KYBERNETIKA — VOLUME 57 (2021), NUMBER 5, PAGES 819–839

LOCAL LINEAR ESTIMATION OF CONDITIONAL CUMULATIVE DISTRIBUTION FUNCTION IN THE FUNCTIONAL DATA: UNIFORM CONSISTENCY WITH CONVERGENCE RATES

Chaima Hebchi and Abdelhak Chouaf

In this paper, we investigate the problem of the conditional cumulative of a scalar response variable given a random variable taking values in a semi-metric space. The uniform almost complete consistency of this estimate is stated under some conditions. Moreover, as an application, we use the obtained results to derive some asymptotic properties for the local linear estimator of the conditional quantile.

Keywords: functional data, local linear estimator, conditional cumulative, conditional quantile, nonparametric regression, small balls probability

Classification: 62G08, 62G20

1. INTRODUCTION

The conditional estimation is an important field in statistics which dates back to Stone [26] and has been widely studied in the real case. It is useful in all domains of statistics, such as time series, survival analysis and growth charts among others, see Koenker [18, 19] for a review. There exist extensive literature and various nonparametric approaches in conditional estimation for independent samples and dependent non-functional or functional observations. Among the lot of papers dealing with conditional estimation in finite dimension, one can refer for example to key works of Stute [27], Samanta [25], Portnoy [21], Koul and Mukherjee [20], Berlinet et al. [3], Honda [16], Gannoun et al. [15] and Yu et al. [33]. This paper is concerned with conditional distribution estimation when the data are both independent and the covariates are of functional nature. We use local polynomials with kernel weights as estimation method and we state the uniform almost complete convergence of F^x which is the estimator of the conditional distribution function of Y given $X = x$.

Noting that, these questions of the modelization statistic of functional data has known a growing interest among theoretical and applied statisticians (see Bosq [5], Ramsay and Silverman [23] for the parametric model, Ferraty and Vieu, [13], for the nonparametric

DOI: 10.14736/kyb-2021-5-0819

case). In this context, of functional statistics, the estimation of the conditional cumulative distribution function has great importance. It is involved in many applications, such as reliability or survival analysis. Moreover, as the conditional distribution and its derivative (the conditional density function) provide information about the relationship between X and Y, they lead to some prediction methods, such as the conditional mode, the conditional median or the conditional quantiles. The literature on this topic is quite important (see Ferraty et al. [13] for a list of references). In this vast variety of papers, the authors use the Nadaraya–Watson techniques as estimation method which is a particular case of (1) by taking $b = 0$. However, it is well known that a local polynomial smoothing has various advantages over the kernel method, in particular, this method has superior bias properties to the previous one (see Fan and Gijbels [11] for an extensive discussion on the comparison between both methods, in the multivariate case). In the nonfunctional case, the local polynomial fitting has been the subject of considerable study. Key references on this topic are Fan [10], Fan and Gijbels [11], Fan and Yao [12] and the references therein.

However, only a few results are available for local linear modeling in functional statistics. Indeed, the first result in this topic were obtained by Baillo and Grané [2].

They studied the local linear estimator of the regression function when the explanatory variable takes values in a Hilbert space. However, The general case where regressors are not Hilbertian has been considered by Barrientos-Marin et al. [1]. Recall that in the i.i.d setting, Barrientos–Marin et al. [1] introduced the local linear estimator of the regression operator of a scalar response Y on an explanatory functional variable X, this method had several advantages like making the estimator computation easy and fast while keeping good predictive performance. In this pioneering work, the authors obtained the almost complete convergence (with rate) of the proposed estimate. We return to Boj et al. [4] for another alternative version for the functional local linear modeling. More recently, Demongeot et al. [7] consider the local polynomial modeling of the conditional density function when the explanatory variable is functional and the quadratic error of this estimator has been treated by Rachdi et al. [22]. Thereafter, the almost-complete convergence with rates of the local linear estimator of the conditional cumulative distribution is stated by Demongeot et al. [9].

In fact, as the conditional distribution and its derivative (the conditional density function) provide information about the relationship between X and Y, they lead to some prediction methods, such as the conditional mode, the conditional median and the conditional quantiles. Recall that in the i.i.d setting, Barrientos-Marin et al. [1] introduced the local linear estimator of the regression operator of a scalar response Y on an explanatory functional variable X, as this method had several advantages like making the estimator computation easy and fast while keeping good predictive performance Demongeot et al. [8] and Rachdi et al. [22] have used this method to estimate the conditional density. Thereafter, the almost-complete convergence with rates of the local linear estimator of the conditional cumulative distribution is stated by Demongeot et al. [9].

In the iid setting, our work focus on the local linear estimation of the conditional cumulative distribution for functional data. In section 2, we started by clarifying our model and under some assumptions mentioned in section 3 the main asymptotic results are stated in section 4. In section 5, we will exploit these results to the conditional quantile estimation.

2. THE MODEL

In the study of the conditional distribution for functional data, we observe n pairs (X_i, Y_i) for $i = 1, \dots, n$ identically distributed as (X, Y), the latter is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d. We suppose for $x \in \mathcal{F}$ that there exists a regular version of conditional probability of Y given $X = x$, which is absolutely continous with respect the Lebesgue measure on $\mathbb{R}$. this work aims to state the uniform almost complete convergence of $\hat{F}^x$ which is the estimator of the conditional distribution function of Y given $X = x$ and this estimator is based on the local linear method. In order to obtain the functional local linear estimator of the conditional distribution function, we consider the function $F^x(\cdot)$ as a regression model with the response variable $H(h_H^{-1}(\cdot - Y_i))$ (see Fan and Gijbels [11]) and by the studies of Barrientos-Marin et al. [1], Demongeot et al. [7] and Demongeot et al. [8], the functional local linear estimator is based on the minimization of the following quantity:

$$(\hat{a}, \hat{b}) = \arg \min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^{n} (h_H^{-1} H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x))^2 K(h_K^{-1} \delta(X_i, x)) \tag{1}$$

where $\hat{a}$ is the estimator of the conditional distribution function $\beta(\cdot, \cdot)$ and $\delta(\cdot, \cdot)$ are two functions defined from $\mathcal{F} \times \mathcal{F}$ to $\mathbb{R}$, such that:

$$\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0, \text{and} \quad d(\cdot, \cdot) = |\delta(\cdot, \cdot)|.$$

K and H are Kernels and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is chosen as a sequence of positive real numbers and each of them converges to 0 when $n \to \infty$.

Clearly, by using a simple algebra, we can obtain explicitly the following definition of $\hat{F}^x$:

$$\hat{F}^x(y) = \frac{\sum_{i,j=1}^{n} W_{i,j}(x) H(h_H^{-1}(y - Y_i))}{\sum_{i,j=1}^{n} W_{i,j}(x)} \tag{2}$$

where

$$W_{i,j}(x) = \beta(X_i, x)(\beta(X_i, x) - \beta(X_j, x)) K(h_K^{-1} \delta(x, X_i)) K(h_K^{-1} \delta(x, X_j))$$

with the convention $0/0 = 0$.

