\mathbf{i}})^2 + 2) + \epsilon_{\mathbf{i}}, \mathbf{i} = (i_1, i_2),$$

where $\epsilon_{(i_1,i_2)}$ follows a normal distribution with mean 0 and variance 0.25. For the sake of simplicity, we consider the same univariate spatial process A_1 used by [15] defined by

$$A_{\mathbf{i}} = \left(\sqrt{(2/m)}\right)\sum_{\varpi=1}^{m} \left(\cos\left(i_1 w_{\varpi,1} + i_2 w_{\varpi,2} + v_{\varpi}\right)\right),$$

where $w_{\varpi,s}, s = 1,2 \quad \varpi = 1,2 \ldots 500$ are independently, identically distributed with the standard normal distribution and are independent of $v_{\varpi}, \varpi = 1, \ldots 500$ which are independently and identically distributed with the uniform distribution on $[-\pi, \pi]$. Recall that as $m \to \infty, A_1$ is a Gaussian spatial ergodic process (see, [15]) which is an example of α-mixing spatial process. We generate the the random field $(A_{\mathbf{1}}, B_{\mathbf{i}})$ at $\hat{\mathbf{n}} = 30 \times 30$ sites.

The spatial correlation is controlled in our theoretical analysis by the strong mixing condition, which is defined as (1)- (3). Therefore, in practise, the functional spatial correlation is incorporated into the estimator's computation by introducing a neighborhood set given by the following equations for each site ℓ:

$$U_\ell = \{\mathbf{i}, \text{suchthat} \varphi(\| \mathbf{i} - \ell \|) \leqslant \| \mathbf{i} - \ell \|^{-\ell}\}.$$

Since we have limited ourselves to the isotropic case in which the spatial dependency is just a function of the distance between locations, we can proceed with the vicinity set defined by:

$$U_\ell = \{\mathbf{i}, \| \mathbf{i} - \mathbf{k} \| \leqslant \xi_{\mathbf{n}}\},$$

where $\xi_{\mathbf{n}}$ is an appropriate sequence of positive real numbers. The quantity $\xi_{\mathbf{n}}$ is optimally selected over the nearest neighbors locations with respect to the Euclidean norm on the coordinates. For more details on the choice of U_ℓ, see [22]. We estimate, for each fixed site ℓ, the kNN conditional hazard function by: $\forall b \in \mathbb{R}$

$$\hat{h}_k^{A_\ell}(b) = \frac{h_J^{-1} \sum_{\mathbf{i},\mathbf{j} \in I_{\mathbf{n}}} L\left(P_k^{-1} d(a, A_\ell)\right) J'\left(\frac{b - B_{\mathbf{i}}}{h_J}\right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}{\sum_{\mathbf{i},\mathbf{j} \in I_{\mathbf{n}}} L\left(P_k^{-1} d(a, A_\ell)\right) J\left(\frac{b - B_{\mathbf{i}}}{h_J}\right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}. \tag{15}$$

where, $\mathbb{I}_{U_\ell}$ is the indicator function of the set U_ℓ.

Second, we need to select a suitable semi-metric $d(.,.)$, kernel $L(.)$. Then, we choose the asymmetrical quadratic kernel defined as $L(u) = \frac{3}{4}(1 - u^2)1_{[0,1]}(u)$. Meanwhile, because of the smoothness of curves $A_i(t)$, we consider the following semimetric based on the first derivative:

$$d^{deriv}(A_i, A_j) = \sqrt{\int_0^{\pi/2} \left(A_i'(t) - A_j'(t)\right)^2 \mathrm{d}t}, \quad \forall A_i, A_j \in \mathcal{G}.$$

Recall that, the classical estimator of the conditional hazard is defined as follows: $\forall b \in \mathbb{R}$

$$\hat{h}_{\mathbf{n}}^{A_\ell}(b) = \frac{h_J^{-1} \sum_{\mathbf{i},\mathbf{j} \in I_{\mathbf{n}}} L\left(h_L^{-1} d(a, A_\ell)\right) J'\left(\frac{b - B_{\mathbf{i}}}{h_J}\right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}{\sum_{\mathbf{i},\mathbf{j} \in I_{\mathbf{n}}} L\left(h_L^{-1} d(a, A_\ell)\right) J\left(\frac{b - B_{\mathbf{i}}}{h_J}\right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}. \tag{16}$$

