Robust equivariant nonparametric regression estimators for strongly mixing data using a k nearest neighbour approach

Somia Guenani [1], Wahiba Bouabsa [2] and Mohammed Kadi Attouch [3]

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

email: (guenanisomia@gmail.com (S. Guenani), wahiba_bouab@yahoo.fr (W. Bouabsa), attou_kadi@yahoo.fr (M. K. Attouch)).

Article History: Do Received: 20 October 2022; Revised: 10 November 2022; Accepted: 15 January 2023; Published online: 8 March 2023.

Abstract: We discuss in this paper the robust equivariant nonparametric regression estimators for strong mixing data with the k Nearest Neighbour (kNN) method. We consider a new robust regression estimator when the scale parameter is unknown. The principal aim is to prove the almost complete convergence (with rate) for the proposed estimator. Furthermore, a comparison study based on simulated data is also provided to illustrate the finite sample performances and the usefulness of the kNN approach.

Keywords: Functional data, ergodic data, kNN estimation, kernel estimate, uniform almost complete convergence rate, entropy.

1. Introduction

It is very well recognized that robust regression in statistics is an attractive research method. It is used to overcome some of the weaknesses of classical regression, namely when outliers contain heteroscedastic data.

The study of the connection between a random variable W and a set of covariates Z is a common problem in statistics. In the literature, these variables are generally known as functional variables. Remember that the robust method is an old statistical issue, this latter was investigated first by [40] who studied an estimation of allocation parameter (see also [23, 45]), for some results containing the multivariate time series case under a mixing or an ergodic condition).

The robust model is an essential alternative regression model that allows overcoming many drawbacks of the classical regression, such as the sensitivity to the outliers or the heteroscedasticity phenomena. Indeed, it was initially proposed by [9] who demonstrated the model’s almost-complete convergence in the independent and identically distributed (i.i.d.) case. Several results on nonparametric robust functional regression have been obtained since this study (for example, [25, 19, 4, 5, 6, 38, 15] and references therein).

Furthermore, it is well known that the kNN method is better than the classical kernel method. Pushed by its attractive features, the functional kNN smoothing approach has received a growing consideration in the last years. The study of [39] is a thorough analysis of kNN estimators in the finite dimensional context. Work in this area was started by [24], and a large number of articles are now available in various estimating contexts, which including regression, discrimination, density and mode estimation, and clustering analysis, we make reference to [22], [28], [47], [52], [27], [29], [10], [44], [18], [59], [48], [7], [8], [43] and [41], [2], [17] for the most recent advances and references. Note that, such a study has a great impact on practice. 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.

All the results involved in the functional kNN estimation above were obtained under i.i.d. case. While in many practical applications, some problems require taking into account the dependence structure that may exist within the dataset. The strong mixing dependence or α-mixing is one of the most general weak dependence modelization in the literature. The research of Nadaraya Watson (NW) kernel method for this dependent functional data analysis has been widely carried out, see for instance, [16, 32] and the bibliographical surveys by

[37] and [49]. However, for the kNN approach, the only paper is, as far as we know, by [53] which studies the kNN estimator under α-mixing sample and states its pointwise almost complete convergence (with rates).

The strong mixing dependence or α-mixing is one of the most general weak dependence modelization in the literature. The research of NW kernel method for this dependent functional data analysis has been widely carried out, see for instance, [16], [32] and the bibliographical surveys by [37] and [49]. However, for the kNN approach, as far as we know, by [53] and [43] which studies the kNN estimator under α-mixing sample and states its pointwise almost complete convergence (with rates). Inspired by all the results above, the purpose of this paper show us that functional kNN approach can be used to further investigate the estimation of functional nonparametric regression opera in the case of strong mixing datasets with unknown scale parameter. This is motivated by the fact that the robust regression 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.

In functional data analysis (FDA), kNN nonparametric robust equivariant regression estimators for strong mixing data is new. This researches's primary goal is to provide generalizations, to the kNN case, the results obtained by [43] in alpha mixing dependency case with the research of [50] and [2]. More precisely, we establish the almost complete convergence with rates of an estimator constructed by combining the ideas of robustness with those of smoothed regression. We point out that the main feature of our approach is to develop an alternative prediction model to the classical regression that is not sensitive to outliers or heteroscedastic data, taking into account the local data structure. The work has not yet been addressed in the literature. We wish that this will be useful to readers who are interested in learning about and comprehending the core idea of functional kNN methods with strong mixing dependence sample and with unknown scale parameter.

This paper's structure is as follows. In Section 2, we find some fundamental concepts and various assumptions. Then in Section 3 we give some technical tools as well as their proofs. The main result is given in Section 4, then we provides all the proofs of the main result in Section . Finally, simulation study is given in Section 6.

2. Principal hypotheses and basic definitions

2.1 Kolmogorov's entropy

The aim of this subsection is to emphasize the topological aspects of our study. Indeed, all asymptotic conclusions in nonparametric statistics for functional variables are intimately connected to the concentration properties of the probability measure of the functional variable Z, as [32] indicated. We must also consider the element of uniformity in this situation. Let $\varepsilon > 0$ be given, and let $\mathcal{T}$ be a subset of a semi-metric space $\mathcal{F}$, a limited set of points $z_1, z_2, \ldots, z_n$ in $\mathcal{F}$ is known as an ε-net for $\mathcal{T}$ if $\mathcal{T} \subset \bigcup_{\ell=1}^{n} B(z_\ell, \varepsilon)$. Kolmogorov's ε-entropy of the set $\mathcal{T}$ is defined as $\Psi_{\mathcal{T}_\mathcal{F}}(\varepsilon) = \log(N_\varepsilon(\mathcal{T}_\mathcal{F}))$, where $N_\varepsilon(\mathcal{T}_\mathcal{F})$ is the minimal number of open balls in $\mathcal{F}$ with radius ε required to cover $\mathcal{F}$. This concept was introduced by [42] and it represents a measure of the complexity of a set, in sense that, high entropy means that much information is needed to describe an element with an accuracy ε. Therefore, the choice of the topological structure (with other words, the choice of the semi-metric) will play a crucial role when one is looking at uniform (over $\mathcal{T}$) asymptotic results. More precisely, a good semi-metric can increase the concentration of the probability measure of the functional variable Z as well as minimize the ε-entropy of the subset $\mathcal{T}_\mathcal{F}$.

To formulate our main result, some additional notations are necessary. Let $z_1, \ldots, z_{N_\varepsilon(\mathcal{T}_\mathcal{F})}$ be an ε-net of $\mathcal{T}_\mathcal{F}$ and for all $z \in \mathcal{T}_\mathcal{F}$, $\ell(x) = \operatorname{argmin}_{\ell \in \{1,2,\ldots,N_\varepsilon(\mathcal{T}_\mathcal{F})\}} d(z, z_\ell)$,

$$\eta_{n,1}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} |cov(\Gamma_z(\frac{W_i - x}{\hat{t}(z)})u_i, \Gamma_z(\frac{W_j - x}{\hat{t}(z)})u_j)|,\ \eta_{n,2}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} |cov(u_i', u_j')|,$$

$$\eta_{n,3}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} |cov(u_i, u_j)|,\ \eta_{n,4}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} |cov(u_i'', u_j'')|,$$

where

$u_i = \mathbb{I}_{B(z_{\ell(z)}, Ch_L + \varepsilon)}(z_i)$ for some $C > 0$ and $0 < h_L \to 0$,

$$u_i' = \frac{\Gamma_z(\frac{W_i - x}{\hat{t}(z)}) L(h_L^{-1}(z_\ell) d(z_\ell, z_i))}{\mathbb{E}L(h_L^{-1}(z_\ell) d(z_\ell, z_1))} - \frac{\mathbb{E}(\Gamma_z(\frac{W_i - x}{\hat{t}(z)}) L(h_L^{-1}(z_\ell) d(z_\ell, z_i)))}{\mathbb{E}L(h_L^{-1}(z_\ell) d(z_\ell, z_1))},$$

$$u_i'' = \frac{L(h_L^{-1}(z_\ell)d(z_\ell,z_i))}{\mathbb{E}L(h_L^{-1}(z_\ell)d(z_\ell,z_1))} - \frac{\mathbb{E}L(h_L^{-1}(z_\ell)d(z_\ell,z_i))}{\mathbb{E}L(h_L^{-1}(z_\ell)d(z_\ell,z_1))}.$$

Indicate

$$\eta_n^2 = \max\{\eta_{n,1}^2, \eta_{n,2}^2, \eta_{n,3}^2, \eta_{n,4}^2\}. \tag{1}$$

2.2 kNN Regression function model

Let $X_i = (Z_i, W_i)_{i=1,\dots,n}$ a strong mixing sample drawn from the pair (Z, W) and is defined in $\mathcal{F} \times \mathbb{R}$. We do not assume the existence of a density for the functional random variable Z since $(\mathcal{F}, d)$ is a semi-metric space, $\mathcal{F}$ is not necessarily of finite dimension. The functional nonparametric regression is defined as

$$W = r(Z) + \varepsilon \text{ with } \mathbb{E}[\varepsilon|Z] = 0.$$

The kNN kernel estimator can be written as for a fixed $z \in E$

$$\hat{r}_{kNN}(z) = \frac{\sum_{i=1}^{n} W_i L(T_{n,k}(z)^{-1} d(z,Z_i))}{\sum_{i=1}^{n} L(T_{n,k}(z)^{-1} d(z,Z_i))}, \tag{2}$$

where L is an asymmetrical kernel and $T_{n,k}(z)$ is defined as follows

$$T_{n,k}(z) = min\left\{h_L \in \mathbb{R}^+ / \sum_{i=1}^{n} \mathbb{I}_{B(z,h_L)}(z_i) = k\right\}.$$

The functional version of the NW kernel type estimator of the nonparametric functional regression is as follows

$$\hat{r}(z) = \frac{\sum_{i=1}^{n} W_i L(h_L^{-1} d(z,Z_i))}{\sum_{i=1}^{n} L(h_L^{-1} d(z,Z_i))}, \tag{3}$$

where $z \in E$ is fixed, and h_L denotes a non-random bandwidth.

2.3 kNN Conditional cumulative distribution function

The conditional cumulative distribution function of W given $\mathrm{Z} = z$, for each $z \in \mathcal{F}$ and for any $w \in \mathbb{R}$ can be written as

$$F(w|\mathrm{Z} = z) = F_{\mathrm{Z}}^{\mathrm{W}}(z, w) = \mathbb{P}(\mathrm{W} \le w|\mathrm{Z} = z) = \mathbb{E}[\mathbb{I}_{(-\infty,w]}(\mathrm{W})|\mathrm{Z} = z].$$

We call the following function the estimator of $F(w|\mathrm{Z} = z)$

$$\hat{F}(w|\mathrm{Z} = z) = \sum_{i=1}^{n} L(d(z,Z_i)/T_{n,k}(z))\left(\sum_{i=1}^{n} L(d(z,Z_i)/T_{n,k}(z))\right)^{-1} \mathbb{I}_{(-\infty,w]}(\mathrm{W_i}). \tag{4}$$

Several authors have studied the estimation of the conditional cumulative distribution function in the real case (see for example [57] and [58]). Then, In the functional case [32] proved the almost complete convergence of a double kernel estimator of the conditional cumulative distribution function.

