FACTA UNIVERSITATIS (NIŠ)
Ser. Math. Inform. Vol. 38, No 3 (2023), 455 - 472
https://doi.org/10.22190/FUMI220322030B
Original Scientific Paper

LIMIT LAWS OF FUNCTIONAL k-NEAREST NEIGHBORS CONDITIONAL MODE ESTIMATE

Oussama Bouanani[1] and Mustapha Mohammedi[1]

[1]**Université Abdelhamid Ibn Badis de Mostaganem, Algérie,**
Laboratory of Stochastic Models, Statistic and Applications,
Dr. Taher Moulay University of Saida, Algeria
[2] **Université Abdelhamid Ibn Badis de Mostaganem, Algérie,**
Laboratory of Statistics and Stochastic Processes,
Djillali Liabes University of Sidi Bel Abbes, Sidi Bel Abbes, Algeria

Abstract. This research study the k-Nearest Neighbors (k-NN) of the conditional mode by using the local linear technique in the case of a scalar response variable that are linked via a single-index structure. The main contribution of this work is to investigate the asymptotic normality of the nonparametric estimator of the conditional mode under some general assumptions. A simulation study was conducted to assess finite sample behavior to exhibit the effectiveness of our method over the standard kernel method.
Keywords: Local constant estimator, Local linear estimation, Asymptotic normality, Conditional mode, Functional data, k-Nearest Neighbors (k-NN), Single functional index model.

1. Introduction

In the last decade, functional data has emerged as one of the primary areas in the field of statisticalo analysis, which is due to the development of modern measuring devices that allow for more extensive data to be collected. The evidence of this interest is the numerous practical applications in medicine, econometrics,

Received March 22, 2022. accepted October 12, 2023.
Communicated by Miroslav Ristić
Corresponding Author: Oussama Bouanani, Université Abdelhamid Ibn Badis de Mostaganem, Algérie, Laboratory of Stochastic Models, Statistic and Applications, Dr. Taher Moulay University of Saida, 20000 Saida, Algeria | E-mail: oussama.bouanani@univ-mosta.dz
2010 *Mathematics Subject Classification.* 62G05. 62G20. 62G30. 62N01. 62N02

chemometrics, etc., as well as the publication of numerous scientific papers on this area. Here we refer to Dauxois et *al.* [18] , Silverman [33], Bosq [11] for parametric models and Ferraty and Vieu [23] for functional nonparametric models.

The estimation of the non parametric conditional mode is an alternative method to estimate the conditional regression and has been extensively used in forecasting field. In view of this, several authors have dealt with this model in their investigations (see Ferraty et al. [22], Ould-Saïd and Tatachak [27] and Khardani et al. [25] and more recently Dabo-Niang et al. [16], for more discussion and motivation).

On the other hand, the Single functional index model (SFIM) is the subject of numerous studies on nonparametric functional data (NFDA), particularly in the field of econometrics. This statistical model is a very simple approach as it reduces the dimension of space (Curse of dimensionality) ensuring some flexibility. The first work in this topic was introduced by Ferraty et al. [21] for regression problems. For more results in this context (see for instance, Ait-Saïdi et al. [1], Attaoui and Ling [7], Goia and Vieu [24], Ling and Xu [30], Sang and Cao [35].)

The novelty of this paper stems from the combination of two approaches, LLE and k-NN, which we will refer to as k-NN-LLE. Furthermore, this method allows us to construct an attractive, robust estimator that converges quickly and is extremely easy to be implemented in practice.

Note that the first technique that we are going to call local linear estimator (LLE) is an alternative to local constant estimator (LCE) (classical kernel estimate or Nadaraya Watson's estimator NWE). However, the latter approach contains some anomalies and this is due to the lack of local variations of the smoothing, which gives us the problem of large bias term at the boundaries of data (boundary effects). (see for instance, Ruppert and Wand [34] and Fan and Gijbels [20] in the multivariate case, and more recently, see Laksaci et al. [29] in the FDA setup among others). For these reasons, the use of the local polynomial method is preferable to the constant local approach. In fact, the variance term of the two methods are the same, however in the infinite-dimensional case, the dispersion term depends on the probability measure of the functionally explanatory variable $\phi_x(h)$ and this function plays an important role in the convergence rate. However, the bias term at the boundaries of the data by using local linear estimation of k-nearest neighbors is of order $O\left(\phi_x^{-1}(k/n)^2\right)$), while by Nadaraya Watson's estimator it is of order $O\left(\phi_x^{-1}(k/n)\right)$) and this comparison indicates that the bais term by (k-NN-LLE) is smaller than the kernel method.

Notice that the local linear smoothing in functional data analysis (FDA) has only recently been discussed. In fact, the almost-complete convergence of the local linear estimator of the conditional mode for independent and identically distributed (i.i.d.) data was obtained by Demongeot et *al.* [17].However, Bouanani et *al.* [9], they established the asymptotic normality of the conditional mode and this last study was extended to the dependent functional data case (under strong mixing conditions) by Bouanani et *al.* [10].

The second technique, that we are going to call k-NN, is an attractive method for nonparametric estimation in infinite dimensional contexts and can take into account the local structure of the data. The motivation to use this method of k-NN is that the local bandwidth of the k-NN is random and depends on the data, and the user has only one parameter to control and take its values in a finite set. The initiator of this approach goes back to [15] in a finite-dimensional, and is generalized in the functional data framework by [12], [13], [32] and [8]. The most recent references are listed in the citations: Chikr Elmezouar et al. [14], Almanjahie et al. [2], [3], [4], [5] Alshahrani et al. [6], Mohammedi et al. [31].

The paper is organized as follows. In the following section, we introduce the model and the k-NN Local Linear Estimator (k-NN-LLE) of the conditional mode. The notations and hypotheses are given in the section 3. Under some conditions, we establish the asymptotic normality of the k-NN-LLE in Section 4. The implementation of this model in practice is given in Section 5. where we compare between the k-nn-LLE and the local constant estimator (LCE) (classical kernel estimate) by simulation data. Finally, the last section 6. is dedicated to the proofs of the results.