The efficiency of the predictors is evaluated by the empirical Mean Squared Error(MSE)

$$MSE\left(\hat{h}_k^{A_\ell}(b)\right) = \mathbf{n}^{-1} \sum_{i=1}^{\mathbf{n}} \left(h_k^{A_{\mathbf{i}}}(B_{\mathbf{i}}) - \hat{h}_k^{A_{\mathbf{i}}}(B_{\mathbf{i}})\right)^2,$$

$$MSE\left(\hat{h}_{\mathbf{n}}^{A_\ell}(b)\right) = \mathbf{n}^{-1} \sum_{i=1}^{\mathbf{n}} \left(h_{\mathbf{n}}^{A_{\mathbf{i}}}(B_{\mathbf{i}}) - \hat{h}_{\mathbf{n}}^{A_{\mathbf{i}}}(B_{\mathbf{i}})\right)^2,$$

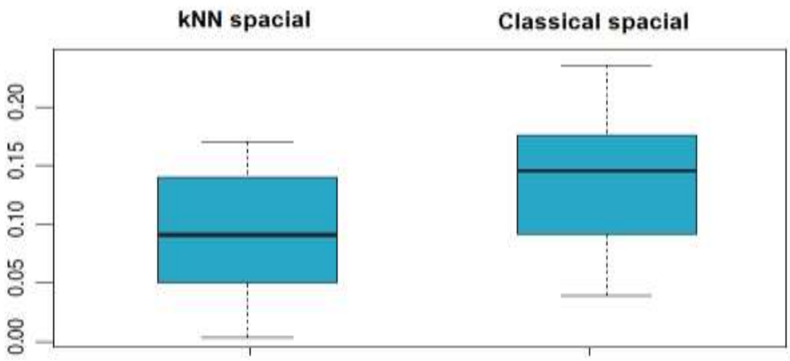

Figure 1: The AE-errors of both models

We use the cross-validation procedure proposed by [66] for which the bandwidth h_L is chosen via the following rule

$$h_{opt} = \underset{h_L}{\arg\min} CV_1(h_L) = \underset{h_L}{\arg\min} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} |B_{\mathbf{i}} - \hat{r}^{(-\mathbf{i})}(A_{\mathbf{i}})|,$$

where $\hat{r}^{(-\mathbf{i})}$ is the leave-one-out-curve estimator of $\left(\hat{h}_{\mathbf{n}}^{A_\ell}\right)$. Concretely, for the robust kNN we select $k_{opt} = \arg\min_k CV_2(k)$,

where $CV_2(k) = \arg\min_k \sum_{\mathbf{i}\in I_{\mathbf{n}}} |B_{\mathbf{i}} - \hat{q}^{(-\mathbf{i})}(A_{\mathbf{i}})|$,

with $\hat{q}^{(-\mathbf{i})}$ is the leave-one-out-curve estimator of $\left(\hat{h}_k^{A_\ell}\right)$.

The box-plot of the MSE of both models is given in Figure 1. We observe that there is no meaningful difference between this spatial predictors. The two predictors are basically equivalent and both show the good behavior. Now, in order to investigate the features of our approach, we introduced some artificial outliers by multiplying 15% values of Y by 10 . We box-plot AE-errors of both models.

Further we see from Figure 2 that the kNN hazard function error is much more better than the classical one in this case. Moreover, looking at both figures, it appears clearly the MSE of the kernel hazard function model has dramatically changed compared to the kNN hazard case. This statement confirms that the kNN hazard function is more robust than the classical regression.

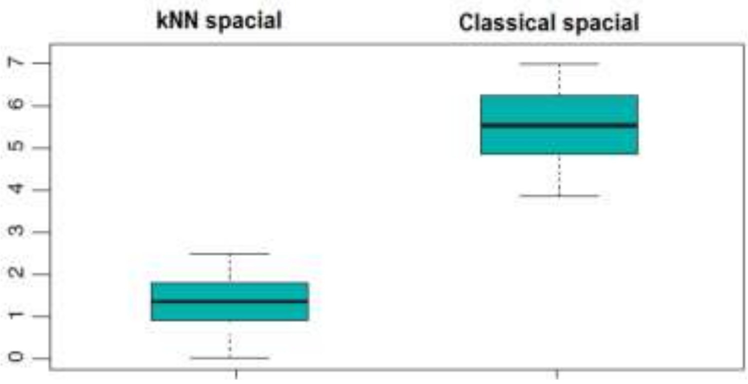

Figure 2: The MSE in presence of outliers

5.5 Real data application

This section's major goal is to apply the theoretical results from the previous section to real data. In particular, we analyze the effectiveness of hazard function in the context of kNN spatial functional prediction via classical hazard function.