2.4 The kNN robust equivariant estimators and their functional relatives function

In this section we define the function of our main problem, we consider estimating a generalized regression function defined as follows

$$\theta(z, x, \tau(z)) = \mathbb{E}[\Gamma_z(\frac{W_i - x}{\tau(z)})/Z_i = z], \tag{5}$$

where Γ_z is a real-valued function, we denoted by $\vartheta(z)$ the unique solution of

$\theta(z) = 0$, where $t(z)$ is a robust measure of the conditional scale. The unique solution of (5) is the so-called robust conditional location functional where Γ_z is a strictly increasing function (see [12]). The conditional scale measure is defined as the conditional median of the absolute deviation from the conditional median, that is,

$$t(z) = med(|W - M(z)|/Z = z) = mad_c(F_W^z(.)), \tag{6}$$

with $M(z) = med(W/Z = z)$ is the median of the conditional distribution.

On the other hand, we note that $t(z)$ which is a robust measure of the conditional scale, always equals $\tau(z)$.

We insert an estimator of $F_W^z(z)$ into (4) to get $\vartheta(z)$ estimators, wich will betaken as

$\hat{F}(w|Z=z)$. A robust estimator of the conditional scale is denoted by $\hat{t}(z)$, for example, $\hat{t}(z) = mad_c(\hat{F}(.|Z=z)$, the scale measure given in (6) measured in $\hat{F}(w|Z=z)$. The solution $\hat{\vartheta}(z)$ of $\hat{\theta}(z)=0$ gives the robust nonparametric estimator of $\vartheta(z)$ in this notation, where

$$\hat{\theta}(z) = \frac{\sum_{i=1}^{n} L(d(z,Z_i)/h_L)\Gamma_z(\frac{W_i-x}{\hat{t}(z)})}{\sum_{i=1}^{n} L(d(z,Z_i)/h_L)}. \tag{7}$$

Hence the kNN estimator of $\theta(.)$ is written as

$$\hat{\theta}_{kNN}(z) = \frac{\sum_{i=1}^{n} L(d(z,Z_i)/T_{n,k}(z))\Gamma_z(\frac{W_i-x}{\hat{t}(z)})}{\sum_{i=1}^{n} L(d(z,Z_i)/T_{n,k}(z))}. \tag{8}$$

2.5 Hypotheses

In this part, we propose the following hypotheses to establish the uniform almost complete convergence of $\hat{\theta}$ on some subset $\mathcal{T}_{\mathcal{F}}$ of $\mathcal{F}$. To do that we denote by C and C' some real generic constants supposed strictly positive and we suppose that:

-(A1) The processes (Z_i, W_i) satisfies

$\forall h_L > 0, \quad \varphi_z(h_L) =: \mathbb{P}(Z \in B(z,h_L)) > 0$ where $\varphi_z(.)$ is continuous in the neighborhood of 0 and $\varphi_z(0) = 0$.

-(A2) $\exists$ function $\phi(.) \geq 0$, a bounded function $f(.) > 0$, $\alpha > 0$ and $\rho > 0$ such that

- $\phi(0) = 0$ and $\lim_{\varepsilon\to\infty}\phi(\varepsilon) = 0$,
- - $\lim_{\varepsilon\to\infty}(\phi(u\varepsilon)/\phi(\varepsilon)) = u^{\alpha}$ with $u > 0$,
- - $\sup_{z\in\mathcal{T}_{\mathcal{F}}}|\varphi_z(\varepsilon)/\phi(\varepsilon) - f(z)| = O(\varepsilon^{\rho})$, as $\varepsilon \to 0$.

-(A3) The kernel $L(.)$ is defined by

- (A3a) is a nonnegative function with support [0,1] such that
  $0 < C\mathbb{I}_{[0,1]}(t) < L(t) < C'\mathbb{I}_{[0,1]}(t) < +\infty$.
- (A3b) - its derivative $L'(.)$ exists on the same support and $-\infty < C < L'(t) < C' < 0$.

-(A4) $\forall (x_1,x_2) \in [\vartheta(z)-\delta, \vartheta(z)+\delta] \times [\vartheta(z)-\delta, \vartheta(z)+\delta], \forall (z_1,z_2) \in \mathcal{T}_{\mathcal{F}}$,

$|\theta(z_1,x,t(z)) - \theta(z_2,x,t(z))| \leq Cd^{\beta_1}(z_1,z_2), \beta_1 > 0$.

-(A5) For each fixed $x \in [\vartheta(z)-\delta, \vartheta(z)+\delta]$, $\forall m \geq 2$,

$\mathbb{E}[\Gamma_z^m(\frac{W_i-x}{t(z)})/Z_i] < \delta(z) < Cm! < \infty$, with $\delta(.)$ continuous on $\mathcal{T}_{\mathcal{F}}$.

-(A6) The functions φ_z and $\Psi_{\mathcal{T}_{\mathcal{F}}}$ are such that

$\exists C > 0,\ \exists \eta_0 > 0,\ \forall \eta < \eta_0, {\varphi_z}'(\eta) < C$ and if $L(1) = 0$ we can seen that

$\exists C > 0,\ \exists \eta_0 > 0,\ \forall 0 < \eta < \eta_0, \int_0^{\eta} \varphi(u)du > C\eta\varphi_z(\eta)$,

if in addition $k/n \to 0$ as $n \to \infty$, $\log^2 n/k < \Psi_{\mathcal{T}_{\mathcal{F}}}(\log n/n) < k/\log n$ and

$0 < C < k/\log n < C' < \infty$ for n large enough.

-(A7) Kolmogorov's ε-entropy of the set $\mathcal{T}_{\mathcal{F}}$ satisfies, for some $\varpi > 1$

$$\sum_{n=1}^{\infty} \exp\{(1-\varpi)\Psi_{\mathcal{T}_{\mathcal{F}}}(\log(n)/n)\} < \infty.$$

-(A8) Consider that $\mathcal{T}_{\mathcal{F}}$ is a compact set of $\mathcal{F}$ such that

- (A8a) The function $F(w|Z=z)$ is uniformly continuous of z in a neighborhood of $\mathcal{T}_{\mathcal{F}}$ for each z fixed.
- (A8b) The equicontinuity condition that follows hold
  $\forall \varepsilon > 0,\ \exists \delta > 0: |\lambda - \nu| < \delta \Longrightarrow \sup_{z\in\mathcal{T}_{\mathcal{F}}}|F(\lambda|Z=z) - F(\nu|Z=z)| < \varepsilon$.

-(A9) $\exists a > 1$ and $p > 2$ such that for some θ large enough $\eta_n^{-((a+1)p/(a+p))} = o(n^{-\theta})$.

Comments on the hypotheses

Our hypotheses are quite light in the context of nonparametric statistics in functional time series.

The latter is used an accordance with assumption (A1) which is less strict than the conditions imposed by [46] because the concentration function $\mathbb{P}(Z_i \in B(z,r))$ and the conditional concentration function $\mathbb{P}(Z_i \in B(z,r)/\mathcal{F}_{i-1})$ do not need to be written as products of two independent nonnegative functions of the center and radius. Assumption(A3) contains two types of kernels which have been utilized in practice box and continuous kernels. Assumptions (A2), (A4) and (A5) are the usual conditions in the nonparametric setting.

About assumption (A6) we can say because the derivative of φ is limited around zero, it can be considered a Lipschitzian function. Assumption (A7) acts on Kolmogorov's ε-entropy of $\mathcal{T}_{\mathcal{F}}$. Assumption (A8) means that there $\exists a, b \in \mathbb{R}$ such that for every $z \in \mathcal{T}_{\mathcal{F}}$, $F(b|Z=z)| > 1-\varepsilon$ and $F(a|Z=z)| < \varepsilon$ which will be used to prove that $t(z) = \mathrm{m}ad_c(F_W^z(.))$ is bounded away from 0 for all $z \in \mathcal{T}_{\mathcal{F}}$. Assumption(A9) demonstrates the dependent sample's covariance structure, for more details see [34] and [32] respectively.

3. Technical tools and their proofs

The first difficulty comes because $T_{n,k}(z)$ is random. To resolve this problem, the idea is to frame sensibly $T_{n,k}(z)$ by two non-random windows. More generally, these technical tools could be useful as long as one has to deal with random bandwidths. So we propose in this part some preliminary Lemmas and their proofs that are necessary to prove our main result. Following the notations in [18] or [43].

Let $(A_i, B_i)_{1\le i\le n}$ be n random pairs valued in $(\Omega \times \mathbb{R}, \mathcal{A} \times \mathcal{B}(\mathbb{R}))$, where $(\Omega, \mathcal{A})$ is a general measurable space. Let $\mathcal{T}_\Omega$ be a fixed subset of Ω, we observe that $G: \mathbb{R} \times (\mathcal{T}_\Omega \times \Omega) \to \mathbb{R}^+$ a function such that, $\forall z \in \mathcal{T}_\Omega$, $G(.,(z,.))$ measurable function such that $\forall t, t' \in \mathbb{R}$,

(K_0): $t \le t' \Longrightarrow G(t,d) \le G(t',d)$ for $\forall d \in \mathcal{T}_\Omega \times \Omega$. Let $c(.): \mathcal{T}_\Omega \to \mathbb{R}$ be a non random function such that $\sup_{z\in\mathcal{T}_\Omega}|c(z)| < \infty$. Moreover, for all $z \in \mathcal{T}_\Omega$ and $n \in \mathbb{N}^*$, with

$$c_{n,z}(t) = \frac{\sum_{i=1}^n \Gamma_z\left(\frac{B_i - x}{\hat{t}(z)}\right) G(t,(z,A_i))}{\sum_{i=1}^n G(t,(z,A_i))}.$$

Lemma 3.1. Let $\{J_n(z)\}_{n\in N^*}$ be a sequence of r.r.v. and let $(v_n)_{n\ge1}$ be a decreasing positive sequence with

$\lim_{n\to\infty} v_n = 0$. If for all increasing sequence $\gamma_n \in (0,1)$ with

$\gamma_n - 1 = O(v_n)$, there exist two sequences of real random variable (r.r.v.) $(J_n^-(\gamma_n,z))_{n\in N^*}$ and $(J_n^+(\gamma_n,z))_{n\in N^*}$ such that:

(K_1) $\forall n \in N^*$, $\forall z \in \mathcal{T}_\Omega, J_n^-(\gamma_n,z) \le J_n^+(\gamma_n,z)$,

(K_2) $\mathbb{I}_{\{J_n^-(\gamma_n,z)\le J_n(z)\le J_n^+(\gamma_n,z),\forall z\in\mathcal{T}_\Omega\}} \longrightarrow 1$, a.co. as $n \to \infty$,

(K_3) $\sup_{z\in\mathcal{T}_\Omega}\left|\frac{\sum_{i=1}^n G(J_n^-(\gamma_n,z),(z,A_i))}{\sum_{i=1}^n G(J_n^+(\gamma_n,z),(z,A_i))} - \gamma_n\right| = O_{a.co.}(v_n)$,

(K_4) $\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}(J_n^-(\gamma_n,z)) - c(z)| = O_{a.co}(v_n)$,

(K_5) $\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}(J_n^+(\gamma_n,z)) - c(z)| = O_{a.co}(v_n)$.