It's clear that if $b = 0$, we obtain from (1) the Nadaraya–Watson estimator studied.

3. ASSUMPTIONS

The purpose of this paper is to establish the uniform almost complete convergence of $\hat{F}$ on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$, such that:

$$S_{\mathcal{F}} \subset \cup_{k=1}^{d_n} B(x_k, r_n)$$

where $x_k \in \mathcal{F}$ and r_n (resp d_n) is a sequence of positive real numbers. All along the paper, x will denote a fixed point in $\mathcal{F}$ and $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$.

Then, we assume that our nonparametric model satisfies the following conditions:

(H1) There exists a differentiable function $\phi(\cdot)$, such that:
$\forall x \in S_{\mathcal{F}}, 0 < C\phi(h) \leq \phi_x(h) \leq C^{'}\phi(h) < \infty$ and $\exists \eta_0 > 0, \forall \eta < \eta_0, \phi^{'}(\eta) < C$, where C and, $C^{'}$ are strictly positive constants and where $\phi^{'}$ denotes the first derivative of ϕ.

(H2) The conditional distribution function F^x is such that: there exist $b_1 > 0, b_2 > 0$, $\forall (y_1, y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}, \forall (x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}}$

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C(d^{b_1}(x_1, x_2) + |y_1 - y_2|^{b_2})$$

where C is a positive constant.

(H3) the function $\beta(\cdot,\cdot)$ is such that:

$$\forall x^{'} \in \mathcal{F}, C_1 d(x, x^{'}) \leq\mid \beta(x, x^{'}) \mid\leq C_2 d(x, x^{'}) \text{ where } C_1 > 0 \text{ and } C_2 > 0$$

and, for some strictly positive constant $C^{'}$, the following Lipschitz's condition:

$$\forall (x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}} : |\beta(x_1, x^{'}) - \beta(x_2, x^{'})| \leq C^{'} d(x_1, x_2).$$

(H4) K is a positive, differentiable function with support $[-1, 1]$ and, for some strictly positive constant C, the following Lipschitz's condition:

$$\mid K(x) - K(y) \mid\leq C \mid \mid x \mid - \mid y \mid \mid .$$

(H5) The kernel H is a differentiable, positive, bounded and Lipschitzian function, such that: H is of class C^2, of compact support and satisfies:

$$\int |t|^{b_2} H^{(1)}(t)\, \mathrm{d}t < \infty.$$

(H6) the bandwidth h_k satisfies: there exists an integer n_0, such that:
$\forall n > n_0, -\dfrac{1}{\Phi_x(h_K)} \displaystyle\int_{-1}^{1} \phi_x(zh_K, h_K) \frac{\mathrm{d}}{\mathrm{d}z}(z^2 K(z))\, \mathrm{d}z > C_3 > 0$
and
$h_K \displaystyle\int_{B(x,h_K)} \beta(u, x)\, \mathrm{d}P(u) = o\Big(\int_{B(x,h_K)} \beta^2(u, x)\, \mathrm{d}P(u)\Big)$
where $B(x, r)$ denotes the closed-ball and $dP(x)$ is the cumulative distribution of X.

(H7) for $r_n = O\Big(\frac{\ln n}{n}\Big)$, and for some $\gamma \in (0, 1)$, for n large enough the sequence d_n satisfies:

$$\frac{(\ln n)^2}{n^{1-\gamma} h_H^2 \phi(h_k)} < \ln d_n < \frac{n^{1-\gamma} h_H^2 \phi(h_k)}{\ln n}$$

and:

$$\sum_{n=1}^{\infty} n^{\dfrac{(3\gamma+1)}{2}} d_n^{1-\beta} < \infty, \text{ for some } \beta > 1.$$

We notice that conditions (H1) and (H7) are linked with the topological structure of the functional variable, hypothesis (H2) is a Lipschitz condition which characterizes the functional space of our model and it's allowed to state the rate of convergence. The assumptions (H4), (H6) are the same as those mentioned in Barrientos-Marin et al. [1] and Demongeot et al. [8]. (H5) is technical condition.

4. ASYMPTOTIC RESULTS

Before giving the asymptotic result,we introduce the following notations

$$\hat{F}_N^x(y) = \frac{1}{n(n-1)EW_{12}} \sum_{i,j=1}^{n} W_{i,j} H_i(h_H^{-1}(y - Y_i)$$

and

$$\hat{f}_D^x = \frac{1}{n(n-1)EW_{12}} \sum_{i,j=1}^{n} W_{i,j}.$$

Theorem 4.1. Under (H1) – (H7), we obtain that:

$$\sup_{x \in S_{\mathcal{F}}} \sup_{x \in S_{\mathbb{R}}} |\hat{F}^x(y) - F^x(y)| = O(h_k^{b_1} + h_H^{b_2}) + O_{a.co}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right).$$

Remark that, the theorem's proof can be deduced directly from the following decomposition

$$\begin{aligned} \hat{F}^x(y) - F^x(y) &= \frac{1}{\hat{f}_D^x}\left\{\left(\hat{F}_N^x(y) - \mathbb{E}[\hat{F}_N^x(y)]\right) - \left(F^x(y) - \mathbb{E}[\hat{F}_N^x(y)]\right)\right\} \\ &+ \frac{\hat{F}^x(y)}{\hat{f}_D^x}\left(1 - \hat{f}_D^x\right) \end{aligned} \tag{3}$$

in addition Lemmas 4.2, 4.4 and 4.5 below (for which the proofs are given in the Appendix) lead us to get Theorem 4.1.

Lemma 4.2. under assumptions (H1), (H3), (H4), (H5) and (H6), we obtain that:

$$\sup_{x \in S_F} |\hat{f}_D^x - 1| = O_{a.co}\left(\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right).$$

Corollary 4.3. Under the assumptions of Lemma 4.2, we have that:

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x \in S_{\mathcal{F}}} \hat{f}_D^x < \frac{1}{2}\right) < \infty.$$

Lemma 4.4. Under assumptions (H1), (H2), (H4) and (H5) we obtain that:

$$\sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathbb{R}}} |F^x(y) - \mathbb{E}[\hat{F}_N^x(y)]| = O(h_K^{b_1}) + O(h_H^{b_2}).$$

Lemma 4.5. under the hypotheses (H1) – (H7) we obtain that:

$$\sup_{x \in S_F} \sup_{y \in S_{\mathbb{R}}} |\hat{F}_N^x(y) - \mathbb{E}\hat{F}_N^x(y)| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right).$$

5. APPLICATION: CONDITIONAL QUANTILE ESTIMATION

The purpose of this section is to state the asymptotic convergence of the conditional functional quantile, denoted by $t_\theta(x)$ for $\theta \in [0,1]$. For this aim, we will need the following assumptions:

(H8) $\forall \epsilon_0 > 0, \exists \eta > 0, \forall r : S_{\mathcal{F}} \to S_{\mathbb{R}}$, we have that:

$$\sup_{x\in S_{\mathcal{F}}} |t_\theta(x) - r(x)| \geq \epsilon_0 \Rightarrow \sup_{x\in S_{\mathcal{F}}} |F^x(r(x)) - F^x(t_\theta(x))| \geq \eta.$$

(H9) there exists some integer $j > 1$ such that $\forall x \in S_{\mathcal{F}}$, the function F^x is j-times continuously differentiable on interior $(S_{\mathbb{R}})$ with respect to y, and:

$$\begin{cases} F^{x(l)}(t_\theta(x)) = 0, \text{ if } 1 \leq l < j \\ \text{and} F^{x(j)}(\cdot) \text{is uniformly continuous on } S_{\mathbb{R}} \\ \text{such that} |F^{x(j)}(t_\theta(x))| > C > 0 \end{cases}$$

where $F^{x(j)}$ denotes the jth order derivative of the conditional distribution F^x.