In this real data example, we're looking for a way to estimate the logarithm of total precipitation based on the monthly maximum temperature curve. The functional predictor A_i is the curve of the monthly maximum temperatures in the ith climatic station (specified by its geographic coordinates) in a period Υ, and B_i is the logarithm of the total precipitations in the same station and period, according to the notations in the previous section.

For this application, we used monthly temperature and precipitation data from 125 stations collected during 2000 and 2010. These observations can be found at the following url: ftp://ftp.ncdc.noaa.gov/pub/data/ushcn/v2/monthly. In Figure 3 The functional covariates are given.

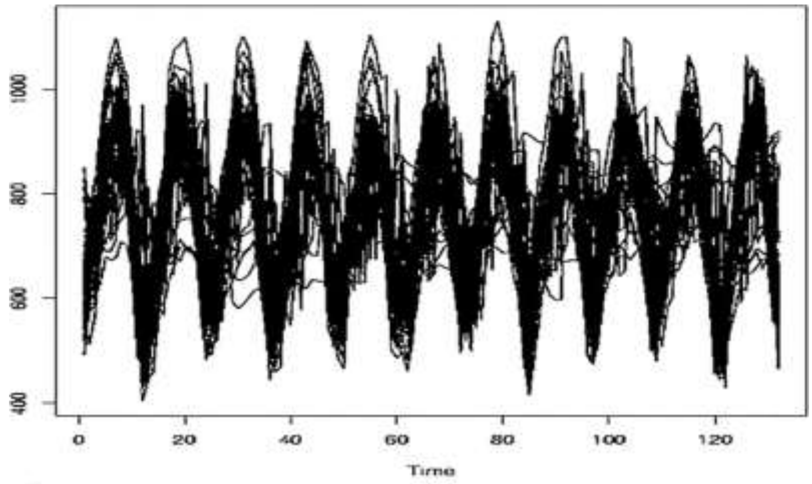

Figure 3: Highest monthly temperatures in 125 weather sites in the United States

In order to evaluate the presented estimator's efficiency, represented by $\hat{h}_{\mathbf{n}}^{A_\ell}(b)$ in Equation (15) and to compared it with the one that does not directly consider distance between positions and is represented by $\hat{h}_{\mathbf{n}}^{A_{\mathbf{k}}}(b)$ in Equation (16), we randomly devide our data $(X_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}}$ into two different subsets:

- learning sample $(X_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}\in I}$ (114 stations),
- test sample $(X_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}\in I'}$ (11 stations).

To calculate the estimator for both methods, we take the same quadratic kernel function and the semi-metric $d^{\text{deriv}}(.,.)$ in the simulation study, Similar to the CV-procedures in the previous section of simulation study we select the parameters k_{opt} (for the kNN) and h_{opt} (for the kernel).

As an accuracy measure, we employ the Mean Square Error (MSE) procedure, which is described as follows:

$$MSE\left(\hat{h}_k^{A_\ell}(b)\right) = \frac{1}{11}\sum_{\mathbf{i}\in I'} \left(h_k^{A_{\mathbf{i}}}(B_{\mathbf{i}}) - \hat{h}_k^{A_{\mathbf{i}}}(B_{\mathbf{i}})\right)^2$$

and

$$MSE\left(\hat{h}_{\mathbf{n}}^{A_\ell}(b)\right) = \frac{1}{11}\sum_{\mathbf{i}\in I'} \left(h_{\mathbf{n}}^{A_{\mathbf{i}}}(B_{\mathbf{i}}) - \hat{h}_{\mathbf{n}}^{A_{\mathbf{i}}}(B_{\mathbf{i}})\right)^2.$$

The results of both methods are plotted against the true values in Figure 4, where the predicted values are plotted against the true values.