2. Model and estimator

Consider n independent pairs of random variables (X_i, Y_i) for $i = 1, ..., n$, we assume that they are drawn from the pair (X, Y). The latter is valued in $\mathbb{F} \times \mathbb{R}$, where $\mathbb{F}$ is a seperable Hilbert space with the norm $|| \ . \ ||$ generated by an inner product $< .,. >$. We consider the semi-metric d_θ , associated to the single-index $\theta \in \mathfrak{F}$ defined by $\forall u, v \in \mathfrak{F} : d_\theta(u, v) := |< u - v, \theta >|$. Under such topological structure and for a fixed functional θ, we suppose that the conditional cumulative distribution function(CDF) of Y given$X = x$ has a single-index structure $\theta \in \mathfrak{F}$ denoted by $F_\theta^x(.)$ exists and is given by :

$$\text{for all } y \in \mathbb{R}, \quad F_\theta^x(y) = \mathbb{P}(Y \leq y \, |< X = x, \theta >).$$

This distribution is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and has bounded density, denoted by g_θ^x.

We consider the conditional distribution function (CDF) $F_\theta^x(y)$ as a regression model with the response variable $\delta(h_\delta^{-1}(y - Y_i))$, where δ is some cumulative distribution function and $(h_\delta = h_{\delta,n})$,is a sequence of positive real numbers. (See for instance, Fan [19] in the finite dimensional case, and Rachid et **al**. [26] in functional setting)), i.e.,

$$\mathbb{E}\left(\delta(h_\delta^{-1}(y - Y_i)) \, |< X_i = x, \theta >\right) \longrightarrow F_\theta^x(y) \ \text{ as } \ h_\delta \longrightarrow 0.$$

In the following part, we define the local linear estimator denoted by $\widehat{F}_\theta^x(y)$ of $F_\theta^x(y)$ obtained by the k-Nearest Neighbors (k-NN) approach based on the single-index model by

$$(2.1)\qquad \widehat{F}_\theta^x(y) = \frac{\sum_{i=1}^{n}\sum_{i\neq j,j=1}^{n}\omega_{\theta,ij}\delta(h_\delta^{-1}(y-Y_j))}{\sum_{i=1}^{n}\sum_{i\neq j,j=1}^{n}\omega_{\theta,ij}} = \frac{\sum_{j=1}^{n}\Omega_{\theta,j}K_j\Gamma_j}{\sum_{j=1}^{n}\Omega_{\theta,j}K_j},$$

where $w_{\theta,ij} = \beta_{\theta,i}\left(\beta_{\theta,i}-\beta_{\theta,j}\right)K_iK_j$, $\Omega_{\theta,j} = K_j^{-1}\left(\sum_{i=1}^{n}w_{\theta,ij}\right) = \sum_{i=1}^{n}\beta_{\theta,i}^2K_i - \left(\sum_{i=1}^{n}\beta_{\theta,i}K_i\right)\beta_{\theta,j}$, with $K_i = K\left(\mathbb{H}_k^{-1}\left(d_\theta(x,X_i)\right)\right)$ and $\delta_j = \delta\left(h_\delta^{-1}(y-Y_j)\right)$.

$\beta_{\theta,i} = \beta_\theta(X_i,x)$ is a known bi-functional operator from $\mathfrak{F}\times\mathfrak{F}$ into $\mathbb{R}^+$, K is a kernel, δ is a distribution function and ($\mathbb{H}_k = \mathbb{H}_{k,n}$) is a positive random variable, defined as follows:

$$(2.2)\qquad \mathbb{H}_k(x) = min\left\{h\in\mathbb{R}^+/\sum_{i=1}^{n}1\!\!1_{B(x,h)}(X_i) = k\right\}.$$

with

$$B(x,h) = \left\{x^{'}\in\mathfrak{F},\ \ d_\theta(x,x^{'}) < h\right\}.$$

Moreover, the estimation of the the conditional density function of Y given X has a single-index structure $\theta\in\mathfrak{F}$, is defined by:

$$\forall y\in\mathbb{R},\ \ g_\theta^x(y) = g(y\,|< X = x,\theta >).$$

According to the same conditions used by Ferraty et al.[21], we suppose that this model is differentiable with respect to x and θ such that $<\theta, e_1>= 1$, where e_1 is the first vector of an orthonormal basis of $\mathfrak{F}$, then under this condition, the identifiability of this model is assured . In other ways , for all $\forall x\in\mathfrak{F}$, we have :

$$g_1(y\,|< .,\theta_1 >) = g_2(y\,|< .,\theta_2 >)\ \ \text{implies that}\ \ g_1 = g_2\ \ \text{and}\ \ \theta_1\equiv\theta_2.$$

From Eq 2.1, we deduce an estimator for the conditional density with (LMM-kNN) approach, defined by:

$$(2.3)\qquad \widehat{g}_\theta^x(y) = \frac{\sum_{i=1}^{n}\sum_{i\neq j,j=1}^{n}w_{\theta,ij}\delta^{(1)}(h_\delta^{-1}(y-Y_j))}{h_\delta\sum_{i=1}^{n}\sum_{i\neq j,j=1}^{n}w_{\theta,ij}} = \frac{\sum_{j=1}^{n}\Omega_{\theta,j}K_j\delta_j^{(1)}}{h_\delta\sum_{j=1}^{n}\Omega_{\theta j}K_j},$$

where $\delta^{(1)}$ denotes the first derivative of δ.

In our context, we focus on the estimation of the conditional mode in the functional single index model denoted by μ_θ. Note that this estimate is not necessarily unique, then for a given x, we chosen the compact set $S = [\mu_\theta(x)-\xi, \mu_\theta(x)+\xi]$ such that the conditional density of Y at X = x has a unique mode $\mu_\theta(x)$ in S, which is defined by

$$\mu_\theta(x) = \sup_{y\in S} g_\theta^x(y)$$

A local linear estimator of $\mu_\theta(x)$ based on the k-Nearest Neighbors approach is defined by

$$\widehat{\mu}_\theta(x) = \sup_{y\in S} \widehat{g}_\theta^x(y). \tag{2.4}$$

3. Hypotheses and notations

Let x (resp. y) be a fixed point in $\mathfrak{F}$ (resp. in $\mathbb{R}$), $\mathcal{N}_x$ (resp. $\mathcal{N}_y$) denotes a fixed neighborhood of x (resp. of y,) and $\phi_{\theta,x}(h) = \mathbb{P}(\mid \mu(x) \mid\leq h)$, where $|\mu(x)| = d_\theta(x,X)$. Morever, we denote the closed-ball the ball of center x and radius r by $B_\theta(x,h) := \{x' \in \mathfrak{F} : \mid < \theta, x-x' > \mid \leq h\}$,
We will denote by C and C' some strictly positive constants.
The first condition is focused on the distribution of the functional variables and the regularity of the model.