Then

$$\sup_{z\in\mathcal{T}_\Omega}\left|c_{n,z}(J_n(z)) - c(z)\right| = O_{a.co}(v_n).$$

Proof. The result for any real valued (r.v.) can be deduced by taking $B_i = B_i^+ - B_i^-$ where

$B_i^+ = \max(B_i, 0)$ and $B_i^- = -\min(B_i, 0)$;

For i=1,...,n, we consider the quantities $\Gamma_z^i(t) = \Gamma_z\left(\frac{B_i - x}{\hat{t}(z)}\right)$.

Under the definition of the r.v $J(n)$, we put

$$J_n^- \le J_n \le J_n^+.$$

It’s clear that

$$G\big(J_n^-(\gamma_n,z),(z,A_i)\big),\leqslant\ G\big(J_n(\gamma_n,z),(z,A_i)\big)\leqslant\ G\big(J_n^+(\gamma_n,z),(z,A_i)\big),$$

$$\sum_{i=1}^{n} G(J_n^-(\gamma),A_i)\leqslant\sum_{i=1}^{n} G\big(J_n(\gamma_n,z),(z,A_i)\big)\leqslant\sum_{i=1}^{n} G\big(J_n^+(\gamma_n,z),(z,A_i)\big).$$

So

$$\frac{1}{\sum_{i=1}^{n} G\big(J_n^+(\gamma_n,z),(z,A_i)\big)}\leqslant\frac{1}{\sum_{i=1}^{n} G\big(J_n(\gamma_n,z),(z,A_i)\big)}\leqslant\frac{1}{\sum_{i=1}^{n} G\big(J_n^-(\gamma_n,z),(z,A_i)\big)}.$$

Under the hypotheses $(A1)-(A5)$, we have

$$\underbrace{\frac{\sum_{i=1}^{n} G\big(J_n^-(\gamma_n,z),(z,A_i)\big)\Gamma_z^i(t)}{\sum_{i=1}^{n} G\big(J_n^+(\gamma_n,z),(z,A_i)\big)}}_{c_{n,z}^+(\gamma_n)}\leqslant\underbrace{\frac{\sum_{i=1}^{n} G\big(J_n(\gamma_n,z),(z,A_i)\big)}{\sum_{i=1}^{n} G\big(J_n(\gamma_n,z),(z,A_i)\big)}}_{c_{n,z}(\gamma_n)}\leqslant\underbrace{\frac{\sum_{i=1}^{n} G\big(J_n^+(\gamma_n,z),(z,A_i)\big)\Gamma_z^i(t)}{\sum_{i=1}^{n} G\big(J_n^-(\gamma_n,z),(z,A_i)\big)}}_{c_{n,z}^-(\gamma_n)}.$$

In the other hand, we can express the r.r.v: $c_{n,z}^-(\gamma_n)$ and $c_{n,z}^+(\gamma_n)$, in the following way

$$c_{n,z}^-(\gamma_n)=c_{n,z}(\gamma_n)\times\frac{\sum_{i=1}^{n} G\big(J_n^-(\gamma_n,z),(z,A_i)\big)}{\sum_{i=1}^{n} G\big(J_n^+(\gamma_n,z),(z,A_i)\big)}$$

and

$$c_{n,z}^+(\gamma_n)=c_{n,z}(\gamma_n)\times\frac{\sum_{i=1}^{n} G\big(J_n^+(\gamma_n,z),(z,A_i)\big)}{\sum_{i=1}^{n} G\big(J_n^-(\gamma_n,z),(z,A_i)\big)},$$

So under (K_2) and (K_3), we have

$$c_{n,z}^-(\gamma_n)\xrightarrow{a.co.}\gamma_n\,c(z),\text{and}c_{n,z}^+(\gamma_n)\xrightarrow{a.co.}c(z)/\gamma_n. \quad (10)$$

For all sequence $\gamma_n\in(0,1)$ with $\gamma_n-1=O(v_n)$, (K_3), (K_4) and (K_5) give

$$\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}^-(\gamma_n)-c(z)|\leq\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}^-(\gamma_n)-\gamma_n c(z)|+|c(z)||\gamma_n-1|=O_{a.co.}(v_n), \quad (11)$$

and

$$\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}^+(\gamma_n)-c(z)|=O_{a.co.}(v_n). \quad (12)$$

For $\varepsilon>0$ we note

$$T_n(\varepsilon)=\left\{\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}(J_n(z))-c(z)|\leq\varepsilon v_n\right\},$$

and for all sequence $\gamma_n\in(0,1)$ with $\gamma_n-1=O(v_n)$,

$$S_n^-(\varepsilon,\gamma_n)=\left\{\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}^-(\gamma_n)-c(z)|\leq\varepsilon v_n\right\},$$

$$S_n^+(\varepsilon,\gamma_n)=\left\{\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}^+(\gamma_n)-c(z)|\leq\varepsilon v_n\right\},$$

$$S_n(\gamma_n)=\left\{c_{n,z}^-(\gamma_n)\leq c_{n,z}(J_n(z))\leq c_{n,z}^+(\gamma_n),\forall z\in\mathcal{T}_\Omega\right\}.$$

It is evident that, for all $\gamma_n\in(0,1)$ with $\gamma_n-1=O(v_n)$,

$$\forall\varepsilon>0,S_n^-(\varepsilon,\gamma_n)\cap S_n^+(\varepsilon,\gamma_n)\cap S_n(\gamma_n)\subset T_n(\varepsilon). \quad (13)$$

Let $G_n(\gamma_n)=\{J_n^-(\gamma_n,z)\leq J_n(z)\leq J_n^+(\gamma_n,z),\forall z\in\mathcal{T}_\Omega\}$, then (K_0) implies that

$G_n(\gamma_n) \subset S_n(\gamma_n)$ and from (13), we obtain

$$\forall \varepsilon > 0, T_n(\varepsilon)^c \subset S_n^-(\gamma_n)^c \cup S_n^+(\gamma_n)^c \cup G_n(\gamma_n)^c,$$

and consequently

$$\mathbb{P}\left(\sup_{z\in\mathcal{T}_\Omega}\left|c_{n,z}\big(J_n(z)\big) - c(z)\right| > \varepsilon v_n\right) \leq \mathbb{P}\left(\sup_{z\in\mathcal{T}_\Omega}\left|c_{n,z}^-(\gamma_n,\varepsilon) - c(z)\right| > \varepsilon v_n\right) +$$

$$\mathbb{P}(\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}^+(\gamma_n,\varepsilon) - c(z)| > \varepsilon v_n) + \mathbb{P}(\mathbb{I}_{\{J_n^-(\gamma_n,z)\leq J_n(z)\leq J_n^+(\gamma_n,z),\forall z\in\mathcal{T}_\Omega\}} = 0).$$

Then, for some $\varepsilon_0 > 0$

$$\sum_{n=1}^{\infty} \mathbb{P}(\sup_{z\in\mathcal{T}_\Omega}|c_{n,z}(J_n(z)) - c(z)| > \varepsilon_0 v_n) < \infty. \tag{14}$$

Lemma 3.2. Let $F_n(w|Z = z)$ be a sequence of conditional distribution functions verifying

$$\sup_{z\in\mathcal{T}_\mathcal{F}}\sup_{w\in\mathbb{R}}|F_n(w|Z = z) - F(w|Z = z)| \longrightarrow 0. \tag{15}$$

Then, if F verifies assumption (A8), there exist positive constants $A \leq B$ such that $t_n(z) = mad_c(F_n(.|Z = z))$ verifies $A \leq t_n(z) \leq B$ for all $z \in \mathcal{T}_\mathcal{F}$ compact and $n \geq n_0$.

Proof. This proof is very close to that of Lemma $A.4$. in Boente et al. [15].

According to (A1) and (A3a), we have that $L(1) > C > 0$,

$$\forall z \in \mathcal{T}_\mathcal{F},\ \exists 0 < C < C' < \infty,\ C\varphi_z(h_L) < E[L(d(z,Z_1)/h_L)] < C'\varphi_z(h_L). \tag{16}$$

In the case when $L(1) = 0$, combining (A1) and (A6) gives the same result.

4 Main result

We start by reminding the uniform asymptotic properties of $\hat{\theta}(z)$ defined in (7). The following Theorem was proved by [3] in the special case when $h_L(z) = h_L$ for all $z \in \mathcal{T}_\mathcal{F}$, but their proof can be followed line by line under (18)). This general condition (18) will be a crucial preliminary tool for us.

Under assumptions (A1)-(A9), if in addition, $h_L(z)$ in (7) satisfies

$$\lim_{n\to\infty}(\varphi_z(T_{n,k}(z)) - \varphi_z(h_L(z))) = 0, a.co. \tag{17}$$

and

$$0 < Ch_L \leq \inf_{z\in\mathcal{T}_\mathcal{F}} h_L(z) \leq \sup_{z\in\mathcal{T}_\mathcal{F}} h_L(z) \leq C'h_L < \infty, \tag{18}$$

where $h_L \to 0 (n \to \infty)$ such that, for n large enough,

$$\frac{\log^2 n}{n\phi(h_L)} < \Psi_{\mathcal{T}_\mathcal{F}}\left(\frac{\log n}{n}\right) < \frac{n\phi(h_L)}{\log n}, \tag{19}$$

and

$$0 < C < \frac{n\varphi(h_L)}{\log^2 n} < C' < +\infty. \tag{20}$$

Then we have

$$\sup_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}(z) - \theta(z)| = O(h_L^{\beta}) + O_{a.co}\left(\sqrt{\frac{\eta_n^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}\right). \tag{21}$$

We can now state our main result, whose proof will be presented in Section 5.