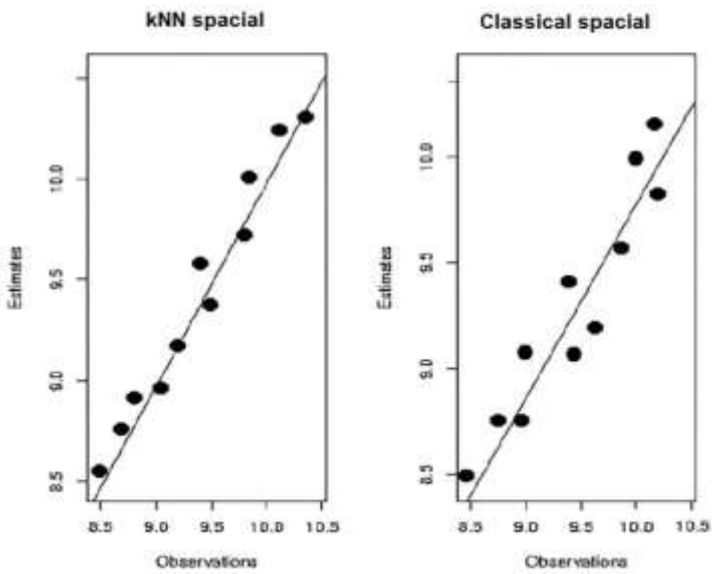

Figure 4: Comparison of the prediction results between the two methods.

The left part of Figure 4 illustrates the case in which the kNN spatial correlation is used, whereas the right half depicts the case in which the classical spatial correlation is used.

The continuous line controls the prediction's effectiveness, in the sense that the efficiency of the prediction method is measured by how close the plotted points are to the continuous line. It is apparent that estimate using kNN spatial correlation is substantially better and more efficient than traditional estimation. The mean squared error obtained in the two situations, respectively, proves this, $MSE\left(\hat{h}_k^{A_\mathbf{k}}(b)\right) = 0.21$ whereas, $MSE\left(\hat{h}_\mathbf{n}^{A_\ell}(b)\right) = 0.36$.

6 Conclusion

This work provides a extension of the study of study of [41] by the kNN method. They are compared to the standard kernel estimator. We showed that the proposed estimator can give better results compared to the kernel estimator in terms of estimation error.

However, the main advantage of this study is that it is considerably faster than the classical one when the outlier data are present.

In conclusion, the proposed estimators allowed us to obtain good results. This is confirmed by the MSE result.

References

[1] Almanjahie I, Aissiri K, Laksaci A, Chiker el mezouar Z. The k nearest neighbors smoothing of the relative-error regression with functional regressor. Communications in Statistics - Theory and Methods. 2020; 356(10):1-14.

[2] Attouch M, Bouabsa W, Chiker el mozoaur Z. The k-nearest neighbors estimation of the conditional mode for functional data under dependency. International Journal of Statistics & Economics. 2018; 19(1):48-60.

[3] Attouch M, Bouabsa W. The k-nearest neighbors estimation of the conditional mode for functional data. Rev. Roumaine Math. Pures Appl. 2013; 58(4):393-415.

[4] Attouch M, Belabed F. The k nearest neighbors estimation of the conditional hazard function for functional data. REVSTAT. 2014; 12:273–97.

[5] Attouch M, Laksaci A, Messabihi N. Nonparametric relative error regression for spatial random variables. Statistical papers. 2015; 58(4):987-1008.

[6] Beirlant J, Berlinet A, Biau G. Higher order estimation at Lebesgue points. Ann. Inst. Statist. Math. 2018; 90(60):651–677.

[7] Biau G, Cadre B. Nonparametric spatial prediction. Stat. Inference Stoch. Process. 2004; 7:327-349.

[8] Bosq D. Linear Processes in Function Spaces: Theory and Applications. Lecture Notes in Statistics. Springer. 2000; 149.

[9] Bouabsa W. Nonparametric Relative Error Estimation via Functional Regressor by the k Nearest Neighbors Smoothing Under Truncation Random. Applications and Applied Mathematics: An International Journal (AAM). 2021; 16(1): 97-116.

[10] Carbon M, Tran L T, Wu B. Kernel density estimation for random fields. Statist. Probab. Lett. 1997; 36:115-125.

[11] Burba F, Ferraty F, Vieu P. k-nearest neighbour method in functional nonparametric regression. J. Nonparametr. Stat. 2009; 21(4):453-469.

[12] Carbon M, Francq C, Tran L T. Kernel regression estimation for random fields. J. Statist. Plann. Inference. 2007; 137:778-798.

[13] Cover T M. Estimation by the nearest neighbor rule. IEEE Trans. Inform. Theory. 1968; 7(14):50-55.