The estimator of $t_\theta(x)$ is the random variable $\hat{t}_\theta(x)$ which defined by

$$\hat{t}_\theta(x) = \inf\{y \in \mathbb{R},\ \hat{F}^x(y) \geq \theta\}$$

and we can get the following corollary

Corollary 5.1. Under the hypotheses of Theorem 4.1 and if conditions (H8) – (H9) are satisfied, we get

$$\sup_{x\in S_{\mathcal{F}}} |\hat{t}_\theta(x) - t_\theta(x)| = O(h_K^{b_1} + h_K^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma\phi(h_K)}}}\right).$$

6. SIMULATION STUDY

We first construct the simulation of the explanatory functional variables. In the second part, we focus on the ability of the nonparametric functional regression to predict responses variable from functional predictors. finally we illustrated the MONTE-CARLO methodology and we will appropriate to test the efficiency of the asymptotic normality results parallel the practical experiment.

For this purpose, we consider the following process explanatory functional variables for $n = 350$:

$$X_i(t) = 1 - \sin(2\Omega_i t)\alpha_i + \Omega_i t\,, \quad \forall t \in [0,3] \tag{4}$$

where α_i and Ω_i are n independent real random variables (r.r.v.) uniformly distributed over $[0.3; 2]$ (resp. $[1; 3]$),t is assumed that these curves are observed on a discretization grid of 100 points in the interval. These functional variables are represented on the Figure 1.

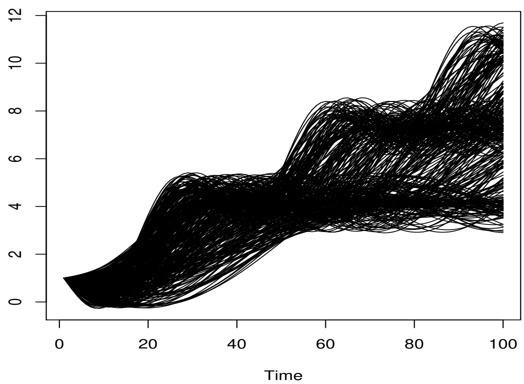

Fig. 1. The curves $X_{i=1,\ldots,200}$.

For response variables Y_i, we consider the following model for all $i = 1, \ldots n$ and $j = 1, \ldots 100$:

$$Y = r(X) + \epsilon$$

where $r(x) = \int_0^{t_j} \frac{1}{1-X_i(v)^2}\, dv$ and ϵ is a centered normal variable and assumed to be independent of $(X_i)_i$. Our goal in this illustration is to show the usefulness of conditional cumulative in a context of forecasting. Thus the use of optimal parameters of the conditional cumulative and without theoretical validity.

Now, we precise the different parameters of our estimators. Indeed, first of all, it is clear that the shape of the curves allows us to use

$$d(x_1, x_2) = \sqrt{\int_0^1 (x_1(t) - x_2(t))^2} \ ; \ \forall x_1, x_2 \in \mathcal{H} \text{ where } \mathcal{H} \text{ is semi-metric.}$$

We choose particularly the quadratic kernels defined by

$$\frac{3}{2}(1 - x^2)\ x \in [0, 1].$$

In this illustration, we select the functional index θ on the set of eigenvectors of the empirical covariance operator.

$$\frac{1}{200} \sum_{i=1}^{200} (X_i - \bar{X})^t ((X_i - \bar{X})).$$

Indeed, we recall that the ideas can be adapted to find a method of practical selection for θ. However, this adaptation in the case of the conditional density require tools and additional preliminary results.

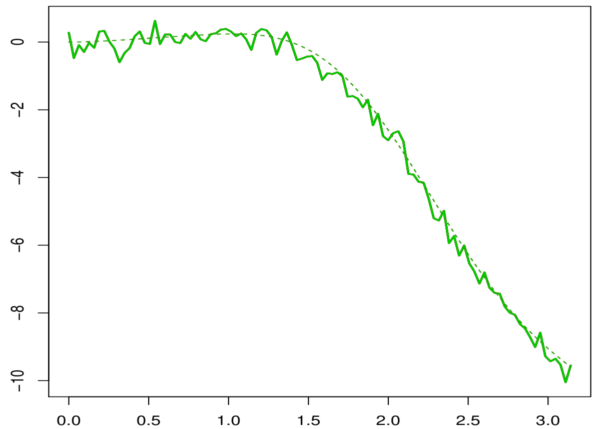

Fig. 2. Predicted functional responses (solid lines); observed functional responses (dashed lines).

For this purpose, we divide our observations on two packets learning sample $(X_i, Y_i)_{i=1,\dots 200}$ and test sample and $(X_i, Y_i)_{i=201,\dots 250}$. For the choice of smoothing parameters h_K and h_H, we will adopt the selection criterion used by Ferraty and Vieu [13] in the case of the kernel method for which h_K and h_H are obtained. by minimizing the next criterion

$$\text{for each} \quad X_i \quad \text{in the sample of the test} \qquad err(h_K, h_H) = |Y_{i^*} - \theta(X_{i^*})| \tag{5}$$

where i^* denotes the index of the nearest curve X_i from all the curves of the learning sample. Now, let us describe the parameter of the estimator:

$$\hat{F}^x(y) = \frac{\sum_{i,j=1}^{n} W_{i,j}(x) H(h_H^{-1}(y - Y_i))}{\sum_{i,j=1}^{n} W_{i,j}(x)}, \forall y \in \mathbb{R}. \tag{6}$$

In this simulation study, we assume the quality of prediction by compared between the predicted functional responses (i. e. $\widehat{F}(x, y)$ for any X in the testing sample) versus the observed functional responses (i. e. $F(x, y)$) as in Figure 2. However, if one wishes to assess the quality of prediction for the whole testing sample, it is much better to see what happens direction by direction.

In other words, displaying the predictions onto the direction $e_{k,n}$ amounts to plotting the 50 points $(< F(x_i, y), e_k >, < \widehat{F}(x_i, y), e_k >)_{i=201,\dots,250}$. Figure 3 proposes a componentwise prediction graph for the two first components $(i.e.k = 1, 2)$. The percentage of variance explained by these 2 components are 99.8% and 0.9% (i. e. $0.998 = (\sum_{k=1}^{2} \widehat{\theta}_k / \sum_{k=1}^{100} \widehat{\theta}_k)$, where $\widehat{\theta}_1 > \widehat{\theta}_2 > \dots$ denotes the eigenvalues of the empirical covariance operator.