Theorem 4.2. Under the assumptions (A1)-(A9), and for n large enough, then we have

$$\sup_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_{kNN}(z) - \theta(z)| = O\left(\varphi^{-1}\left(\frac{k}{n}\right)^{\beta_1}\right) + O_{a.co}\left(\sqrt{\frac{\eta_n^2\Psi_{\mathcal{T}_\mathcal{F}}\left(\frac{\log n}{n}\right)}{n^2}}\right).$$

5 Proofs

5.1 Proofs of Theorem4.1

The proof is based on the following decomposition

$$\widehat{\theta}(z)-\theta(z)=\frac{1}{\widehat{\theta}_N^1(z)}[\widehat{\theta}_N^2(z)-\mathbb{E}\widehat{\theta}_N^2(z)]+\frac{1}{\widehat{\theta}_N^1(z)}[\mathbb{E}\widehat{\theta}_N^2(z)-\theta(z)]+[1-\widehat{\theta}_N^1(z)]\frac{\theta(z)}{\widehat{\theta}_N^1(z)},$$

where,

$$\widehat{\theta}_N^1(z)=\frac{1}{n\mathbb{E}[L(h_L^{-1}d(z,Z_1))]}\sum_{i=1}^n L(h_L^{-1}d(z,Z_i)),$$

$$\widehat{\theta}_N^2(z)=\frac{1}{n\mathbb{E}[L(h_L^{-1}d(z,Z_1))]}\sum_{i=1}^n L(h_L^{-1}d(z,Z_i))\Gamma_z(\frac{W_i-x}{\hat{t}(z)}).$$

Thus, the proof of the proposition is valid as soon as the following three Lemmas can been checked respectively.

Lemma 5.1. Under the assumptions (A1), (A3), (A4) and (A7), we have

$$\sup_{z\in\mathcal{T}_{\mathcal{F}}}|\mathbb{E}\widehat{\theta}_N^2(z)-\theta(z)|=O(h_L^{\beta}).$$

Proof. We have

$$|\mathbb{E}\widehat{\theta}_N^2(z)-\theta(z)|=|\frac{1}{n\mathbb{E}[L(h_L^{-1}d(z,Z_1))]}\mathbb{E}[\sum_{i=1}^n L(h_L^{-1}d(z,Z_i))\Gamma_z(\frac{W_i-x}{\hat{t}(z)})]-\theta(z)|,$$

$$\leq|\frac{1}{\mathbb{E}[L(h_L^{-1}d(z,Z_1))]}\mathbb{E}[L(h_L^{-1}d(z,Z_1))\Gamma_z(\frac{W_1-x}{\hat{t}(z)})]-\theta(z)|,$$

$$\leq\frac{1}{\mathbb{E}[L(h_L^{-1}d(z,Z_1))]}[\mathbb{E}[L(h_L^{-1}d(z,Z_1))|\theta(Z_1,x,\hat{t}(z))-\theta(z)|]].$$

As a result we obtain

$$\forall z\in\mathcal{T}_{\mathcal{F}},|\mathbb{E}\widehat{\theta}_N^2(z)-\theta(z)|\leq\frac{1}{\mathbb{E}[L(h_L^{-1}d(z,Z_1))]}[\mathbb{E}[L(h_L^{-1}d(z,Z_1))|\theta(Z_1,x,\hat{t}(z))-\theta(z)|]].$$

Accordingly, with the assumptions (A1), (A4) and (16) we get

$$\forall z\in\mathcal{T}_{\mathcal{F}},\left|\mathbb{E}\widehat{\theta}_N^2(z)-\theta(z)\right|\leq C\frac{1}{\mathbb{E}\left[L\left(h_L^{-1}d(z,Z_1)\right)\right]}\left[\mathbb{E}\left[L\left(h_L^{-1}d(z,Z_1)\right)\mathbb{I}_{B(z,h_L)}(Z_1)d^{\beta}(Z_1,z)\right]\right]\leq Ch_L^{\beta},$$

which gives the result.

Lemma 5.2. Under the assumptions of Theorem 4, we have

$$\sup_{z\in\mathcal{T}_{\mathcal{F}}}|\widehat{\theta}_N^2(z)-\mathbb{E}\widehat{\theta}_N^2(z)|=O_{a.co}(\sqrt{\frac{\eta_n^2\Psi_{\mathcal{T}_{\mathcal{F}}}(\frac{\log n}{n})}{n^2}}).$$

Proof. We consider the following decomposition

$$\sup_{z\in\mathcal{T}_{\mathcal{F}}}|\widehat{\theta}_N^2(z)-\mathbb{E}\widehat{\theta}_N^2(z)|\leq\underbrace{\sup_{z\in\mathcal{T}_{\mathcal{F}}}|\widehat{\theta}_N^2(z)-\widehat{\theta}_N^2(z_{\ell(z)},x,\hat{t}(z))|}_{G_1}+\underbrace{\sup_{z\in\mathcal{T}_{\mathcal{F}}}|\widehat{\theta}_N^2(z_{\ell(z)},x,\hat{t}(z))-\mathbb{E}\widehat{\theta}_N^2(z_{\ell(z)},x,\hat{t}(z))|}_{G_2}$$

$$+\underbrace{\sup_{z\in\mathcal{T}_{\mathcal{F}}}|\mathbb{E}\widehat{\theta}_N^2(z_{\ell(z)},x,\hat{t}(z))-\mathbb{E}\widehat{\theta}_N^2(z)|.}_{G_3}$$

By assumption (A3) which follows that $0<C\leq\mathbb{E}[L(d(z,z_1)/h_L(z))]\leq C'<+\infty$, we treat the term G_1 as following

$$G_1=\sup_{z\in\mathcal{T}_{\mathcal{F}}}\frac{1}{n}\sum_{i=1}^n|\frac{1}{\mathbb{E}[L(h_L^{-1}d(z,Z_1))]}L(h_L^{-1}d(z,Z_i))\Gamma_z(\frac{W_i-x}{\hat{t}(z)})$$
$$-\frac{1}{\mathbb{E}[L(h_L^{-1}d(z_{\ell(z)},Z_1))]}L(h_L^{-1}d(z_{\ell(z)},Z_i))\Gamma_z(\frac{W_i-x}{\hat{t}(z)})|$$

$$\leq \sup_{z\in\mathcal{T}_\mathcal{F}} \frac{1}{n\varphi(h_L)} \sum_{i=1}^{n} |\Gamma_z(\frac{W_i-x}{\hat{t}(z)})||L(h_L^{-1}d(z,Z_i)) - L(h_L^{-1}d(z_{\ell(z)},Z_i))|\mathbb{I}_{B(z,h_L)\cup B(z_{\ell(z)},h_L)}(Z_i).$$

Hence, for $\zeta = \sqrt{2\varpi} > 0$, we have

$$\mathbb{P}(G_1 > \zeta\sqrt{\frac{\eta_{n,1}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}) \leq$$

$$\mathbb{P}(\sup_{z\in\mathcal{T}_\mathcal{F}} \frac{1}{n\varphi(h_L)} \sum_{i=1}^{n} |\Gamma_z(\frac{W_i-x}{\hat{t}(z)})||L(h_L^{-1}d(z,Z_i)) - L(h_L^{-1}d(z_{\ell(z)},Z_i))|\mathbb{I}_{B(z,h_L)\cup B(z_{\ell(z)},h_L)}(Z_i) > \zeta\sqrt{\frac{\eta_{n,1}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}),$$

$$\leq N_\varepsilon(\mathcal{T}_\mathcal{F}) \max_{\ell\in 1,\dots,N_\varepsilon(\mathcal{T}_\mathcal{F})} \mathbb{P}(\sup_{z\in\mathcal{T}_\mathcal{F}} \frac{1}{n\varphi(h_L)} \sum_{i=1}^{n} |\Gamma_z(\frac{W_i-x}{\hat{t}(z)})|\mathbb{I}_{B(z,h_L)\cup B(z_{\ell(z)},h_L)}(Z_i) > \zeta\sqrt{\frac{\eta_{n,1}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}).$$

And we have that

$$\mathbb{E}[\Gamma_z^m(\frac{W-x}{\hat{t}(z)})] = \mathbb{E}[\mathbb{E}[\Gamma_z^m(\frac{W-x}{\hat{t}(z)})|Z]] = \int \delta(z)dP_Z < C < \infty,$$

which implies that

$$\mathbb{E}[|\frac{\varepsilon\Gamma_z(\frac{W_i-x}{\hat{t}(z)})}{\varphi(h)} \sup_{z\in\mathcal{T}_\mathcal{F}} \mathbb{I}_{B(z,h_L)\cup B(z_{\ell(z)},h_L)}|] \leq \frac{C\varepsilon^m}{h_L^m\varphi(h_L)^{m-1}}.$$

Moreover,by the assumption (A5), Markov's inequality and Proposition A.11.(i) in [32], it follows that

$$\mathbb{P}(\sup_{z\in\mathcal{T}_\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n} |\Gamma_z(\frac{W_i-x}{\hat{t}(z)})|\mathbb{I}_{B(z,h_L)\cup B(z_{\ell(z)},h_L)}(z_i) > \zeta\sqrt{\frac{\eta_{n,1}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}})$$

$$\leq \left\{C(1+\frac{\zeta^2\Psi_{\mathcal{T}_\mathcal{F}}(\varepsilon)}{r})^{-(r/2)} + \frac{n}{r}(\frac{r}{\zeta\sqrt{\eta_{n,1}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}})^{(a+1)p/(a+p)}\right\},$$

where $a > 1,\ p > 2,\ r = \log^2 n$. Then, by assumptions (A9) and (20), we have

$$\mathbb{P}(G_1 > \zeta\sqrt{\frac{\eta_{n,1}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}) \leq CN_\varepsilon(\mathcal{T}_\mathcal{F})\{CN_\varepsilon(\mathcal{T}_\mathcal{F})^{-\varpi} + n^{1-\theta}(2\varpi)^{-((a+1)p/2(a+p))}C'(\log n)^{2(a+1)p/(a+p)-2}\}$$

$$\leq CN_\varepsilon(\mathcal{T}_\mathcal{F})\{CN_\varepsilon(\mathcal{T}_\mathcal{F})^{-\varpi}n^{1-\theta+\varepsilon_0}\},$$

for some $\varepsilon_0 > 0$ and large enough θ. On the contrary, by the definition of Kolmogorov's ε-entropy and (19), it follows that $n^{1-\theta+\varepsilon_0} = O(N_\varepsilon(\mathcal{T}_\mathcal{F})^{\varpi})$. Hence, by assumption (A9), we obtain

$$G_1 = O_{a.co}(\sqrt{\frac{\eta_{n,1}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}). \quad (22)$$