[14] Collomb G. Estimation non paramétrique de la régression: Revue bibliographique. Internat. Statist. Rev. 1981; 18(49):75-93.

[15] Cressie N A. Statistics for Spatial Data. Wiley, New York. 1993.

[16] Daoudi H, Mechab B, Chikr Elmezouar Z. Asymptotic normality of a conditional hazard function estimate in the single index for quasi-associated data. 2018; 4(22):513-530.

[17] Dabo-Niang S, Yao A F. Kernel regression estimation for continuous spatial processes, Math. Methods Statist. 2007; 16(15):298-317.

[18] Devroye L, Györfi L, Krzyzak A, G.Lugosi. On the strong universal consistency of nearest neighbor regression function estimates. Ann. Statist. 1994; 20(22):1371-1385.

[19] Devroye L, Wagner T. The strong uniform consistency of nearest neighbor density. Ann. Statist. 1977; 10(5):536–540.

[20] Devroye L, Wagner T. Nearest neighbor methods in discrimination. In Classification, pattern recognition and reduction of dimensionality. Handbook of Statistics. North-Holland, Amsterdam. 1982; 2.

[21] Dabo-Niang S, Thiam B. Robust quantile estimation and prediction for spatial processes. Stat Probab Lett. 2010; 80:1447–1458.

[22] Dabo-Niang S, Yao A F. Kernel regression estimation for continuous spatial processes. Mathematical Methods of Statistics. 2007; 16(4):298–317.

[23] Dabo-Niang S, Ould-Abdi S, Ould-Abdi A, Diop A. Consistency of a nonparametric conditional mode estimator for random fields. Stat Methods Appl. 2014; 16(23):1–39.

[24] Dony J, Einmahl U. Uniform in Bandwidth Consistency of Kernel Regression Estimators at a Fixed Point. High Dimensional Probability. 2009; 53(5):308-325.

[25] Einmahl U, Mason D. Uniform in bandwidth consistency of kernel-type function estimators. The Annals of Statistics. 2005; 33(3):1380–1403.

[26] El Machkouri M, Stoica R. Asymptotic normality of kernel estimates in a regression model for random fields. J. Nonparametric Stat. 2010; 22(15):955–971.

[27] Estévez-Pérez G, Quintela-del-Ro A, Vieu P. Convergence rate for crossvalidatory bandwidth in kernel hazard estimation from dependent samples. J. Statist. Plann. Inference. 2002; 16(16):1-30.

[28] Ferraty F, Peuch A, Vieu P. Modèle à indice fonctionnel simple. C. R. A. S.Mathématiques Paris. 2003; 336:1025-1028.

[29] Ferraty F, Vieu P. Nonparametric functional data analysis. Theory and practice. New York: Springer-Verlag. 2006.

[30] Ferraty F, Laksaci A, Vieu P. Estimating some characteristics of the conditional distribution in nonparametric functional models. Stat. Inference Stoch. Process. 2006; 9(20):47–76.

[31] Ferraty F, Rabhi A, Vieu P. Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. Rev. Roumaine Math. Pures Appl. 2008; 53(15):1–18.

[32] Ferraty F, Tadj A, Laksaci A, Vieu P. Rate of uniform consistency for nonparametric estimates with functional variables. J. Statist. Plann. Inference. 2010; 140(15):335-352.

[33] Gagui A, Chouaf A. On the nonparametric estimation of the conditional hazard estimator in a single functional index. STATISTICS IN TRANSITION new series. 23(2):89–105.

[34] Gheriballah A, Laksaci A, Rouane R. Robust nonparametric estimation for spatial regression. J Stat Plan Inference. 2010; 140:1656–1670

[35] Guyon X. Estimation d'un champ par pseudo-vraisemblance conditionnelle: Etude asymptotique et application au cas Markovien. Proceedings of the Sixth Franco-Belgian Meeting of Statisticians. 1987.

[36] Györfi L, Kohler A, Krzyzak A, Walk H. A Distribution-Free theory of Nonparametric Regression. Springer. New York. 2002.

[37] Hallin M, Lu Z, Yu K. Local linear spatial quantile regression. Bernoulli. 2009; 15:659–686.

[38] Härdle W, Marron J S. Optimal Bandwidth Selection in Nonparametric Regression Function Estimation. Annals of Statistics. 1985; 13(15):1465-1481.