The quality of the prediction for the whole sample is also presented in Figure 3. such that the percentage of variance explained by two components are 99.8% and 0.9%. Finally we conclude that the quality of componentwise predictions is quite good for each component.

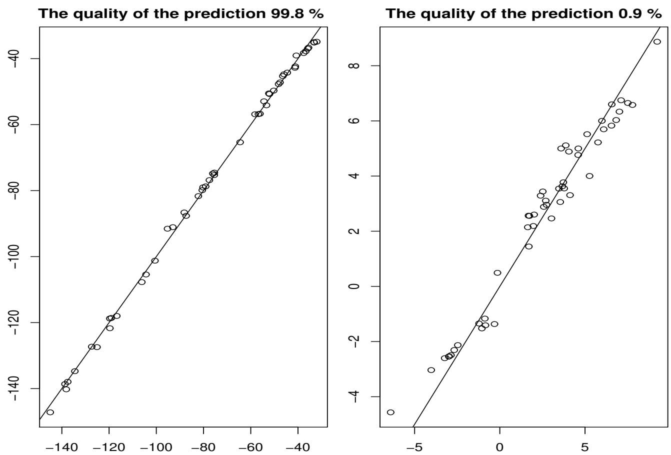

Fig. 3. Representation of the quality of the prediction.

For the next simulation algorithm we used:

- Simulate a sample of size n.
- Calculate the smoothing parameters h_K and h_H that are varied over an interval [0,1] and which minimizes by 5
- We compute the quantities

$$(nh_H\phi_x)^{1/2}(\widehat{F}(x,y) - F(x,y))$$

  where $\widehat{F}(x,y)$ is the functional kernel estimator from the sample $(X_i, Y_i)_{i=1,\dots,200}$, and given in relation (6).
- Compute a standard cumulative estimator by the local linear method.
- Compares the estimated $\widehat{F}(x,y)$ with the corresponding estimated $F(x,y)$).

The obtained results are shown in Figure 4.

It can be seen that, both distributions are very close and have good behaviors with respect to the standard normal distribution.

6.1. Real data application

In this section, we examine the performance of our local linear modeling in functional prediction context. More precisely, we examine its superiority over the classical kernel method in prediction on meteorological data. It should be noted that, the prediction of meteorological data is of interest, in particular, for studying the microclimate conditions in mountainous terrain, resource management or calibration of satellite sensors.

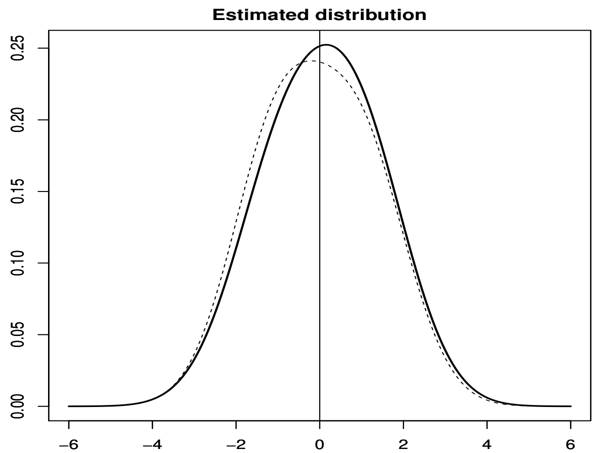

Fig. 4. Representation of the estimated cumulative function with the local linear method.

In this real data example, we are interested in the prediction of the logarithm of the total precipitations given the curve of the monthly maximum temperatures. Specifically, according to the notations of the previous section, the functional predictor X_i is the curve of the monthly maximum temperatures in the ith climatic station (defined by its geographic coordinates) in a period T and Y_i is the logarithm of the total precipitations in the same station and in the same period. The real data considered here are available on the ftp address: `ftp://ftp.ncdc.noaa.gov/pub/data/ushcn/v2/monthly`, and are collected in 100 climatic stations in USA from 2004 to 2010. The functional covariates are given in Figure 5.

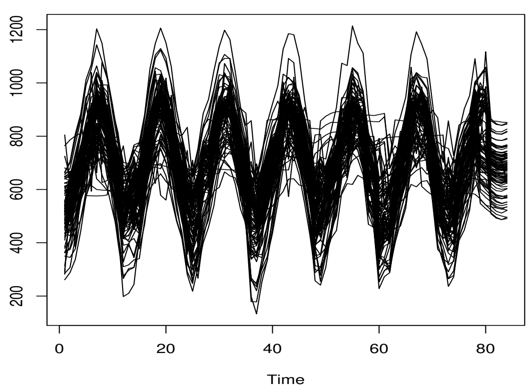

Fig. 5. Monthly maximum temperatures in 100 climatic stations in USA.

For this comparison study, we treat the both estimators in the same conditions and we use are used in the computation of the estimators we proceed by the following estimators: the functional local linear estimator (FLLE) is

$$\text{for each } k, \quad \widehat{F}_1(X_k) = \frac{\sum_{i,j=1}^{n} W_{ij}(X_k) H(h_H^{-1}(y - Y_i))}{\sum_{i,j=1}^{n} W_{ij}(X_k)}$$

and the kernel estimator (KE) is

$$\text{for each } k, \quad \widehat{F}_2(X_k) = \frac{\sum_{i=1}^{n} K(h^{-1} d((X_k), X_i)) H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h^{-1} d((X_k), X_i))}$$

where the smoothing parameters h is locally chosen by cross-validation on the k-nearest neighbors with respect the criterium (5). The kernel K is chosen to be quadratic on $(-1, 1)$. As discussed in the previous section, the shape of the curves has great consideration in the choice of β, δ and d. In this example, we take

$$d(x, x') = \delta(x, x') = \|x - x'\|_{\mathcal{F}} \quad \text{and} \quad \beta(x, x') = < \theta, x - x' >_{\mathcal{F}} .$$

These quantities are computed by using the semi-metric based on the functional principal components analysis (see Ferraty and Vieu [13]) with $q = 4$.

Now, in order to compare the both predictors model, we split our data into two subsets: learning sample (80 stations) and test sample (20 stations) and we compute the mean squared prediction errors (MSE), defined by the following quantities:

$$MSE(FLLE) = \frac{\sum_{i=1}^{20} \left(Y_i - \widehat{F}_1(X_i)) \right)^2}{20} \text{ and } MSE(Kernel) = \frac{\sum_{i=1}^{20} \left(Y_i - \widehat{F}_2(X_i)) \right)^2}{20}.$$

Clearly, the comparison of both scatterplots (see Figure 6) indicates that the local linear estimate method gives better results than those given by the kernel method. Furthermore, our results are also comparable to other tools of forecast such as the Conditional quantile and the median (cf. Ferraty and Vieu [13]).