Now, we deal with G_2. For $\zeta > 0$, we have

$$\mathbb{P}\left(G_2 > \zeta\sqrt{\frac{\eta_{n,2}^2\Psi_{\mathcal{T}_\mathcal{F}}\left(\frac{\log n}{n}\right)}{n^2}}\right) \leq$$

$$\mathbb{P}(\sup_{z\in\mathcal{T}_\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n} \left|\frac{\Gamma_z\left(\frac{W_i-\varrho}{\hat{t}(z)}\right)L\left(h_L^{-1}d\left(z_{\ell(z)},Z_i\right)\right)}{\mathbb{E}L\left(h_L^{-1}d\left(z_{\ell(z)},Z_1\right)\right)} - \frac{\mathbb{E}\Gamma_z\left(\frac{W_i-\varrho}{\hat{t}(z)}\right)L\left(h_L^{-1}d\left(z_{\ell(z)},Z_i\right)\right)}{\mathbb{E}L\left(h_L^{-1}d\left(z_{\ell(z)},Z_1\right)\right)}\right|) > \zeta(\sqrt{\frac{\eta_{n,2}^2\Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}).$$

Let

$$\Lambda_{\ell_i}^p = \frac{1}{\mathbb{E}L(h_L^{-1}d(z_{\ell(z)},Z_1))}(L(h_L^{-1}d(z_{\ell(z)},Z_i))\Gamma_z(\frac{W_i-\varrho}{\hat{t}(z)}) - \mathbb{E}[L(h_L^{-1}d(z_{\ell(z)},Z_i))\Gamma_z(\frac{W_i-x}{\hat{t}(z)})]).$$

And we have

$$\mathbb{P}(G_2 > \zeta \sqrt{\frac{\eta_{n,2}^2 \Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}) \leq N_\varepsilon(\mathcal{T}_\mathcal{F}) \max_{\ell \in 1,\dots,N_\varepsilon(\mathcal{T}_\mathcal{F})} \mathbb{P}(\frac{1}{n}|\sum_{i=1}^n \Lambda_{\ell_i}| > \zeta \sqrt{\frac{\eta_{n,2}^2 \Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}),$$

by using the same result as with the treatment of G_1 we get

$$G_2 = O_{a.co}(\sqrt{\frac{\eta_{n,2}^2 \Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}). \quad (23)$$

In the end, we treat G_3. It is evident that $G_3 \leq \mathbb{E}[\sup_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N(z) - \hat{\theta}_N(z_{\ell(z)}, x, \hat{t}(z_{\ell(z)}))|]$, similar steps of treating G_1 allow to get

$$G_3 = O_{a.co}(\sqrt{\frac{\eta_{n,1}^2 \Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}). \quad (24)$$

Lemma 5.3. Under the assumptions (A1) and (A4)-(A7), we have

$$\sup_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N^1(z) - 1| = O_{a.co}(\sqrt{\frac{\eta_n^2 \Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}). \quad (25)$$

Proof. The steps in this proof are the same as in the proof of Lemma 5.1 . For this, we conserve these notations and use the decomposition that follows

$$\sup_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N^1(z) - \mathbb{E}\hat{\theta}_N^1(z)| \leq \underbrace{\sup_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N^1(z) - \hat{\theta}_N^1(z_{\ell(z)})|}_{F_1} + \underbrace{\sup_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N^1(z_{\ell(z)}) - \mathbb{E}\hat{\theta}_N^1(z_{\ell(z)})|}_{F_2}$$

$$+ \underbrace{\sup_{z\in\mathcal{T}_\mathcal{F}}|\mathbb{E}\hat{\theta}_N^1(z_{\ell(z)}) - \mathbb{E}\hat{\theta}_N^1(z)|}_{F_3}.$$

Then, using the same proofs as (22) to (24), we get

$$F_1 = O_{a.co}\left(\sqrt{\frac{\eta_{n,3}^2 \Psi_{\mathcal{T}_\mathcal{F}}\left(\frac{\log n}{n}\right)}{n^2}}\right), F_2 = O_{a.co}(\sqrt{\frac{\eta_{n,4}^2 \Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}),$$

and

$$F_3 = O_{a.co}(\sqrt{\frac{\eta_{n,3}^2 \Psi_{\mathcal{T}_\mathcal{F}}(\frac{\log n}{n})}{n^2}}).(26)$$

Finally, the claimed result is obtained from the last decomposition and (26).

Corollary 5.4. Under the assumptions of Lemma 5.1, we have

$$\sum_{n=1}^{\infty} \mathbb{P}(\inf_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N^1(z)| < \frac{1}{2}) < \infty.$$

Proof. We have that

$\inf_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N^1(z)| \leq \frac{1}{2} \Rightarrow \exists z \in \mathcal{T}_\mathcal{F}$ such that $1 - \hat{\theta}_N^1(z) \geq \frac{1}{2} \Rightarrow \sup_{z\in\mathcal{T}_\mathcal{F}}|1 - \hat{\theta}_N^1(z)| \geq \frac{1}{2}$.

We conclude that

$$\mathbb{P}(\inf_{z\in\mathcal{T}_\mathcal{F}}|\hat{\theta}_N^1(z)| \leq \frac{1}{2}) \leq \mathbb{P}(\sup_{z\in\mathcal{T}_\mathcal{F}}|1 - \hat{\theta}_N^1(z)| \geq \frac{1}{2}).$$

As a result

$\sum_{n=1}^{\infty} \mathbb{P}(\inf_{z\in\mathcal{T}_{\mathcal{F}}} |\hat{\theta}_N^1(z)| < \frac{1}{2}) \leq \infty$.

5.2 Proofs of main result

Similar to the proof of ([43], Theorem 2), we must to investigate the conditions of Lemma 3.

For that, we denote: $\mathcal{T}_\Omega = \mathcal{T}_{\mathcal{F}}, A_i = z_i, B_i = w_i, G(t,(z,A_i)) = L(d(z,A_i)/t), J_n(z) = T_{n,k}(z), c_{n,z}(T_{n,k}(z)) = \hat{\theta}_{kNN}(z)$ and $c(z) = \theta(z)$. Let $\gamma_n \in (0,1)$ be an increasing sequence such that $\gamma_n - 1 = O(v_n)$, where

$$v_n = \phi^{-1}(\frac{k}{n})^{\beta} + \sqrt{\frac{\eta_n^2 \Psi_{\mathcal{T}_{\mathcal{F}}}(\frac{\log n}{n})}{n^2}}.$$

Let $h_L = \phi^{-1}(\frac{k}{n})^{\beta}$, we choose $J_n^-(\gamma_n, z)$ and $J_n^+(\gamma_n, z)$ such that

$$\varphi_z(J_n^-(\gamma_n, z)) = \varphi_z(h_L(z))\gamma_n^{1/2}, \quad (27)$$

and

$$\varphi_z(J_n^+(\gamma_n, z)) = \varphi_z(h_L(z))\gamma_n^{-(1/2)}. \quad (28)$$

Checking (K_4) and (K_5): we note that the local bandwidth $J_n^-(\gamma_n, z)$ satisfies (17), (18) and (20), we have now

$$\sup_{z\in\mathcal{T}_{\mathcal{F}}} |c_{n,z}(J_n^-(\gamma_n, z)) - c(z)| = O_{a.co}(\phi^{-1}(\frac{k}{n})^{\beta} + \sqrt{\frac{\eta_n^2 \Psi_{\mathcal{T}_{\mathcal{F}}}(\frac{\log n}{n})}{n^2}})$$

$$= O_{a.co}(v_n).$$

Consequently, (K_4) in Lemma 3 is valid. We use the same steps for $J_n^+(\gamma_n, z)$, we obtain

$$\sup_{z\in\mathcal{T}_{\mathcal{F}}} |c_{n,z}(J_n^+(\gamma_n, z)) - c(z)| = O_{a.co}(v_n),$$

Therefore (K_5) is also correct.

We check (K_1) and (K_2): with (27) and (28), we have

$\varphi_z(J_n^-(\gamma_n, z)) \leq \varphi_z(h_L(z)) \leq \varphi_z(J_n^+(\gamma_n, z))$. Thus, with (17) and the property of $\varphi_z(.)$, we have as a result

$J_n^-(\gamma_n, z) \leq T_{n,k}(z) \leq J_n^+(\gamma_n, z)$, a.co

and

$\mathbb{I}_{\{J_n^-(\gamma_n, z) \leq J_n(z) \leq J_n^+(\gamma_n, z), \forall z \in \mathcal{T}_\Omega\}} \to 1$, a.co. as $n \to \infty$.

Consequently, (K_1) and (K_2) in Lemma 5.1 are valid. Checking (K_3): Same as Kudraszow and Vieu [43], $\forall z \in \mathcal{T}_{\mathcal{F}}$, indicate

$$f^*(z, h_L(z)) =: \mathbb{E}[L(\frac{d(z,z_1)}{h_L(z)})], H_1 =: \frac{f^*(z, J_n^-(\gamma_n, z))}{f^*(z, J_n^+(\gamma_n, z))},$$

$$H_2 =: \frac{\hat{\theta}_N^1(z, J_n^-(\gamma_n, z))}{\hat{\theta}_N^1(z, J_n^+(\gamma_n, z))} - 1, H_3 =: \frac{f^*(z, J_n^-(\gamma_n, z))}{f^*(z, J_n^+(\gamma_n, z))}\gamma_n - 1.$$

After that, we come at

$$\left| \frac{\sum_{i=1}^n L\left(\frac{d(z,z_i)}{J_n^-(\gamma_n, z)}\right)}{\sum_{i=1}^n L\left(\frac{d(z,z_i)}{J_n^+(\gamma_n, z)}\right)} - \gamma_n \right| \leq |H_1||H_2| + |H_1||H_3|. \quad (29)$$

With (A3), we get

$$\sup_{z\in\mathcal{T}_{\mathcal{F}}} |H_1| \leq C. \quad (30)$$

In addition, with (25), we have that

$$\sup_{z\in\mathcal{J}_{\mathcal{F}}}|H_2| \le \frac{\sup_{z\in\mathcal{J}_{\mathcal{F}}}|\hat{\theta}_N^1(z,J_n^-(\gamma_n,z))-1| + \sup_{z\in\mathcal{J}_{\mathcal{F}}}|\hat{\theta}_N^1(z,J_n^+(\gamma_n,z))-1|}{\inf_{z\in\mathcal{J}_{\mathcal{F}}}|\hat{\theta}_N^1(z,J_n^+(\gamma_n,z))|} = O_{a.co}(\sqrt{\frac{\eta_n^2\Psi_{\mathcal{J}_{\mathcal{F}}}(\frac{\log n}{n})}{n^2}}).$$

Then, we use the (Lemma 1, p. 10) of [30] with (A2) and also the fact that

$$\varphi_z(J_n^-(\gamma_n,z))/\varphi_z(J_n^+(\gamma_n,z)) = \gamma_n,$$

we get

$$\sup_{z\in\mathcal{J}_{\mathcal{F}}}|H_3| = O(\phi(h_L)h_L^{\beta}) = O((\sqrt{\gamma_n}\phi^{-1}(\frac{k}{n}))^{\beta}). \quad (31)$$

In consequence, with the fact that $\gamma_n \to 1$ and integrating (29)-(31), we have that

$$\sup_{z\in\mathcal{J}_{\mathcal{F}}}|\frac{\sum_{i=1}^n L(\frac{d(z,z_i)}{J_n^-(\gamma_n,z)})}{\sum_{i=1}^n L(\frac{d(z,z_i)}{J_n^+(\gamma_n,z)})} - \gamma_n| = O_{a.co}(v_n).$$

It should be noted that (K_0) is clearly satisfied because of (A3a), additionally (L_1) is also easily satisfied by building of $J_n^-(\gamma_n,z)$ and $J_n^+(\gamma_n,z)$. So, one can therefore apply Lemma 3, and (9) is precisely the consequence of Theorem 4.