(H1) (i) For any $h>0$, $\phi_{\theta,x}(h) := P(X \in B_\theta(x,h)) > 0$, is an invertible function and there exists a function $\Psi_{\theta,x}(\cdot)$ such that:

$$\forall t \in [0,1],\ \lim_{h\to 0} \frac{\phi_{\theta,x}(th)}{\phi_{\theta,x}(h)} = \Psi_{\theta,x}(t).$$

(ii) For $j = \{0,1,2\}$, $g_\theta^{x(j)}(y)$ satisfies that there exist some positive constants b_1 and b_2, such that:

$$\mid g_\theta^{x_1(j)}(y_1) - g_\theta^{x_2(j)}(y_2) \mid\leq C_{\theta,x}\left(d_\theta^{b_1}(x_1,x_2)+\mid y_1-y_2\mid^{b_2}\right).$$

(H2) On the flatness of the conditional density $g_\theta^{(x)}$

(i) There exists some integer $j>0$, such that, g_θ^x is j-times continuously differentiable with respect to y on $(\mu_\theta(x)-\xi, \mu_\theta(x)+\xi)$ for $\xi>0$.

(ii) $$\begin{cases} g_\theta^{x(l)}(\mu_\theta(x)) = 0 \quad \text{if} \quad 1 \leq l < j, \\ \\ |g_\theta^{x(j)}(\mu_\theta(x))| > 0. \end{cases}$$

(H3) On the locating function $\beta_\theta(.,.)$.

$$\begin{cases} \text{for all } z \in \mathfrak{F}, \;\; C \mid d_\theta(x,z) \mid \leq \mid \beta_\theta(x,z) \mid \leq C' \mid d_\theta(x,z) \mid, \\ \\ \sup_{v \in B_\theta(x,r)} |\beta_\theta(v,x) - d_\theta(x,v)| = o(r), \\ \\ h \int_{B_\theta(x,h)} \beta_\theta(v,x) d\mathbb{P}_X(v) = o\left(\int_{B_\theta(x,h)} \beta_\theta^2(v,x)\, d\mathbb{P}_C(v) \right), \end{cases}$$

where $dP(v)$ is the probability distribution of X.

(H4) On the kernels K and δ.

(i) The kernel K is a positive bounded kernel of class C^1 over its support $[0,1]$ and for which the first derivative K' satisfies: $K(1) > 0$, $K'(t) < 0$, for $t \in [0,1]$.

(ii) The kernel function δ is a 3-times continuously differentiable function which is monotonous on its compact support and such that

$$\int \delta^{(1)}(t) = 1, \int t\delta^{(1)}(t) = 0 \text{ and } \int \mid t \mid^{b_2} \delta^{(1)}(t) dt < \infty.$$

(H5) On the bandwidth $h_\delta = h_{\delta,n}$

$$\lim_{n\to\infty} h_\delta = 0, \quad \lim_{n\to\infty} n h_\delta \phi_{\theta,x}(h) = \infty \;\; and \;\; \lim_{n\to\infty} \frac{log(n)}{n h_\delta \phi_{\theta,x}(h)} = 0.$$

(H6) On the parameter $k = k_n$ of neighbors:
The sequence of positive real numbers k satisfies,

$$\lim_{n\to\infty} \phi_{\theta,x}^{-1}\left(\frac{k}{n}\right) = 0, \quad \lim_{n\to\infty} n\, \phi_{\theta,x}^{-1}\left(\frac{k}{n}\right) = \infty \;\; and \;\; \lim_{n\to\infty} \frac{log(n)}{n\phi_{\theta,x}^{-1}(k/n)} = 0.$$

4. Main Results

Theorem 4.1. *Under assumptions (H1)-(H6) and if the k-NN parameter k satisfies* $\lim_{n\to\infty} \sqrt{k h_\delta^3} \left(\phi_x^{-1} \left(\frac{k}{n}\right)^{b_1} + h_\delta^{b_2} \right) = 0$, *we have*

$$\left(\frac{k\, h_\delta^3 (g_\theta^{x(2)}(\mu_\theta(x)))^2}{\sigma_\theta^2(x,\mu_\theta(x))} \right)^{1/2} (\widehat{\mu}_\theta(x) - \mu_\theta(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{4.1}$$

where

$$\sigma_\theta^2(x,\mu_\theta(x)) = \frac{M_2\, g_\theta^x(\mu_\theta(x))}{M_1^2}\int(\delta^{(2)}(t))^2)^2dt,$$

withe $M_j = K^j(1) - \int_0^1 (K^j(u))'\Psi_{(\theta,x)}(u)du,$ *for* $j=1,2,$

$$\mathcal{A} = \{x\in\mathfrak{F}, g_\theta^{x(2)}(\mu_\theta(x))g_\theta^x(\mu_\theta(x))\neq 0\},$$

$\xrightarrow{\mathcal{D}}$ *denoting the convergence in distribution.*

Proof. To simplify the proofs of our results let us note

$$\widehat{g}_{\theta,N}^x(y) = \frac{1}{n\,h_\delta\;\mathbb{E}(\Omega_{\theta,1}K_1)}\sum_{j=1}^n \Omega_{\theta,j}K_j\delta_j^{(1)},\quad \widehat{g}_{\theta,D}^x = \frac{1}{n\,\mathbb{E}(\Omega_{\theta,1}K_1)}\sum_{j=1}^n \Omega_{\theta,j}K_j,$$

where

$$\widehat{g}_\theta^x(y) = \frac{\widehat{g}_{\theta,N}^x(y)}{\widehat{g}_{\theta,D}^x}.$$

Based on the Taylor expansion of $\widehat{g}_\theta^{x(1)}(.)$ in the neighborhood of $\mu_\theta(x)$ and according to the assumptions (H2), we have

$$\widehat{\mu}_\theta(x) - \mu_\theta(x) = -\frac{\widehat{g}_\theta^{x(1)}(\mu_\theta(x))}{\widehat{g}_\theta^{x(2)}(\bar{\mu_\theta}(x))} = -\frac{\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x))}{\widehat{g}_{\theta,N}^{x(2)}(\bar{\mu_\theta}(x))}, \tag{4.2}$$

where $\bar{\mu_\theta}(x)$ is between $\widehat{\mu_\theta}(x)$ and $\mu_\theta(x)$. Then, we can write :

$$\sqrt{k\,h_\delta^3}\,(\widehat{\mu}_\theta(x) - \mu_\theta(x)) \quad = \quad -\frac{\sqrt{k\,h_H^3}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x)) - \mathbb{E}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x))\right)\right) + \sqrt{k\,h_H^3}\mathbb{E}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x))\right)}{\widehat{g}_{\theta,N}^{x(2)}(\bar{\mu}_\theta(x))}.$$

Then, the rest of the proof of this theorem is based on the following lemmas for which proofs are given in the appendix 6.