7. APPENDIX

In what follows, when no confusion is possible, we put for any $x \in \mathcal{F}$, and for all $i = 1, \ldots, n$:

$$K_i(x) = K(h^{-1} \delta(x, X_i)), \beta_i(x) = \beta(X_i, x) \text{ and } H_i(y) = H(h_H^{-1}(y - Y_i)).$$

P r o o f of Lemma 4.2: One starts by using the same decomposition's in Barrientos–

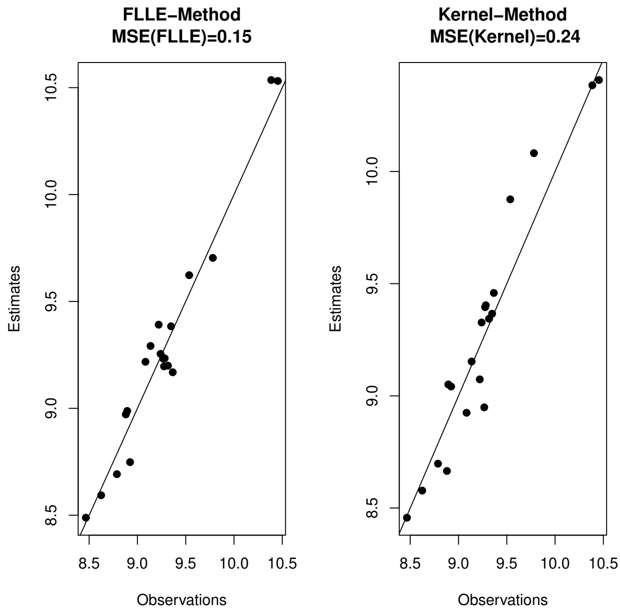

Fig. 6. Comparison of the prediction results between the FLLE-Method and the Kernel-Method.

Marin et al. [1]. Indeed,

$$\hat{f}_D^x = Q_1 \left[\underbrace{\left(\frac{1}{n} \sum_{j=1}^{n} \frac{K_j(x)}{\phi_x(h_K)} \right)}_{P_2(x)} \underbrace{\left(\frac{1}{n} \sum_{i=1}^{n} \frac{K_i(x)\beta_i^2(x)}{h_K^2 \phi_x(h_K)} \right)}_{P_4(x)} - \underbrace{\left(\frac{1}{n} \sum_{j=1}^{n} \frac{K_j(x)\beta_j(x)}{h_K \phi_x(h_K)} \right)}_{P_3(x)} \underbrace{\left(\frac{1}{n} \sum_{i=1}^{n} \frac{K_i(x)\beta_i(x)}{h_K \phi_x(h_K)} \right)}_{P_3(x)} \right].$$

Thus, all it remains to show are the following uniform convergences:

$$\sup_{x \in S_{\mathcal{F}}} |P_k(x) - \mathbb{E}[P_k(x)]| = O\left(\sqrt{\frac{\ln d_n}{n\phi(h_K)}} \right) \quad a.co. \text{ for } k = 2, 3, 4, \tag{7}$$

and

$$\sup_{x \in S_{\mathcal{F}}} |\mathbb{E}[P_2(x)]\mathbb{E}[P_4(x)] - \mathbb{E}[P_2(x)P_4(x)] - \text{var}[P_3(x)]| = o\left(\sqrt{\frac{\ln d_n}{n\phi(h_K)}} \right) \quad a.co$$

and, also that, uniformily on $x \in S_{\mathcal{F}}$:

$$Q_1 = O(1) \text{ and } |\mathbb{E}[P_k(x)]| = O(1), \text{ for } k = 2, 3, 4.$$

Clearly, it suffices to consider (H1) to get the results of the last tow equations. Furthermore, the proof of (7) follows the same ideas as in Ferraty et al. [14]. Indeed, by noting: $j(x) = \arg\min_{j\in\{1,2,\ldots,d_n\}} |\delta(x, x_k)|$, we consider the following decomposition:

$$\begin{aligned}
&\sup_{x\in S_{\mathcal{F}}} |P_k(x) - \mathbb{E}[P_k(x)]| \\
\leq\ & \underbrace{\sup_{x\in S_{\mathcal{F}}} |P_k(x) - P_k(x_{j(x)})|}_{T_1^k} + \underbrace{\sup_{x\in S_{\mathcal{F}}} |P_k(x_{j(x)}) - \mathbb{E}[P_k(x_{j(x)})]\ |}_{T_2^k} \\
&+ \underbrace{\sup_{x\in S_{\mathcal{F}}} |\mathbb{E}[P_k(x_{j(x)})] - \mathbb{E}[P_k(x)]\ |}_{T_3^k}.
\end{aligned}$$

Firstly, we consider that the terms T_1^k and T_3^k have almost the same treatment. let us analyse the first term T_1^k for $k = 2, 3, 4$. Since K is supported in $[-1, 1]$, we can write for all $k = 2, 3, 4$ that:

$$\begin{aligned}
T_1^k \leq\ & \frac{C(k-2)}{nh_K^{k-2}\phi_x(h_K)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^n K_i(x)\mathbb{1}_{B(x,h_K)}(X_i)|\beta_i^{k-2}(x) - \beta_i^{k-2}(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h_K)}(X_i)| \\
&+ \frac{1}{nh_K^{k-2}\phi_x(h_K)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^n \beta_i^{k-2}(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h_K)}(X_i)|K_i(x)\mathbb{1}_{B(x,h_K)}(X_i) - K_i(x_{j(x)})|.
\end{aligned}$$

The Lipschitz condition on K allows us directly to write

$$\begin{aligned}
\mathbb{1}_{B(x_{j(x)},h_K)}(X_i)|K_i(x)\mathbb{1}_{B(x,h_K)}(X_i) - K_i(x_{j(x)})| \leq\ & C\epsilon\mathbb{1}_{B(x,h_K)\cap B(x_{j(x)},h_K)}(X_i) \\
&+ C\mathbb{1}_{B(x_{j(x)},h_K)\cap\overline{B(x,h_K)}}(X_i).
\end{aligned}$$

Once again, one can apply Lipschitz condition on β to get

$$\begin{aligned}
\mathbb{1}_{B(x,h_K)}(X_i)|\beta_i(x) - \beta_i(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h_K)}(X_i)| \leq\ & \epsilon\mathbb{1}_{B(x,h_K)\cap B(x_{j(x)},h_K)}(X_i) \\
&+ h_K\mathbb{1}_{B(x,h_K)\cap\overline{B(x_{j(x)},h_K)}}(X_i) \\
\mathbb{1}_{B(x,h_K)}(X_i)|\beta_i^2(x) - \beta_i^2(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h_K)}(X_i)| \leq\ & \epsilon h_K\mathbb{1}_{B(x,h_K)\cap B(x_{j(x)},h_K)}(X_i) \\
&+ h_K^2\mathbb{1}_{B(x,h_K)\cap\overline{B(x_{j(x)},h_K)}}(X_i)
\end{aligned}$$

which implies that, for $k = 3, 4$:

$$\begin{aligned}
\mathbb{1}_{B(x,h_K)}(X_i)|\beta_i^{k-2}(x) - \beta_i^{k-2}(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h_K)}(X_i)| \leq\ & \epsilon h_K^{k-3}\mathbb{1}_{B(x,h_K)\cap B(x_{j(x)},h_K)}(X_i) \\
&+ h_K^{k-2}\mathbb{1}_{B(x,h_K)\cap\overline{B(x_{j(x)},h_K)}}(X_i).
\end{aligned}$$