6 Simulated data application

This section compares the Robust kernel estimator introduced by [9] with the finite sample behavior of the proposed Robust kNN estimator for strong mixing dependency samples with unknown scale parameter.

First, we consider the following nonparametric functional regression model:

$$W_i = m(Z_i) + \varepsilon_i, \qquad i = 1,2,\dots,200,$$

where $m(Z_i) = \left(\int_0^{\pi/3} Z_i'(t)dt\right)^2$, $Z_i'(t)$ is the first derivative of $Z_i(t)$, $\{\varepsilon_i\}_{i=1}^{200}$ are i.i.d. according to $N(0,0.05)$ and the curves $Z_i(t)$ are generated as follows:

$$Z_i(t) = x_i t^2 + \sin\left(y_i\left(t - \frac{\pi}{3}\right)\right), \quad i = 1,2,\dots,200, t \in \left[0, \frac{\pi}{3}\right],$$

where $\{x_i\}_{i=1}^{200}$ are i.i.d. according to $U(0,\pi/9)$, $y_i = \frac{1}{4}y_{i-1} + \xi_i$, $\{\xi_i\}_{i=1}^{200}$ are i.i.d. according to $N(0,1)$ and independent to $\{x_i\}_{i=1}^{200}$ and $\{b_i\}_{i=1}^{200}$ respectively, b_0 is from $N(0,1)$.

Figure 1 displays the curves of the sample sizes $n = 200$.

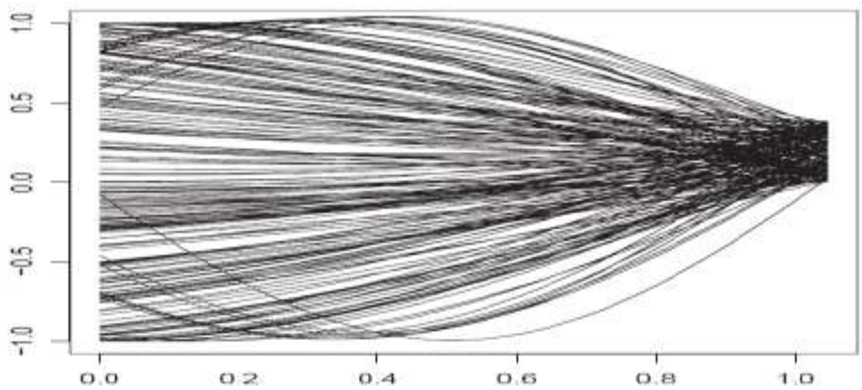

Figure 1: The curves $Z_i(t), t \in [0,\pi/3]$ for $i = 1,\dots,200$.

Second, we need to select a suitable semi-metric $d(.,.)$, kernel $L(.)$, smoothing parameter k_{opt} for functional kNN estimator. For that purpose, we choose the asymmetrical quadratic kernel defined as $L(u) = \frac{3}{4}\left(\frac{12}{11} - u^2\right)1_{[0,1]}(u)$. Meanwhile, because of the smoothness of curves $Z(t)$, we consider the following semimetric based on the first derivative:

$$d^{\text{deriv}}(Z_i, Z_j) = \sqrt{\int_0^{\pi/3} \left(Z_i'(t) - Z_j'(t)\right)^2 \mathrm{d}t}, \quad \forall Z_i, Z_j \in \mathcal{F}.$$

Our main goal is to compare our estimator (Robust Equivariant Estimator REE) $\hat{\theta}(z)$ with Robust Kernel Estimator (RKE) $\tilde{\theta}(z)$ introduced by [9], where $\hat{\theta}(z)$, $\tilde{\theta}(z)$ are define as following

$\hat{\theta}(z)$ is the zero with respect to x of $\dfrac{\sum_{i=1}^{n} L\left(\frac{d(z,Z_i)}{T_{n,k}(z)}\right)\Gamma_z\left(\frac{W_i - x}{\hat{t}(z)}\right)}{\sum_{i=1}^{n} L\left(\frac{d(z,Z_i)}{T_{n,k}(z)}\right)} = 0,$

$\tilde{\theta}(z)$ is the zero with respect to x of $\dfrac{\sum_{i=1}^{n} L\left(\frac{d(z,Z_i)}{h_L(z)}\right)\Gamma_z(W_i - x)}{\sum_{i=1}^{n} L\left(\frac{d(z,Z_i)}{h_L(z)}\right)} = 0,$

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

$$MSE_{\hat{\theta}} = n^{-1}\sum_{i=1}^{n} \left(\theta(Z_i) - \hat{\theta}(Z_i)\right)^2, MSE_{\tilde{\theta}} = n^{-1}\sum_{i=1}^{n} \left(\theta(Z_i) - \tilde{\theta}(Z_i)\right)^2,$$

In what follows, we randomly split the 200-sample into two parts: one is a training sample $\{Z_i, W_i\}_{i=1}^{150}$ which is used to model, and the other is a testing sample $\{Z_i, W_i\}_{i=151}^{200}$ which is used to verify the prediction effect. On the one hand, by the training sample, we can select the optimal parameter k_{opt} for kNN estimator by the following cross-validation procedures.

Concretely, for the robust kNN we select $k_{opt} = \operatorname{argmin}_k CV_2(k)$, where $CV_2(k) = \sum_{i=1}^{n} \left(W_i - \hat{\theta}^{(-i)}(Z_i)\right)^2$ and

$$\hat{\theta}^{(-i)}(Z) = \underset{t}{\operatorname{argmin}} \frac{\sum_{j=1, j\neq i}^{n} \Gamma_z\left(\frac{W_i - x}{\hat{t}(z)}\right) L\left(\frac{d(Z_j, Z)}{T_{n,k}(Z)}\right)}{\sum_{j=1, j\neq i}^{n} L\left(\frac{d(Z_j, Z)}{T_{n,k}(Z)}\right)}.$$

Then the robust kernel estimator by $k_{opt} = \operatorname{argmin}_k CV_3(k)$, where $CV_3(k) = \sum_{i=1}^{n} \left(W_i - \tilde{\theta}^{(-i)}(Z_i)\right)^2$ and

$$\tilde{\theta}^{(-i)}(Z) = \underset{t}{\operatorname{argmin}} \frac{\sum_{j=1, j\neq i}^{n} \Gamma_z(W_i - x) L\left(\frac{d(Z_j, Z)}{T_{n,k}(Z)}\right)}{\sum_{j=1, j\neq i}^{n} L\left(\frac{d(Z_j, Z)}{T_{n,k}(Z)}\right)}.$$

On the other hand, by the testing sample, we can calculate the prediction values of the response variables denoted by $\{\widehat{W}_i\}_{i=151}^{200}$ for Robust kNN method and $\{\widetilde{W}_i\}_{i=151}^{200}$ for Robust kernel method respectively. Thus, MSE of the predicted responses for the two methods are illustrated in Figure 3 where we see clearly that the forecasting of Robust kNN estimator is more accurate than that of Robust kernel one under the strong mixing functional dependent sample. The similar results are also shown when the sample sizes are $n = 300$ and $n = 500$.

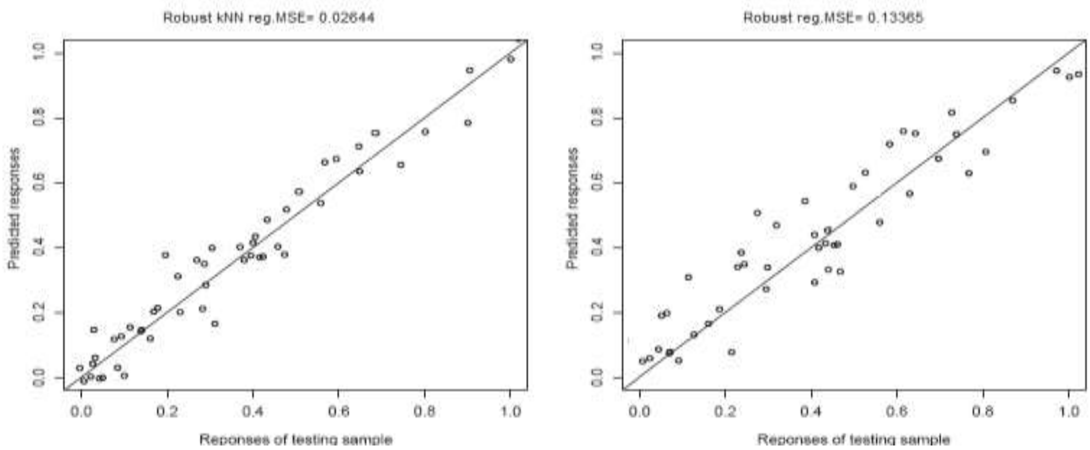

Figure 2: MSE of the Robust kNN method and the Robust kernel method respectively

To further explore the performances of the two methods, we carry out $M = 100$ independent replications of the experiment for Robust kNN estimator, Robust regression estimator when the sample sizes are $n = 200, n = 300$ and $n = 500$ respectively. In each case, let the testing sample sizes be fifty.