Lemma 4.1. *Under the condition of Theorem 4.1 and by assumptions (H2)-(H4), we have:*

$$\sqrt{k\,h_H^3}\mathbb{E}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x))\right)\xrightarrow{\mathcal{P}}0,\ \ as\ \ n\longmapsto\infty,$$

Lemma 4.2. *Under the assumptions of Theorem 4.1, we have*

$$\sqrt{k\,h_H^3}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x)) - \mathbb{E}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x))\right)\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\sigma_\theta^2(x,\mu_\theta(x)),\ \ as\ \ n\longmapsto\infty,$$

Lemma 4.3. *Under the conditions of Theorem 4.1, we have:*

$$\widehat{g}_{\theta,N}^{x(2)}(\bar{\mu}_\theta(x))\xrightarrow{\mathcal{P}}g_\theta^{x(2)}(\mu_\theta(x)),\ \ as\ \ n\longmapsto\infty,$$

5. Simulation

Our main goal of this application is to display the usefulness of the conditional mode in the functional single index model of Y given $< x, \theta >$ in a prediction context. More precisely, we illustrate the performance and the superiority of our estimator by using the criteria the mean square error (MSE). For this aim we compare the (MSE) of the kNN local linear approach (the k-NN-LLE) is defined in 2.1 , over the local constant estimator (L.C.E) (classical kernel method) is defined by

$$\mu(x) = \sup_{y \in S} g^x_{L.C.E}(y),$$

where

$$\widehat{g}^x_{L.C.E}(y) = \frac{h_\delta^{-1} \sum_{i=1}^{n} K(\mathbb{H}_{k,n}^{-1} d(x, X_i)) \delta^{(1)}(h_\delta^{-1}(Y_i - y))}{\sum_{i=1}^{n} K(\mathbb{H}_{k,n}^{-1} d(x, X_i))}. \tag{5.1}$$

For this purpose, we generated the functional covarite X on the interval $[0, \pi]$ (see figure 5.1) by the following process:

$$X_i(t) = cos(W_i t) - (W_i + t), \text{ for } i = 1, 2, .., 100.$$

where $W_i \rightsquigarrow N(0, 0.3)$. All the curves X_i 's are generated from 100 equidistant values in $[0, \pi]$.

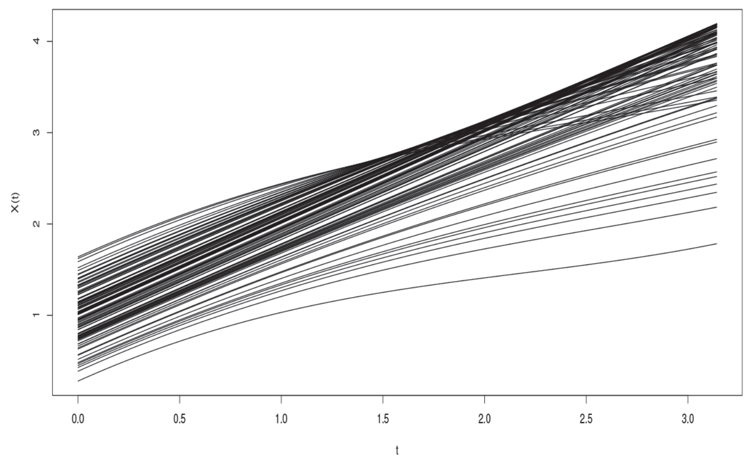

Fig. 5.1: A sample of 100 curves

5.1. On the smoothing parameters (h_Γ and k)

The bandwidth parameters are very crucial in nonparametric estimation because, as in the case we are studying, these parameters interfere with all asymptotic properties, especially in improving the convergence rate. In this application, we use the cross-validation (CV) method to select the h and k bandwidths. For the similar technique used by Rachdi et al.[19] is used. We consider minimizing the squared error in the local linear estimation of the conditional density for functional data defined by the following criterion

$$\frac{1}{n}\sum_{i=1}^{n} W_1(X_i)\int\left[\widehat{\mu}_\theta^{(-i)}(y|X_i)\right]^2 W_2(y)dy - \frac{2}{n}\sum_{i=1}^{n}\left[\widehat{\mu}_\theta^{(-i)}(y|X_i)\right] W_1(X_i)W_2(Y_i),$$

where, $\widehat{\mu}_\theta^{(-l)}(x) = \sup_{y\in S}\widehat{g}_\theta^{(-l)}(y \mid x)$ is called the leave-one-out curve estimator, which is expressed as the following

$$\widehat{g}_\theta^{(-l)}(y|X_l) = \frac{h_\delta^{-1}\displaystyle\sum_{j=1,j\neq l}^{n}\sum_{i=1,i\neq l}^{n}\omega_{\theta,ij}(X_l)\delta^{(1)}(h_\Gamma^{-1}(y-Y_j))}{\displaystyle\sum_{j=1,j\neq l}^{n}\sum_{i=1,i\neq l}^{n}\omega_{\theta,ij}(X_k)}. \tag{5.2}$$

In our simulation study, we take $W_1() = 1$ and $W_2(y) = \mathbb{I}_{[0.9\times min_{i=1,..,n}(Y_i);1.1\times min_{i=1,..,n}(Y_i)]}$, (see, for instance, Laksaci et al. [28]) for more discussions on the choice about weight function.)

5.2. On the the single functional index model

The single functional index can be found in many areas of applied science, particularly econometrics, where it has demonstrated advantages such as reducing the dimension of space (curse of dimensionality) and proposing an interpretation of the results by estimating this function parameter. In practice, this parameter is unknown, then by following the same idea as those in Attaoui et *al.* [7], we select the single functional index model θ as follows :

- Step 1. We compute the mean curve of $(X_i)_{i=1,..,100}$ by the routine func.mean of the R-package fda.usc.