Therefore

$$T_1^k \leq C \sup_{x\in S_{\mathcal{F}}} \left(T_{11}^k + T_{12} + T_{13}^k + T_{14}\right),$$

where

$$T_{11}^k = \frac{C(k-2)}{n\phi_x(h_K)}\sum_{i=1}^{n} \mathbb{1}_{B(x,h_K)\cap\overline{B(x_{j(x)},h_K)}}(X_i), \qquad T_{12} = \frac{C\epsilon}{n\phi_x(h_K)}\sum_{i=1}^{n} \mathbb{1}_{B(x,h_K)\cap B(x_{j(x)},h_K)}(X_i)$$

$$T_{13}^k = \frac{C(k-2)\epsilon}{nh_K\phi_x(h_K)}\sum_{i=1}^{n} \mathbb{1}_{B(x,h_K)\cap B(x_{j(x)},h_K)}(X_i), T_{14} = \frac{C}{n\phi_x(h_K)}\sum_{i=1}^{n} \mathbb{1}_{B(x_{j(x)},h_K)\cap\overline{B(x,h_K)}}(X_i).$$

Now, to evaluate the terms T_{11}^k, T_{12}, T_{13}^k and T_{14}, one can apply a standard inequality for sums of bounded random variables with Z_i is identified such that:

$$Z_i = \begin{cases} \dfrac{1}{\phi_x(h_K)} \sup_{x\in S_{\mathcal{F}}} \left[\mathbb{1}_{B(x,h_K)\cap\overline{B(x_{j(x)},h_K)}}(X_i)\right] & \text{for } T_{11}^k \\ \dfrac{\epsilon}{h_K\ \phi_x(h_K)} \sup_{x\in S_{\mathcal{F}}} \left[\mathbb{1}_{B(x,h_K)\cap B(x_{j(x)},h_K)}(X_i)\right] & \text{for } T_{12} \text{ and } T_{13}^k \\ \dfrac{1}{\phi_x(h_K)} \sup_{x\in S_{\mathcal{F}}} \left[\mathbb{1}_{B(x_{j(x)},h_K)\cap\overline{B(x,h_K)}}(X_i)\right] & \text{for } T_{14}. \end{cases}$$

Clearly, under the second part of (H1), we have for the first and the last case:

$$Z_1 = O\left(\frac{1}{\phi(h_K)}\right), \ \mathbb{E}[Z_1] = O\left(\frac{\epsilon}{\phi(h_K)}\right) \text{ and } \ \text{var}(Z_1) = O\left(\frac{\epsilon}{(\phi(h_K))^2}\right).$$

So that, we get:

$$T_{11}^k = O\left(\frac{\epsilon}{\phi(h_K)}\right) + O_{a.co.}\left(\sqrt{\frac{\epsilon\ \ln\ n}{n\ \phi(h_K)^2}}\right).$$

In the same way, assumption (H7) allows to get, for T_{12} or T_{13}^k case

$$Z_1 = O\left(\frac{\epsilon}{h_K\phi(h_K)}\right), \ \mathbb{E}[Z_1] = O\left(\frac{\epsilon}{h_K}\right) \text{ and } \ \text{var}(Z_1) = O\left(\frac{\epsilon^2}{h_K^2\phi(h_K)}\right)$$

which implies that:

$$T_{12}^k = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\ \phi(h)}}\right).$$

To achieve the study of the term T_1^k, it suffices to put togather all the intermediate results and to use (H7) to obtain:

$$T_1^k = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right). \tag{8}$$

Furthermore, since:

$$T_3^k \leq \mathbb{E}\left[\sup_{x\in S_{\mathcal{F}}} |P_k(x) - P_k(x_{j(x)})|\right]$$

we get:

$$T_3^k = O\left(\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right). \tag{9}$$

Secondly, about the term T_2^k. For all $\eta > 0$, we have that:

$$\mathbb{P}\left(T_2^k > \eta\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right) = \mathbb{P}\left(\max_{j\in\{1,\dots,d_n\}}|P_k(x_{j_{(x)}}) - \mathbb{E}[P_k(x_{j_{(x)}})]| > \eta\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right)$$
$$\leq d_n \max_{j\in\{1,\dots,d_n\}} \mathbb{P}\left(|P_k(x_{j_{(x)}}) - \mathbb{E}[P_k(x_{j_{(x)}})]| > \eta\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right)$$

we use a Bernstein-type inequality on

$$\Delta_{ki} = \frac{1}{nh_K^{k-2}\phi(h_K)}\left(K_i(x_k)\beta_i^{k-2}(x_k) - \mathbb{E}[K_i(x_k)\beta_i^{k-2}(x_k)]\right),\ \text{for } k = 2,3,4,$$

we have for all $j = 1,\dots,d_n$ and $i = 1,\dots,n$, that:

$$\mathbb{E}|\Delta_{ki}|^m = O(\phi(h_K)^{-m+1}),\ \text{for}\ \ k = 2,3,4$$

and we get

$$\mathbb{P}\left(|P_i(x_k) - \mathbb{E}[P_i(x_k)]| > \eta\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right) = \mathbb{P}\left(\frac{1}{n}|\sum_{i=1}^{n}\Delta_{lki}| > \eta\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right)$$
$$\leq 2\exp\{-C\eta^2\ln d_n\}.$$

Thus, by choosing η such that $C\eta^2 = \beta$, we obtain that:

$$d_n \max_{k\in\{1,\dots,d_n\}} \mathbb{P}\left(|P_i(x_k) - \mathbb{E}[P_i(x_k)]| > \eta\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right) \leq C' d_n^{1-\beta}. \tag{10}$$

Since $\sum_{i=1}^{n} d_n^{1-\beta} < \infty$, we get

$$T_2 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\ \phi(h_K)}}\right).$$

□

P r o o f of Corollary 4.3: Clearly, we have that: $\inf_{x\in S_{\mathcal{F}}} \hat{f}_D(x) \leq \frac{1}{2}$ implies that exists $x \in S_{\mathcal{F}}$ such that

$$|1 - \hat{f}_D(x)| \geq \frac{1}{2}$$

which allows to write

$$\sup_{x\in S_{\mathcal{F}}} |1 - \hat{f}_D(x)| \geq \frac{1}{2}.$$

According to the Lemma 4.2, we obtain

$$\sum_n \mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}} \hat{f}_D(x) < \frac{1}{2}\right) \leq \sum_n \mathbb{P}\left(\sup_{x\in S_{\mathcal{F}}} |1 - \hat{f}_D(x)| \geq \frac{1}{2}\right) < \infty.$$

□

P r o o f of Lemma 4.4: Since the pairs (X_i, Y_i) are identically distributed then:

$$\mathbb{E}[\hat{F}_N^x(y)] = \frac{1}{\mathbb{E}[W_{12}(x)]}\mathbb{E}[W_{12}(x)\mathbb{E}[H_2|X_2]].$$