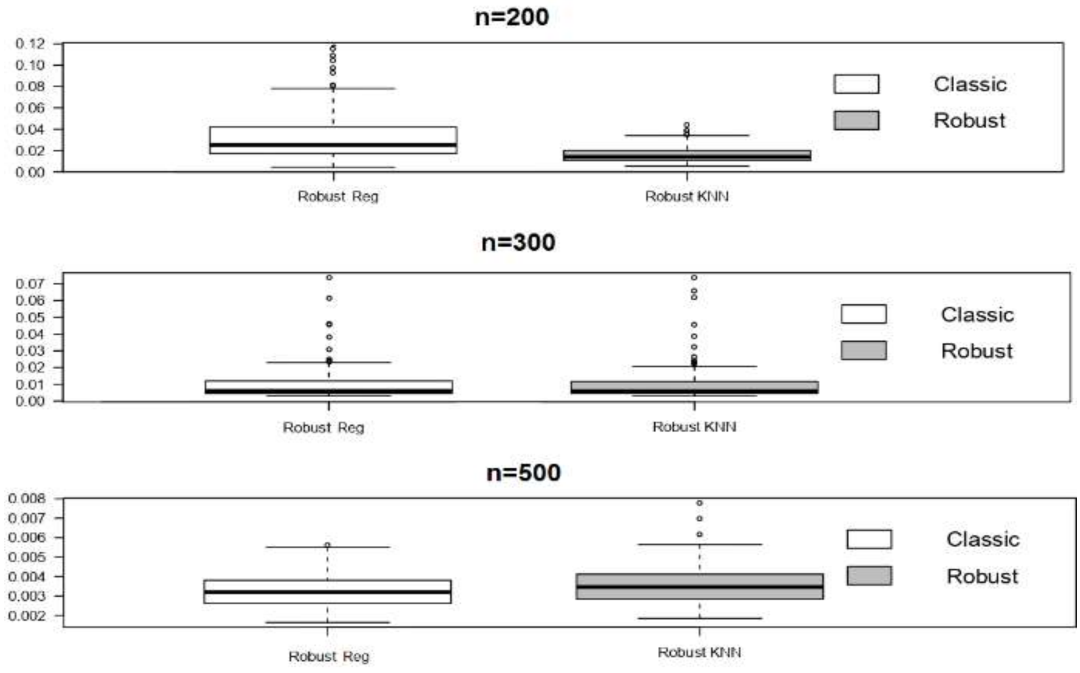

Figure 3: The box plots of the MSE of the prediction values by the two methods for the different sample sizes when the Number of repetitions of experiments is M=100

Meanwhile, let us calculate the average of MSE of Robust kNN estimator and the Robust regression estimator. The results are reported in Table 1.

Table 1: The comparison of different methods

n	200	300	500
MSE Robust reg.	0.01388555	0.016432	0.1159574
MSE k NN Robust reg.	0.003120220	0.00877745	0.04466478

We can see that, for fixed sample sizes n, Robust kNN estimator is much more smaller than the Robust regression estimator, or the prediction accuracy of Robust kNN regression method is better than that of the other method even in the case of the strong mixing dependence.

7.Conclusions

The kNN approach is a smoothing alternative that allows for the development of an adaptive estimator for a variety of statistical problems, including bandwidth choice.
The assumption that the bandwidth parameter in the kNN method is a random variable adds to the complexity of this problem.
In the situation of equivariant robustification results, the key innovation of this approach is to estimate the regression function by mixing two essential statistical techniques: the regression estimators for strong data when the scale parameter is unknown with the kNN method. This strategy allowed for the development of a new estimator that combines the benefits of both methods.
Another unanswered concern is how to treat a more general case in which data are generated from a functional *alpha*-mixing dependency and the scale parameter is unknown . Precisely, we can obtain the almost complete convergence of the same constructed estimator under standard conditions allowing us to explore different structural axes of the topic. We emphasize that, contrary to the usual case when the scale parameter is fixed, it must be estimated, which makes it more difficult to establish the complete convergence of the estimator.

To summarize, the behavior of the developed estimator is unaffected by the number of outliers in the data collection. In comparison to the classical kernel method, the mixture of the kNN algorithm and the robust method allows for a reduction in the impact of outliersin results.

References

[1] Aerts, M., Claeskens, G., Hens, N. and Molenberghs, G. Local multiple imputation. *Biometrika regressor 89*(2002), 375-388.

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

[3] Almanjahie, I., Attouch, M., Kaid, Z. and Louab, H. Robust equivariant nonparametric regression estimators for functional ergodic data. *Comm. Statist. Theory and Methods 50*(2020), 3505-3521.

[4] Attouch, M., Laksaci A. and Ould Sa***i***d, E. Asymptotic distribution of robust estimator for functional non parametric models. *Comm. Statist. Theory and Methods 45*(2009), 230-287.

[5] Attouch, M., Laksaci A. and Ould Sa***i***d, E. Asymptotic normality of a robust estimator of the regression function for functional time series data. *J. of the Korean Statist.Soc. 39*(2010), 489-500.

[6] Attouch, M., Laksaci A. and Ould Sa***i***d, E. Strong uniform convergence rate of a robust estimator of the regression function for functional and dependant processes. *Journal of Japan Statistical Society*, 2012.

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

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

[9] Azzedine, N., Laksaci, A. and Ould Sa***i***d, E. On the robust nonparametric regression estimation for functional regressor. *Statistics & Probability Letters 78*(2008), 3216-3221.

[10] Beirlant, J., Berlinet, A. and Biau, G. Higher order estimation at Lebesgue points. *Ann. Inst. Statist. Math. 90,* 60(2008), 651-677.

[11] Benziadi, F., and Laksaci, A. Recursive kernel estimate of the conditional quantile for functional ergodic data. *Communications in Statistics - Theory and Methods 45*(2015), 3097-3113.

[12] Boente, G. and Fraiman, R. Robust nonparametric regression estimation for dependant observations. *Ann. Statst. 17*(1989), 1242-1256.

[13] Boente, G. and Fraiman, R. Strong order of convergence and asymptotic distribution of nearest neighbor density estimates from dependent observations. ***Sankyä****: The Indian Journal of Statistics 53*(1991), 194-205.

[14] Boente, G. and Rodriguez, D. Robust estimators of high order derivatives of regression function. *Statist. Probab. Lett. 7*(2006), 1335-1344.

[15] Boente, G. and Vahnovan, A. Strong convergence of robust equivariant nonparametric functional regression estimators. *Statistic and Probability Letters 100*(2015), 1-11.

[16] Bosq, D. Linear process in function spaces. Theory and application. Lecture Notes in Statistics, Springer-Verlag 149, 2000.

[17] 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) 16*, 1(2021), 97-116.

[18] Burba, F., Ferraty, F. and Vieu, P. k-nearest neighbour method infunctional nonparametric regression. *J. Nonparametric Statistics 21*, 4(2009), 453-469.

[19] Cheng, J. and Zhang, L. Asymptotic properties of nonparametric M-estimation for mixing functional data. *J. Statist. Plann. Inference 139*(2009), 533-546.

[20] Collomb, G. Estimation nonparamétrique de la régression par la méthode du noyau. PhD dissertation, Universite Paul Sabatier, Toulouse, France, 1976.

[21] Collomb, G. Estimation de la régression par la méthode des k points les plus proches: propriétés de convergence ponctuelle. *Comptes Rendus Acad. Sci. Paris 289, Séries A,*(1979).

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

[23] Collomb, G. and Härdale, H. Strong uniform convergence rates in robust nonparametric time series analysis and prediction: Kernel regression estimation from dependent observations. *Stochastic Processes and their Application 23*(1979), 77-89.

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

[25] Crambes, C. and Laksaci, A. Robust nonparametric estimation for functional data. *Journal of Nonparametric Statistics 20*(2008), 573-598.

[26] Delsol, L. Advances on Asymptotic Normality in Nonparametric Functional Time Series Analysis, Statistics, 43(1) (2009),, 13–33.

[27] Devroye, L. and Wagner, T. The strong uniform consistency of nearest neighbor density. *Ann. Statist. 10*, 5 (1977), 536-540.

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

[29] Devroye, L., Györfi, L., Krzyzak, A. and Lugosi, G. On the stronguniversal consistency of nearest neighbor regression function estimates. *Ann. Statist. 20*, 22(1994), 1371-1385.

[30] Ezzahrioui, M. and Ould Sa$\boldsymbol{i}$d, E. *Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data*, Journal of Nonparametric Statistics, 20 (1)(2008), 3-18.

[31] Ezzahrioui, M., and Ould-Sa$\boldsymbol{i}$d, E. Some Asymptotic Results of a Non-Parametric Conditional Mode Estimator for Functional Time-Series Data, Statistica Neerlandica, 64(2) (2010), 171–201.

[32] Ferraty, F. and Vieu, P. *Nonparametric functional data analysis. Theory and Practice*, Springer-Verlag, New York, 2006.

[33] Ferraty, F., Rabhi, A., and Vieu, P. Conditional Quantiles for Dependent Functional Data with Application to the Climatic El Nino Phenomenon, Sankhya, 67(2) (2005),378–398.

[34] Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. *Rate of uniform consistency for nonparametric estimates with functional variables*, Journal of Statistical Planning and Inference 140 (2010), 335-352.

[35] Ferraty, F. and Romain, Y. *The Oxford handbook of functional data analysis*, Oxford University Press, 2011.

[36] Fraiman Maus, R.J. *Estimadores robustos para regresion no lineal*, Facultad de Ciencias Exactas Naturales. Universidas de Buenos Aires, Phd Thesis, 1980.

[37] Geenens, G. *Curse of dimensionality and related issues in nonparametric functional regression*, Statistics Surveys 5 (2011), 30-43.

[38] Gheriballah, A., Laksaci, A. and Sekkal, S. *Nonparametric M-regression for functional ergodic data*, Journal of Statistics and Probability Letters 83 (2013), 902-908.

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

[40] Hurber, P. Robust estimation of a location parameter, Ann. Math. Statist. 35 (1964), 73-101,

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

[42] Kolmogorov, A. N. and Tikhomirov, V. M. $\boldsymbol{\varepsilon}$*-entropy and* $\boldsymbol{\varepsilon}$*-capacity of sets in function spaces*, Uspekhi Mat. Nauk, 14, 3-86.

[43] Kudraszow, N. L. and Vieu, P. *Uniform consistency of kNN regressors for functional variables*, Journal of Statistics and Probability Letters 83 (2013), 1863-1870.

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

[45] La$\boldsymbol{i}$b, N. and Ouled-Sa$\boldsymbol{i}$d, E. *A robust nonparametric estimation of the autoregression function under an ergodic hypothesis*, Canad. J. Statist. 28 (2000), 817-828.

[46] Laib, N. and Louani, D. *Strong consistency of the regression function estimator for functional stationary ergodic data*, J. Statist. Plann. Inference, 141 (2011), 359-372..

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

[48] Lian, H. *Convergence of functional k-nearest neighbor regression estimate with functional responses*, Electronic Journal of Statistics 5 (133) (2011), 31-40.

[49] Ling, N. and Vieu, P. *Nonparametric Modelling for Functional Data : Selected Survey and Tracks for Future*, *Statistics*, Journal of Nonparametric Statistics 31 (2018), 20-30.