- Step 2. We Compute the covariance operator $\mathbb{E}[(X-\mathbb{E}(X) < X, . >_{\mathfrak{F}}]$ by using the empirical covariance operator in the sample $\varpi = \{1, ..., 100\}$ and $|\varpi| = 100$.

$$\frac{1}{|\varpi|}\sum_{i\in\Lambda}(X_i(s)-\overline{X}(s))^t(X_i(s)-\overline{X}(s)),\ \ where\ \overline{X}(s) = \frac{1}{|\varpi|}\sum_{i\in\Lambda}X_i(s), \tag{5.3}$$

- Step 3. We calculate the eigenvectors of (5.3) (empirical covariance operator).

The obtained results are shown in the following graphs 5.2

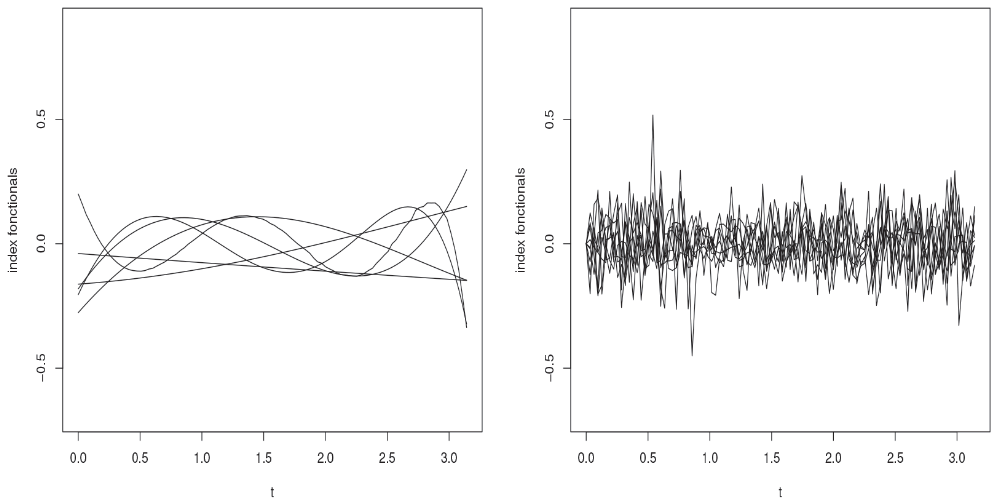

FIG. 5.2: The curve on the left represents the eigenfunctions $\theta_i(t_j),\ t_j \in [0,\pi],$ for $i = 7, 8.., 15$, (resp. for$i = 1, 2, .., 6$)

To illustrate the performance of our estimator, we proceed the following algorithm

- Step 1. We choose θ^* the first eigenfunction corresponding to the first higher eigenvalue.
- Step 2. We Compute the inner product

$$< \theta^*, X_1 >, < \theta^*, X_2 >, ..., < \theta^*, X_{100} >$$

.

- Step 3. We generate the response variables Y_i by the following relation

$$Y_i = r\left(< \theta^*, X_i >\right) + \epsilon_i, \text{ where } r\left(< \theta^*, X_i >\right) = e^{(< \theta^*, X_i >)},$$

  and ϵ_i simulate independently and follow the normal distribution N (0, 0.1).
- Step 4. Because of the nature of the data (the shape of the curves.(5.1), we chose the following family of locating function β_θ

$$\beta_\theta(x_1, x_2) = \sqrt{\int_0^1 \theta^*(t)(x_1(t) - x_2(t))^2 dt},$$

and

$$d(x_1, x_2) = \sqrt{\int_0^1 (x_1(t) - x_2(t))^2 dt}.$$

- Step 5. We divide our observations into two subsets :
  - $(X_i, Y_i)_{i=1,..,80}$, training sample.
  - $(X_j, Y_j)_{j=81,..,100}$, test sample.
- Step 6. We choose a quadratic kernel K on $[0,1]$ and take $K = \delta^{(1)}$.
- Step 7. For each j in the test sample, we compute $\widehat{Y}_j = \widehat{\mu}_{\theta^*}(X_j)$ by using the two approach (L.C.E) and (k-NN-LLE.
- Step 8. We present our results by plotting the boxplot of the prediction error are represented in (Figure 5.3) and we compute the empirical mean square error with k-NN-LLE (resp. L.C.E) :
  - MSE= $\dfrac{1}{20}\displaystyle\sum_{i=1}^{20}(Y_i - \widehat{\mu}_{\theta^*}^{X_i}(Y_i))^2 = 0.007.$
  - MSE $= \dfrac{1}{20}\displaystyle\sum_{i=1}^{20}(Y_i - \widehat{\mu}_{LCE}^{X_i}(Y_i))^2 = 0.074.$

Based on the Figure (5.3), the kernel method results are given on the center, while the right part of Figure (5.3) presents the kNN local linear method (k-NN-LLE). Then, we remark that the performance of the prediction is controlled by the continuous line, in the sense that the efficiency of the prediction method is quantified by the closeness of the dark point to this continuous line. Clearly, the comparison results in (5.3) indicate that the method based on the local linear polynomial estimation is much better and more efficient than the kernel method (LCE) . This is confirmed by the mean squared error MSE(LCE) = 0.074 whereas MSE(k-NN-LLE) = 0.007.

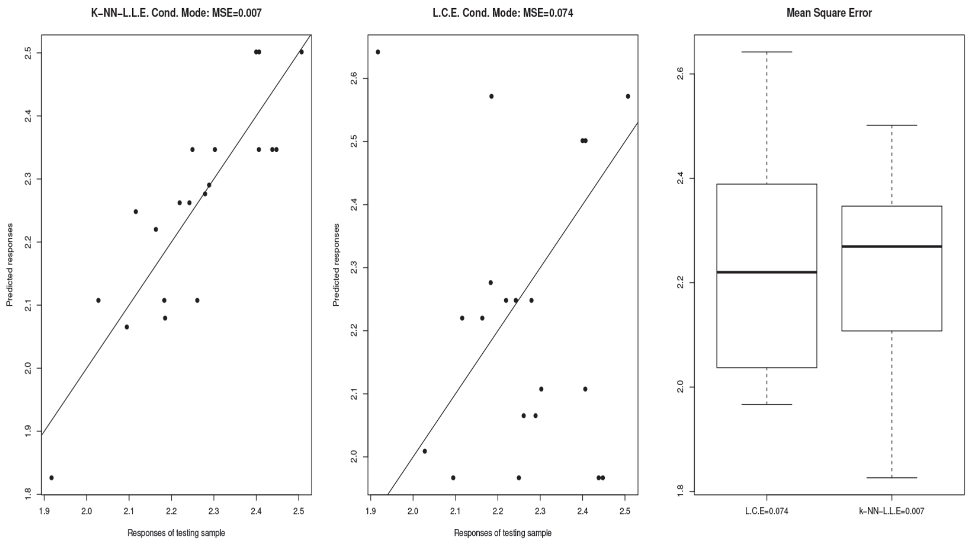

FIG. 5.3: Comparison results betwenn the k-NN-local linear estimator and the classical estimator (LCE)

6. Appendix

In our context, the smoothing parameter $\mathbb{H}_k$ is a random variable, which complicates the proofs of our results. To solve this problem, the idea is to frame $\mathbb{H}_k$ meaningfully by two non-random bandwidth parameters.