[39] Kara Z, Laksaci A, Vieu P. Data-driven kNN estimation in nonparametric functional data analysis. Journal of Multivariate Analysis. 2017; 153(85):176–188.

[40] Kudraszow N L, Vieu P. Uniform Consistency of kNN Regressors for Functional Variables. Statistics and Probability Letters. 2013; 83(153):1863–1870.

[41] Laksaci A, Mechab B. Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionelle: cas des données spaciales. REV. ROUMAINE MATH. PURES APPL. 2010; 55(1):35–51.

[42] Laksaci A, Mechab B. Conditional hazard estimate for functional random fields. Journal of Statistical -Theory and Practice. 2014; 8: 192-200.

[43] Laloë T. A k-nearest neighbor approach for functional regression. Statistics & Probability Letters. 2008; 78(10): 1189-1193.

[44] Lecoutre J P, Ould-Said E. Estimation de la densité et de la fonction de hasard conditionnelle pour un processus fortement mélangeant avec censure. C.R. Math. Acad. Sci. Paris. 1992; 314(15):295-300.

[45] Li J P. Strong convergence rates of error probability estimation in the nearest neighbor discrimination rule. J. Math. 1985; 15(5): 113-118.

[46] Li J, Tran L T. Hazard rate estimation on random fields. J. Multivariate Anal. 2007; 98:1337-1355.

[47] Li J, Tran L T. Nonparametric estimation of conditional expectation. J. Statist. Plann. Inference. 2009; 139(36): 164-175.

[48] Li J, Tran L T. Nonparametric estimation of conditional expectation. J Stat Plan Inference. 2009; 139(35):164–175

[49] Lian H. Convergence of Functional k-Nearest Neighbor Regression Estimate with Functional Responses. Electronic Journal of Statistics. 2011; 5(133):31-40.

[50] Lu Z, Chen X. Spatial kernel regression: weak consistency. Stat Probab Lett. 2004; (68)30:125–136.

[51] Massim I, Mechab B. Local linear estimation of the conditional hazard function. International Journal of Statistics and Economics. 2016; 17:1-11.

[52] Masry E. Recursive probability density estimation for weakly dependent stationary processes. IEEE Trans. Inform. Theory. 1986; (32)16:254-267.

[53] Moore D, Yackel J. Consistency properties of nearest neighbor density function estimators. Ann. Statist. 1977; (20)5:143-154.

[54] Quintela-del-Ro A. Nonparametric estimation of the maximum hazard under dependence conditions. Statist. Probab. Lett. 2006; (76)15, 1117-1124.

[55] Quintela-del-Ro A. Hazard function given a functional variable: Non-parametric estimation under strong mixing conditions. J. Nonparametr. Stat. 2008; (20)15, 413-430.

[56] Rabhi A, Benaissa S, Hamel E, Mechab B. Mean square error of the estimator of the conditional hazard function. Appl. Math. (Warsaw). 2013; (40)4:405-420.

[57] Ramsay J O, Silverman B W. Applied Functional Data Analysis; Methods and Case Studies. Springer-Verlag. New York. 2002.

[58] Robinson P M. Asymptotic theory for nonparametric regression with spatial data. J Econom. 2011; 165(15): 5–19.

[59] Roussas G. Hazard rate estimation under dependence conditions. J. Statist. Plann. Inference. 1989; 22(15): 81-93.

[60] Tran L T. Kernel density estimation on random fields. J. Multivariate Anal. 1990; (34)25:37-53.

[61] Tabti H, Ait Saïdi A. Estimation and simulation of conditional hazard function in the quasi-associated framework when the observations are linked via a functional single-index structure. Commu. Stat. Theory and Methods. 2018; 47: 816-838.

[62] Torkia M, Mechab Massim I. Quadratic error of the conditional hazard function in the local linear estimation for functional data. Afr. Stat. (3)13, 1759-1777.

[63] Tran T, Wehrens R, Buydens L. kNN-kernel density-based clustering for high-dimensional multivariate data. Comput. Statist. Data Anal. 2006; (51)81:513-525.

[64] Tran L T. Kernel density estimation on random fields. Journal of Multivariate Analysis. 1990; (34)30:37–53.

[65] Watson G S, Leadbetter M R. Hazard analysis. Sankhyia. 1964; (26)25:101-116.

[66] Xu. R, Wang. J . L^1- estimation for spatial nonparametric regression. J Nonparametric Stat. 2008; 20:523–537.