[50] Ling, N., Meng, S. and Vieu, P. *Uniform consistency rate of kNN regression estimation for functional time series data*, Journal of Nonparametric Statistics 31 (2019), 451-468.

[51] Marron, J. M. *Optimal rates on convergence to Bayes risk in nonparametric discrimination*, Ann. Statist. 11 (1983), 1142-1155.

[52] Moore, D. and Yackel, J. *Consistency properties of nearest neighbor density function estimators*, Ann. Statist. 20 (5) (1977), 143–154.

[53] Müller, S. and Dippon, J. *kNN kernel estimate for nonparametric functional regression in time series analysis*, 2011.

[54] Rachdi, M. and Vieu, P. *Nonparametric regression for functional data: automatic smoothing parameter selection*, J. Stat. Plan. Inf. 137 (2006), 2784-2801.

[55] Ramsay, J. and Silverman, B. *Applied functional data analysis: methods and case studies*, Springer, New York, 2002.

[56] Ramsay, J. and Silverman, B *Functional data analysis,* Second Edition, Springer, New York, 2006.

[57] Roussas, G. G. Kernel estimates under association: strong uniform consistency, Statistics & Probability Letters. 12 (9) (1991), 215-224.

[58] Samanta, M. *Non-parametric estimation of conditional quantiles*, Statistics & Probability Letters. 7(5) (1989), 407-412.

[59] Tran, T., Wehrens, R. and Buydens, L. *kNN-kernel density-basedclustering for high-dimensional multivariate data*, Comput. Statist. Data Anal. 51 (81) (2008), 513-525.

[1] Aerts, M., Claeskens, G., Hens, N. and Molenberghs, G. Local multiple imputation. *Biometrika regressor 89*(2002), 375-388.

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

[3] Almanjahie, I., Attouch, M., Kaid, Z. and Louab, H. Robust equivariant nonparametric regression estimators for functional ergodic data. *Comm. Statist. Theory and Methods 50*(2020), 3505-3521.

[4] Attouch, M., Laksaci A. and Ould Sa***i***d, E. Asymptotic distribution of robust estimator for functional non parametric models. *Comm. Statist. Theory and Methods 45*(2009), 230-287.

[5] Attouch, M., Laksaci A. and Ould Sa***i***d, E. Asymptotic normality of a robust estimator of the regression function for functional time series data. *J. of the Korean Statist.Soc. 39*(2010), 489-500.

[6] Attouch, M., Laksaci A. and Ould Sa***i***d, E. Strong uniform convergence rate of a robust estimator of the regression function for functional and dependant processes. *Journal of Japan Statistical Society*, 2012.

[7] Attouch, M. and Bouabsa, W. The $\boldsymbol{k}$-nearest neighbors estimation of the conditional mode for functional data. *Rev. Roumaine Math. Pures Appl. 58,* 4(2013), 393-415.

[8] Attouch, M. and Bouabsa, W. and Chiker el mozoaur, Z. The $\boldsymbol{k}$-nearest neighbors estimation of the conditional mode for functional data under dependency, *International Journal of Statistics & Economics 19,* 1(2018), 48-60.

[9] Azzedine, N., Laksaci, A. and Ould Sa**13̆053′f**d, E. On the robust nonparametric regression estimation for functional regressor. *Statistics & Probability Letters 78*(2008), 3216-3221.

[10] Beirlant, J., Berlinet, A. and Biau, G. Higher order estimation at Lebesgue points. *Ann. Inst. Statist. Math. 90,* 60(2008), 651-677.

[11] Benziadi, F., and Laksaci, A. Recursive kernel estimate of the conditional quantile for functional ergodic data. *Communications in Statistics - Theory and Methods 45*(2015), 3097-3113.

[12] Boente, G. and Fraiman, R. Robust nonparametric regression estimation for dependant observations. *Ann. Statst. 17*(1989), 1242-1256.

[13] Boente, G. and Fraiman, R. Strong order of convergence and asymptotic distribution of nearest neighbor density estimates from dependent observations. ***Sank□yä****: The Indian Journal of Statistics 53*(1991), 194-205.

[14] Boente, G. and Rodriguez, D. Robust estimators of high order derivatives of regression function. *Statist. Probab. Lett. 7*(2006), 1335-1344.

[15] Boente, G. and Vahnovan, A. Strong convergence of robust equivariant nonparametric functional regression estimators. *Statistic and Probability Letters 100*(2015), 1-11.

[16] Bosq, D. Linear process in function spaces. Theory and application. Lecture Notes in Statistics, Springer-Verlag 149, 2000.

[17] 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) 16*, 1(2021), 97-116.

[18] Burba, F., Ferraty, F. and Vieu, P. k-nearest neighbour method infunctional nonparametric regression. *J. Nonparametric Statistics 21*, 4(2009), 453-469.

[19] Cheng, J. and Zhang, L. Asymptotic properties of nonparametric M-estimation for mixing functional data. *J. Statist. Plann. Inference 139*(2009), 533-546.

[20] Collomb, G. Estimation nonparamétrique de la régression par la méthode du noyau. PhD dissertation, Universite Paul Sabatier, Toulouse, France, 1976.

[21] Collomb, G. Estimation de la régression par la méthode des k points les plus proches: propriétés de convergence ponctuelle. *Comptes Rendus Acad. Sci. Paris 289, Séries A*,(1979).

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

[23] Collomb, G. and Härdale, H. Strong uniform convergence rates in robust nonparametric time series analysis and prediction: Kernel regression estimation from dependent observations. *Stochastic Processes and their Application 23*(1979), 77-89.

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

[25] Crambes, C. and Laksaci, A. Robust nonparametric estimation for functional data. *Journal of Nonparametric Statistics 20*(2008), 573-598.

[26] Delsol, L. (2009), Advances on Asymptotic Normality in Nonparametric Functional Time Series Analysis, Statistics, 43(1), 13–33.

[27] Devroye, L. and Wagner, T. The strong uniform consistency of nearest neighbor density. *Ann. Statist. 10*, 5 (1977), 536-540.

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

[29] Devroye, L., Györfi, L., Krzyzak, A. and Lugosi, G. On the stronguniversal consistency of nearest neighbor regression function estimates. *Ann. Statist. 20*, 22(1994), 1371-1385.

[30] Ezzahrioui, M. and Ould Sa$\boldsymbol{i}$d, E. *Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data*, Journal of Nonparametric Statistics, 20 (1), 3-18, 2008.

[31] Ezzahrioui, M., and Ould-Sa$\boldsymbol{i}$d, E. (2010), Some Asymptotic Results of a Non-Parametric Conditional Mode Estimator for Functional Time-Series Data, Statistica Neerlandica, 64(2), 171–201.

[32] Ferraty, F. and Vieu, P. *Nonparametric functional data analysis. Theory and Practice*, Springer-Verlag, New York, 2006.

[33] Ferraty, F., Rabhi, A., and Vieu, P. Conditional Quantiles for Dependent Functional Data with Application to the Climatic El Nino Phenomenon, Sankhya, 67(2), 378–398, 2005.

[34] Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. *Rate of uniform consistency for nonparametric estimates with functional variables*, Journal of Statistical Planning and Inference 140, 335-352, 2010.

[35] Ferraty, F. and Romain, Y. *The Oxford handbook of functional data analysis*, Oxford University Press, 2011.

[36] Fraiman Maus, R.J. *Estimadores robustos para regresion no lineal*, Facultad de Ciencias Exactas Naturales. Universidas de Buenos Aires, Phd Thesis, 1980.

[37] Geenens, G. *Curse of dimensionality and related issues in nonparametric functional regression*, Statistics Surveys 5, 30-43, 2011.

[38] Gheriballah, A., Laksaci, A. and Sekkal, S. *Nonparametric M-regression for functional ergodic data*, Journal of Statistics and Probability Letters 83, 902-908, 2013.

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

[40] Hurber, P. *Robust estimation of a location parameter*, Ann. Math. Statist. 35, 73-101, 1964.

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

[42] Kolmogorov, A. N. and Tikhomirov, V. M. *$\boldsymbol{\varepsilon}$-entropy and $\boldsymbol{\varepsilon}$-capacity of sets in function spaces*, Uspekhi Mat. Nauk, 14, 3-86.

[43] Kudraszow, N. L. and Vieu, P. *Uniform consistency of kNN regressors for functional variables*, Journal of Statistics and Probability Letters 83, 1863-1870, 2013.

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

[45] La$\boldsymbol{i}$b, N. and Ouled-Sa$\boldsymbol{i}$d, E. *A robust nonparametric estimation of the autoregression function under an ergodic hypothesis*, Canad. J. Statist. 28, 817-828, 2000.

[46] La$\boldsymbol{i}$b, N. and Louani, D. *Strong consistency of the regression function estimator for functional stationary ergodic data*, J. Statist. Plann. Inference, 141, 359-372, 2011.

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

[48] Lian, H. *Convergence of functional k-nearest neighbor regression estimate with functional responses*, Electronic Journal of Statistics 5 (133), 31-40, 2011.

[49] Ling, N. and Vieu, P. *Nonparametric Modelling for Functional Data : Selected Survey and Tracks for Future, Statistics*, Journal of Nonparametric Statistics 31, 20-30, 2018.

[50] Ling, N., Meng, S. and Vieu, P. *Uniform consistency rate of kNN regression estimation for functional time series data*, Journal of Nonparametric Statistics 31, 451-468, 2019.

[51] Marron, J. M. *Optimal rates on convergence to Bayes risk in nonparametric discrimination*, Ann. Statist. 11, 1142-1155, 1983.

[52] Moore, D. and Yackel, J. *Consistency properties of nearest neighbor density function estimators*, Ann. Statist. 20 (5), 143–154, 1977.

[53] Müller, S. and Dippon, J. *kNN kernel estimate for nonparametric functional regression in time series analysis*, 2011.

[54] Rachdi, M. and Vieu, P. *Nonparametric regression for functional data: automatic smoothing parameter selection*, J. Stat. Plan. Inf. 137, 2784-2801, 2006.

[55] Ramsay, J. and Silverman, B. *Applied functional data analysis: methods and case studies*, Springer, New York, 2002.

[56] Ramsay, J. and Silverman, B *Functional data analysis*, Second Edition, Springer, New York, 2006.

[57] Roussas, G. G. Kernel estimates under association: strong uniform consistency, Statistics & Probability Letters. 12 (9), 215-224, 1991.

[58] Samanta, M. *Non-parametric estimation of conditional quantiles*, Statistics & Probability Letters. 7(5), 407-412, 1989.

[59] Tran, T., Wehrens, R. and Buydens, L. *kNN-kernel density-basedclustering for high-dimensional multivariate data*, Comput. Statist. Data Anal. 51 (81), 513-525, 2006.