Proof. of Lemma 4.1

By applying lemma (3.2) of burba et *al.* [12], we denote:

$$C_{n,\theta}(D_n) = \widehat{g}_\theta^{x(1)}(y),\ \ c = g_\theta^{x(1)}(y) \text{ and } D_n = \mathbb{H}_{n,k},$$

then we can write:

$$C_{n,\theta}(D_n) = \frac{C^1_{n,\theta}(D_n)}{C^0_{n,\theta}(D_n)},$$

where

$$C^r_{n,\theta}(D_n) = \frac{1}{nh_\delta^{2r}\,\mathbb{E}(\Omega_{\theta,1}K_1)} \sum_{j=1}^{n} \Omega_{\theta,j} K_j \left(\delta_j^{(2)}\right)^r, \quad \text{for } \ r \in \{0,1\}$$

By using the fact that $\mathbb{1}_{D_n^- \leq D_n \leq D^+} \overset{a.co.}{\longrightarrow} 1$ when $\dfrac{k}{n} \longrightarrow 0$ (see [12]), we have:

$$\mathcal{C}^j_{n,\theta}(D_n^+) \leq \mathcal{C}^j_{n,\theta}(x, D_n) \leq \mathcal{C}^j_{n,\theta}(D_n^-), \ \text{ for } \ j \in \{0,1\}. \tag{6.1}$$

Then

$$\mathbb{E}(C^1_{\theta,n}(D_n^+)) \le \mathbb{E}(C^1_{\theta,n}(D_n)) \le \mathbb{E}(C^1_{\theta,n}(D_n^-))$$

For the expectation term we obtain:

$$\mathbb{E}(C^1_{\theta,n}(D_n^-)) = \frac{1}{\mathbb{E}(\Omega_{\theta,1}K_1)}\mathbb{E}\left(\Omega_{\theta,1}K_1\left(\mathbb{E}\left(h_\delta^{-2}\delta^{(2)}(y) \mid < X = x, \theta >\right)\right)\right).$$

Morever, by using an integration by parts and by a change of variable $t = \dfrac{\mu_\theta(x) - u}{h_\delta}$, we obtain:

$$\begin{aligned}\mathbb{E}\left(\delta^{(2)}\left(\frac{\mu_{\mu_{\mu_\theta}}(x) - Y_1}{h_\delta}\right) \mid < X_1, \theta >\right) &= \int_{\mathbb{R}} \delta^{(2)}\left(\frac{\theta(x) - u}{h_\delta}\right) g_\theta^x(u)du \\ &= h_\delta^2 \int_{\mathbb{R}} \delta^{(1)}(t) g_\theta^{x(1)}(\mu_\theta(x) - th_\delta)dt.\end{aligned}$$

Now, by using a two order Taylor expansion of $g_\theta^{x(1)}$ around $\theta(x)$ we can deduce that

$$\begin{aligned}\sqrt{kh_\delta^3}E(g_N^{x(1)}(\mu_\theta(x))) &= \sqrt{kh_\delta^3}\int_{\mathbb{R}}\delta^{(2)}(t)\left[g_\theta^{x(1)}(\mu_\theta(x))dt\right] \\ &-\sqrt{kh_\delta^3}\int_{\mathbb{R}}\delta^{(2)}(t)\left[th_\delta g_\theta^{x(2)}(\mu_\theta(x))dt\right] \\ &+\sqrt{kh_\delta^3}\int_{\mathbb{R}} h_\delta^2 t^2 \delta^{(2)}(t) g_\theta^{x(3)}(\bar{\mu}_\theta(x))dt.\end{aligned}$$

Then 4.1 is a simple consequence of this last result with the condition of Theorem 4.1 and under assumptions (H2) and (H4).

Proof. of Lemma 4.2.

$$\begin{aligned}\sqrt{kh_\delta^3}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x)) - \mathbb{E}(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x)))\right) &= \sqrt{kh_\delta^3}\left(C^1_{n,\theta}(D_n) - \mathbb{E}(C^1_{n,\theta}(D_n))\right) \\ \text{(6.2)} \qquad &= \underbrace{\sqrt{kh_\delta^3}\left(C^1_{n,\theta}(D_n^+) - \mathbb{E}(C^1_{n,\theta}(D_n^+))\right)}_{I_{1,\theta}} \\ &- \underbrace{\sqrt{kh_\Gamma}\left(\mathbb{E}(C^1_{n,\theta}(D_n) - \mathbb{E}(C^1_{n,\theta}(D_n^+))\right)}_{I_{2,\theta}} \\ &+ \underbrace{\sqrt{kh_\delta^3}\left(C^1_{n,\theta}(D_n) - C^1_{n,\theta}(D_n^+)\right)}_{I_{3,\theta}}.\end{aligned}$$

Then, the decomposition (6.2) becomes:

$$\sqrt{kh_\delta^3}\left(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x)) - \mathbb{E}(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x)))\right) = I_{1,\theta} + I_{2,\theta} + I_{3,\theta}.$$

In order to establish the asymptotic normality of formula (6.2), we apply Slutsky's Theorem which gives us the following formulas

$$I_{1,\theta} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_\theta(x, \mu_\theta(x))), \quad as \;\; n \longrightarrow \infty. \tag{6.3}$$

$$I_{2,\theta} \xrightarrow{\mathcal{P}} 0, \;\; , as \;\; n \longrightarrow \infty. \tag{6.4}$$

$$I_{3,\theta} \xrightarrow{\mathcal{P}} 0, \quad as \;\; n \longrightarrow \infty. \tag{6.5}$$

Proof. of Eq 6.3

Was proven in Corollary 4.1 of Bouanani et *al* . [9] by choosing the bandwidth parameters as as $h_K = D_n^+$.