Next, we use an integration by part to show that:

$$\mathbb{E}[H_2|X_2] = h_H^{-1}\int_{\mathbb{R}} H^{(1)}(h_H^{-1}(y-z))F^x(z)\,\mathrm{d}z.$$

In order to evaluate $\mathbb{E}[H_2|X_2]$, we use the usual change of variables $t = \dfrac{y-z}{h_H}$, we get:

$$|\mathbb{E}[H_2|X_2] - F^x(y)| \leq \int_{\mathbb{R}} H^{(1)}(t)|F^X(y-h_Ht) - F^x(y)|\,\mathrm{d}t.$$

Thus, by using (H_2) and (H_4), we obtain:

$$\mathbb{1}_{B(x,h_K)}|\mathbb{E}[\hat{F}_N^x(y)] - F^x(y)| \leq \int_{\mathbb{R}} H^{(1)}(t)(h_K^{b_1} + h_H^{b_2}|t|^{b_2})\,\mathrm{d}t.$$

Since $H^{(1)}$ is a probability density, and by (H5), we get:

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|F^x(y) - \mathbb{E}[\hat{F}_N^x(y)]| \leq c(h_K^{b_1} + h_H^{b_2}).$$

□

P r o o f of Lemma 4.5: The proof of this lemma follows the same steps as the proof of Lemmma 4.2, where $P_2(x)$, $P_3(x)$ and $P_4(x)$ are replaced by:

$$\begin{cases} M_2^x(y) &= \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\frac{K_j(x)H_j(y)}{\phi_x(h_K)} \\ M_3^x(y) &= \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\frac{K_j(x)\beta_j(x)H_j(y)}{h_K\phi_x(h_K)} \\ M_4^x(y) &= \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\frac{K_j(x)\beta_j^2(x)H_j(y)}{h_K^2\phi_x(h_K)} \end{cases}$$

by using the compactness property of $S_{\mathbb{R}}$, we can write that: there exists a sequence of real numbers $(t_k)_{k=1,\dots,s_n}$, such that: $S_{\mathbb{R}} \subset \cup_{k=1}^{s_n}(t_k - l_n, t_k + l_n)$ where: $l_n = n^{-\frac{3}{2}\gamma-\frac{1}{2}}$ and $s_n = O(l_n^{-1})$.
Taking: $t_y = \arg\min_{t\in\{t_1,\dots,t_{s_n}\}}|y-t|$ and by the same previously notation of $j(x)$ we consider the following decomposition:

$$
\begin{aligned}
|M_i^x(y) - \mathbb{E}[M_i^x(y)]| \leq & \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{y\in S_{\mathbb{R}}} |M_i^x(y) - M_i^{x_{j(x)}}(y)|}_{A_1} \\
& + \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{y\in S_{\mathbb{R}}} |M_i^{x_{j(x)}}(y) - M_i^{x_{j(x)}}(t_y)|}_{A_2} \\
& + \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{y\in S_{\mathbb{R}}} |M_i^{x_{j(x)}}(t_y) - \mathbb{E}[M_i^{x_{j(x)}}(t_y)]|}_{A_3} \\
& + \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{y\in S_{\mathbb{R}}} |\mathbb{E}[M_i^{x_{j(x)}}(t_y)] - \mathbb{E}[M_i^{x_{j(x)}}(y)]|}_{A_4} \\
& + \underbrace{\sup_{x\in S_{\mathcal{F}}} \sup_{y\in S_{\mathbb{R}}} |\mathbb{E}[M_i^{x_{j(x)}}(y)] - \mathbb{E}[M_i^x(y)]|.}_{A_5}
\end{aligned}
$$

Similarly to the study of the term T_1 in Lemmma 4.2, we obtain:

$$
A_1 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \text{ and } A_5 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right). \tag{11}
$$

To treat the term A_2, we use the Lipschitz's condition on the kernel H and, we can write: $|M_i^{x_{j(x)}}(y) - M_i^{x_{j(x)}}(t_y)| \leq C\frac{1}{nh_K^l\phi(h_K)}\sum_{i=1}^{n} K_i(x_{j(x)})\beta_i^l(x_{j(x)})|H_i(y) - H_i(t_y)| \leq \frac{l_n}{h_H}P_i(x_{j(x)})$ where $P_i(\cdot)$ for $i = 2, 3, 4$ is treated in proof of Lemma 4.2. Thus by using the facts that: $l_n = n^{-\frac{3}{2}\gamma-\frac{1}{2}}$ and (H7), we obtain:

$$
A_2 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \text{ and } A_4 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right). \tag{12}
$$

Finally, for the term A_3, we have for all $\eta > 0$ that:

$$
\begin{aligned}
&\mathbb{P}\left(A_3 > \eta\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \\
&= \mathbb{P}\left(\max_{j\in\{1,2,\ldots,S_n\}} \max_{k\in\{1,2,\ldots,d_n\}} |M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right) \\
&\leq s_n d_n \max_{j\in\{1,2,\ldots,S_n\}} \max_{k\in\{1,2,\ldots,d_n\}} \mathbb{P}\left(|M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right).
\end{aligned}
$$

To do this last probability, we use the classical Bernstein's inequality such that, we put:

$$
Z_i^l = \frac{1}{h_K^l\phi(h_K)}\left(K_i(x_k)H_i(t_j)\beta_i^l(x_k) - \mathbb{E}[K_i(x_k)H_i(t_j)\beta_i^l(x_k)]\right), \text{ for } l = 0, 1, 2.
$$

By the assumption (H3), we have that $\frac{1}{h_K^l}(K_i\beta_i^l) < C$ and since $H < 1$ then, we can write:

$$|Z_i^l| \leq \frac{C}{\phi_x(h_K)} \quad \text{and} \quad \mathbb{E}|Z_i^l| \leq \frac{C^{'}}{\phi_x(h_K)}.$$

So, the use of the classical Bernstein's inequality allows us to write for all $\eta \in (0, C^{'}/C)$:

$$\mathbb{P}\left(|\sum_{i=1}^{n} Z_i^l| \geq \eta\sqrt{\frac{\ln\ d_n}{n\phi_x(h_K)}}\right) \leq C^{'} d_n^{-C\eta^2}.$$

Therefore, the last inequality allows to get:

$$\forall j \leq S_n, \mathbb{P}\left(|M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right) \leq 2\exp\{-C\eta^2 \ln d_n\}$$

since: $S_n = O(n^{-\frac{3}{2}\gamma - 1/2})$, $\frac{\ln d_n}{n\phi(h_K)} > \frac{1}{n\phi(h_K)}$ and by choosing $C\eta^2 = \beta$, one gets:

$$s_n d_n \max_{j\in\{1,2,\ldots,S_n\}} \max_{k\in\{1,2,\ldots,d_n\}} \mathbb{P}\left(|M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right) \leq C^{''} S_n d_n^{1-\beta}.$$

By using the second part of condition (H7) we obtain

$$E_3 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right). \tag{13}$$

Thus, Lemma 4.5's result can be deduced from (11), (12), (13). □

Proof of Corollary 5.1: With a Taylor development of $F^x(\hat{t}_\theta(x))$ around $t_\theta(x)$, we get:

$$F^x(\hat{t}_\theta(x)) = F^x(t_\theta(x)) + \sum_{i=1}^{j-1}\frac{1}{i!}(\hat{t}_\theta(x) - t_\theta(x))^i F^{x(i+1)}(t_\theta(x)) + \frac{1}{j!}(\hat{t}_\theta(x) - t_\theta(x))^j F^{x(j+1)}(t_\theta^{'}(x))$$

because of (H8), we have:

$$F^x(\hat{t}_\theta(x)) = F^x(t_\theta(x)) + \frac{1}{j!}(\hat{t}_\theta(x) - t_\theta(x))^j F^{x(j)}(t_\theta^{'}(x))$$

where $t_\theta^{'}(x)$ is lying between $t_\theta(x)$ and $\hat{t}_\theta(x)$.