Proof. of Eq 6.4

For the term $I_{\theta,2}$ with $D_n = \mathcal{H}_{n,k}$, we use again equation 6.1. This leads:

$$\begin{aligned}
\sqrt{kh_\delta^3}|\mathbb{E}(C^1_{n,\theta}(D_n)) - \mathbb{E}(C^1_{n,\theta}(D_n^+))| \;\; &\leq \;\; \sqrt{kh_\delta^3}|\mathbb{E}(C^1_{n,\theta}(D_n^-)) - \mathbb{E}(C^1_{n,\theta}(D_n^+))| \\
&\leq \;\; \underbrace{\sqrt{kh_\delta^3}|\mathbb{E}(C^1_{n,\theta}(D_n^-)) - c\,|}_{E_1} + \underbrace{\sqrt{kh_\Gamma}|\mathbb{E}(C^1_{n,\theta}(D_n^+)) - c|}_{E_2}.
\end{aligned}$$

Morever, by applynig Lemma1 of Laksaci et *al.* [29], we have

$$|\; g_\theta^{x(1)}(y) - \mathbb{E}(\widehat{g}_{\theta,N}^{x(1)}(y)) \;| = O(h_K^{b_1}) + O(h_\delta^{b_2})$$

By chossing the bandwidth parameter as $h_K = D_n^+$, where (D_n^-) and (D_n^+) such that:

$$\begin{cases}
\phi_{\theta,x}((D_n^+)) = \sqrt{\zeta_n}\phi_{\theta,x}(h) = \sqrt{\zeta_n}\dfrac{k}{n} \\[2ex]
\phi_{\theta,x}((D_n^-)) = \dfrac{1}{\sqrt{\zeta_n}}\phi_{\theta,x}(h) = \dfrac{1}{\sqrt{\zeta_n}}\dfrac{k}{n}
\end{cases} \tag{6.6}$$

besides that $\zeta_n \in]0,1[$ is bounded by 1, we get:

$$|\; g_\theta^{x(1)}(\mu_\theta(x)) - \mathbb{E}(\widehat{g}_{\theta,N}^{x(1)}(\mu_\theta(x)) \;| = O\left(\phi_{\theta,x}^{-1}\left(\frac{k}{n}\right)^{b_1} + h_\delta^{b_2}\right).$$

Thus, we get

$$\sqrt{kh_\delta^3}(E_1 + E_2) = 2\sqrt{kh_\delta^3}\; O\left(\phi_{\theta,x}^{-1}\left(\frac{k}{n}\right)^{b_1} + h_\delta^{b_2}\right).$$

Consequently , by using the condition of Theorem 4.1, we get

$$\sqrt{kh_\delta^3}\ \mathbb{E}(C^1_{n,\theta}(D_n)) - \mathbb{E}(C^1_{n,\theta}(D_n^+)) \longrightarrow 0 \ \ as \ \ n \longrightarrow \infty.$$

The proof of Claim2 is therefore complete. □

To show the required result of Eq 6.5 , we following the same ideas as those used in Eq 6.4.

Proof. of Lemma 4.3

we have $\bar{\mu}_\theta(x)$ between $\widehat{\mu}_\theta(x)$ and $\mu_\theta(x)$, then it suffices to prove that

$$\widehat{\mu}_\theta(x) \overset{\mathcal{P}}{\longrightarrow} \mu_\theta(x), \ \ \text{as} \quad n \longrightarrow \infty.$$

Due to g^x_θ is a continuous function, we have that: for all $\xi > 0$, there exists $\eta(\xi) > 0$ such that $\forall y \in [\mu_\theta(x) - \xi, \mu_\theta(x) + \xi]$, we get

$$\mid g^x_\theta(\mu_\theta(x)) - g^x_\theta(y) \mid \leq \eta(\xi) \Rightarrow \mid \mu_\theta(x) - y \mid \leq \xi,$$

which implies that

$$P\left(\mid \mu_\theta(x) - y \mid > \xi\right) \leq P\left(\mid g^x_\theta(\mu_\theta(x)) - g^x_\theta(y) \mid > \eta(\xi)\right).$$

By following the same ideas as those used in lemma burba et *al.* [12], we denot:

$$C_{n,\theta}(D_n) = g^{x(2)}_{\theta,N}(\mu_\theta(x)) \text{ and } c = g^{x(2)}_\theta(\mu_\theta(x)).$$

By using again the relationship 6.1, we get

$$\begin{aligned} (6.7)\ \mid C_{n,\theta}(D_n) - c \mid \ &\leq \ \mid C_{n,\theta}(D_n^-) - c \mid \\ &\leq \ \underbrace{\mid C_{n,\theta}(D_n^-) - \mathbb{E}(C_{n,\theta}(D_n^-)) \mid}_{E_1'} + \underbrace{\mid c - \mathbb{E}(C_{n,\theta}(D_n^-) \mid}_{E_2'}. \end{aligned}$$

Then, by applying Lemma 6 and 7 of Laksaci et *al.* [29] and under condition of theorme and assumption (H6), that the term E_1' and E_2' converges in probability to 0.

7. Conclusion and Perspectives

In this research paper, we have investigated the k-NN method in a single index of the nonparametric estimation of the conditional mode function using the local linear method. Under mild regularity conditions, we establish asymptotic normality of the k-NN single index of our estimator by giving an explicit expression of the terms of bias and its the variance.

The advantage of our method is double. On one hand, the functional local linear estimator can improve the estimation accuracy of the conditional mode function by using a high-order kernel. On the other hand, the k-nearest neighbor estimator takes into account the local structure of the data and gives better predictions when the functional data are heterogeneously concentrated.

The current research opens some crucial tracks for the future. In particular, it would be interesting to extend our work to: the local polynomial setting, $\alpha-$mixing or ergodic data, the response and the covariate are both of functional structure, censored, truncated or missing data in the case of censored data.

Acknowledgments
The authors want to thank the referees for their constructive comments.

REFERENCES

1. A. Ait-Saïdi, F. Ferraty, R. Kassa, and P. Vieu: *Cross-validated estimations in the single-functional index model.* Statistics . **42**, (2008), 475-494.
2. I. Almanjahie, Z. Kaid, A. Laksaci, and M. Rachdi: *Predicting temperature curve based on fast kNN local linear estimation of the conditional distribution function.* PeerJ . **9**,(2021), pp. e11719 .
3. I. Almanjahie, W. Alahmari, A. Laksaci, and M. Rachdi: *Computational aspects of the k NN local linear smoothing for some conditional models in high dimensional statistics.* Communications In Statistics-Simulation And Computation. (2021), pp. 1-32.
4. I. Almanjahie, K. A . Assiri, A. Laksaci, Z. Chikr Elmezouar *The k nearest neighbors smoothing of the relative-error regression with functional regressor,* Commun. Stat. Theory Methods, **51**. (2022), 4196–4209. https://doi.org/10.1080/03610926.2020.1811870.
5. İ. Almanjahie , W. Mesfer and L. Ali , *The k nearest neighbors local linear estimator of functional conditional density when there are missing data*, Hacettepe Journal of Mathematics and Statistics, **51** , (2022) no. 3, pp. 914-931.
6. F. Alshahrani, W. Bouabsa, I. Almanjahie and M. Kadi, *kNN local linear estimation of the conditional density and mode for functional spatial high dimensional data.* AIMS Mathematics. **8**, (2023), 15844-15875.
7. S. Attaoui, and N. Ling : *Asymptotic results of a nonparametric conditional cumulative distribution estimator in the single functional index modeling for time series data with applications.* Metrika. **79**, (2016), 485-511.
8. M. Attouch and B. Tawfik: *Asymptotic distribution of robust k-nearest neighbour estimator for functional nonparametric models.* Matematichki Vesnik. **64**, (2012), 275-285.
9. O. Bouanani, A. Laksaci, M. Rachdi and S. Rahmani : *Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics.* Behaviormetrika. **46** ,(2019), pp. 199-233.
10. O. Bouanani, S. Rahmani, A. Laksaci, andM. Rachdi : *Asymptotic normality of conditional mode estimation for functional dependent data.* Indian Journal Of Pure And Applied Mathematics . **51**, (2020), 465-481.

11. D. Bosq *Linear processes in function spaces: theory and applications.* (Springer Science and Business Media, (2000).

12. F. Burba, F. Ferraty and P. Vieu *k-Nearest Neighbour method in functional nonparametric regression.* Journal Of Nonparametric Statistics. **21**, (2009), 453-469.

13. F. Cérou, and A. Guyader: *Nearest neighbor classification in infinite dimension.* ESAIM: PS. **10**, (2006), pp. 340-355.

14. Z. Chikr-Elmezouar, I.M. Almanjahie, A. Laksaci and M. Rachdi: *FDA: strong consistency of the kNN local linear estimation of the functional conditional density and mode.* Journal of Nonparametric Statistics. **31**(1), (2019), 175-195.

15. G. Collomb: *Estimation de la régression par la méthode des k points les plus proches avec noyau: quelques propriétés de convergence ponctuelle.* Nonparametric Asymptotic Statistics (Proc. Conf., Rouen, 1979) (French). **821**, (1980), pp. 159-175.

16. S. Dabo-Niang, Z. Kaid, and A. Laksaci *On spatial conditional mode estimation for a functional regressor.* Statist. Probab. Lett. , **82**,(2012), 1413-1421.

17. J. Demongeot, A. Laksaci , F. Madani, and M. Rachdi: *Functional data: local linear estimation of the conditional density and its application, Statistics.* **47**, (2013), 26-44.

18. J. Dauxois, A. Pousse, and Y. Romain: *Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference.* Journal Of Multivariate Analysis . **12**,(1982), 136-154.

19. J. Fan: *Local Polynomial Modelling and Its Applications.* Pub. Location Boca Raton, 1996.

20. J. Fan, and I. Gijbels *Local polynomial modelling and its applications: monographs on statistics and applied probability* 66. (Routledge,2018).

21. F. Ferraty, A. Peuch, and P. Vieu: *Modèle à indice fonctionnel simple.* Comptes Rendus Mathematique. **336**, (2003), 1025-1028

22. F. Ferraty, A. Laksaci, and F. Vieu: *Functional time series prediction via the conditional mode estimation.* C. R., Math., Acad. Sci. Paris, **340**, (2005), 389-392.

23. F. Ferraty and P. Vieu. *Nonparametric functional data analysis.* Springer, New York, Theory and practice, 2006

24. A. Goia, and P. Vieu, *A partitioned single functional index model.* Computational Statistics. **30**, (2015), 673-692

25. S. Khardani, M. Lemdani, and E. Ould-Saïd: *Uniform rate of strong consistency for a smooth kernel estimator of the conditional mode for censored time series.* J. Statist. Plann. Inference , **141**, (2011), 3426-3436.

26. M. Rachdi, A. Laksaci, J. Demongeot, A. Abdali and F. Madani *Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data.* Comput. Stat. Data Anal. **73**,(2014), pp. 53-68.

27. E. Ould-Saïd and Tatachak, A. A nonparametric conditional mode estimate under RLT model and strong mixing condition. *Int. J. Stat. Econ.*, **6**, 76-92. (2011)

28. A. Laksaci, F. Madani and M. Rachdi: *Kernel conditional density estimation when the regressor is valued in a semi-metric space*. Communcations in Statistics-Theory and Methods. **42**(19),(2013), 3544-3570.

29. A. Laksaci, M. Rachdi and S. Rahmani: *Spatial modelization: local linear estimation of the conditional distribution for functional data*. Spatial Statistics. **6** ,(2013), pp. 1-23.

30. N. Ling, and Q. Xu: *Asymptotic normality of conditional density estimation in the single index model for functional time series data*. Statistics and Probability Letters . **82**, (2012), 2235-2243.

31. M. Mohammedi, S. Bouzebda, A. Laksaci and O. Bouanani *Asymptotic normality of the k-NN single index regression estimator for functional weak dependence data*, Communications in Statistics - Theory and Methods,(2022), DOI: 10.1080/03610926.2022.2150823

32. S. Muller, and J. Dippon *k-NN kernel estimate for nonparametric functional regression in time series analysis.* Fachbereich Mathematik, Fakultat Mathematik Und Physik (Pfaffenwaldring 57) . **14**,(2011), pp. 2011.

33. J. Ramsay, and B. Silverman . *Functional Data Analysis Springer-Verlag.* New York . (1997)

34. D. Ruppert, and M. Wand *Multivariate locally weighted least squares regression.* The Annals Of Statistics. (1994), pp. 1346-1370

35. P. Sang, and J. Cao: *Functional single-index quantile regression models.* Statistics And Computing, (2020), pp. 1-11