The condition (H9) and from the definition of $t_\theta(x)$ and $\hat{t}_\theta(x)$ allow to write:

$$|F^x(\hat{t}_\theta(x)) - F^x(t_\theta(x))| \leq 2\sup_{y\in S_{\mathbb{R}}} |\hat{F}^x(y) - F^x(y)|. \tag{14}$$

It is clear that, from conditions (H8), (14) and from Theorem 4.1, we obtain that:

$$\sup_{x \in S_{\mathcal{F}}} |\hat{t}_\theta(x) - t_\theta(x)| \to 0, \ a.co.$$

And by using (H9) again, we obtain that:

$$\sup_{x \in S_{\mathcal{F}}} |F^{x(j)}(\theta^{'}(x)) - F^{x(j)}(t_\theta(x))| \to 0, \ a.co.$$

Consequently, we can get $\tau > 0$ such that:

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x \in S_{\mathcal{F}}} F^{x(j)}(\theta^{'}(x)) < \tau \right) < \infty$$

and we have:

$$\sup_{x \in S_{\mathcal{F}}} |\hat{t}_\theta(x) - t_\theta(x)|^j \leq C \sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{R}}} |\hat{F}^x(y) - F^x(y)|, \ a.co.$$

So, this last inequality together with Theorem 4.1 lead us to get the claimed result.

□

ACKNOWLEDGEMENT

We show our gratitude to the referees for their constructive comments, which help greatly to improve the presentation of this paper.

(Received March 31, 2020)

REFERENCES

[1] J. Barrientos-Marin, F. Ferraty, and P. Vieu: Locally modelled regression and functional data. J. Nonparametr. Statist. *22* (2010), 617–632. DOI:10.1080/10485250903089930

[2] A. Baillo and A. Grané: Local linear regression for functional predictor and scalar response. Journal of Multivariate Analysis. *100* (2009), 102–111. DOI:10.1016/j.jmva.2008.03.008

[3] A. Berlinet, A. Gannoun, and E. Matzner-Lober: Asymptotic normality of convergent estimators of the conditional mode. Canadian J. Statistics. *26* (1998), 365–380. DOI:10.2307/3315517

[4] E. Boj P. Delicado J. Fortiana: Distance-based local linear regression for functional predictors. Computational Statistics and Data Analysis *54* (2010), 429–437. DOI:10.1016/j.csda.2009.09.010

[5] D. Bosq: Linear processes in function spaces. Theory and Applications. Lecture Notes in Statistics. *149* (2000), Springer-Verlag.

[6] O. Bouanani, A. Laksaci, R. Rachdi, and S. Rahmani: Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. Behaviormetrika *46* (2019), 199–233. DOI:10.1007/s41237-018-0057-9

[7] J. Demongeot, A. Laksaci, F. Madani, and M. Rachdi: A fast functional locally modeled conditional density and mode for functional time-series. Recent Advances in Functional Data Analysis and Related Topics, Contributions to Statistics, Physica-Verlag/Springer 2011, pp. 85–90.

[8] 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. DOI:10.1080/02331888.2011.568117

[9] J.Demongeot, A. Laksaci, M. Rachdi, and S. Rahmani: On the local modalization of the conditional distribution for functional data. Sankhya A *76* (2014), 2, 328–355. DOI:10.1007/s13171-013-0050-z

[10] J. Fan: Design-adaptive nonparametric regression. J. Amer. Statist. Assoc. *87* (1992), 998–1004. DOI:10.1080/01621459.1992.10476255

[11] J. Fan and I. Gijbels: Local Polynomial Modelling and its Applications. Chapman and Hall, London 1996.

[12] J. Fan and Q. Yao: Nolinear Time Series : Nonparametric and Parametric Methods. (2003) Springer–Verlag, New York.

[13] F. Ferraty and P. Vieu: Nonparametric Functional Data Analysis. Theory and Practice. Springer Series in Statistics. New York 2006.

[14] F. Ferraty, A. Laksaci, A. Tadj, and P. Vieu: Rate of uniform consistency for nonparametric estimates with functional variables. J. Statist. Plann. Inference *140* (2010), 335–352. DOI:10.1016/j.jspi.2009.07.019

[15] A. Gannoun, J. Saracco, and K. Yu: Nonparametric prediction by conditional median and quantiles. J. Statist. Plann. Inference *117* (2003), 207–223. DOI:10.1016/S0378-3758(02)00384-1

[16] T. Honda: Nonparametric estimation of a conditional quantile for α-mixing processes. Annals of the Institute of Statistical Mathematics *52* (2000), 459–470. DOI:10.1023/A:1004113201457

[17] C. Kima, M. Ohb, S. Yangc, and H. Choi: A local linear estimation of conditional hazard function in censored data. J. Korean Statist. Soc. *39* (2010), 347–355. DOI:10.1016/j.jkss.2010.03.002

[18] R. Koenker: Galton, Edgeworth, Frisch, and prospects for quantile regression in econometrics. J. Econometrics *95* (2000), 2, 347–374. DOI:10.1016/S0304-4076(99)00043-3

[19] R. Koenker: Quantile Regression. Cambridge University Press, 2005.

[20] H. L Koul and Mukherjee: Regression quantiles and related processes under long range dependence. Journal of Multivariate Analysis *51* (1994), 318–337. DOI:10.1006/jmva.1994.1065

[21] S. L. Portnoy: Asymptotic behavior of regression quantiles in nonstationary dependent cases. Journal of Multivariate Analysis *38* (1991), 100-113. DOI:10.1016/0047-259X(91)90034-Y

[22] 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.l Statist. and Data Anal. *73* (2014), 53–68. DOI:10.1016/j.csda.2013.11.011

[23] J. O. Ramsay and B. W. Silverman: Functional Data Analysis Second edition. Springer, New York 2005.

[24] G. G. Roussas: On some properties of nonparametric estimates of probability density functions. Bull. Soc. Math. Greece (N.S.) *9* (1968), 29–43. DOI:10.1111/j.1600-0447.1968.tb07928.x

[25] M. Samanta: Non-parametric estimation of conditional quantiles. Statistics and Probability Letters. *vol. 7, issue 5* (1989), 407–412. DOI:10.1016/0167-7152(89)90095-3

[26] C. J. Stone: Consistent Nonparametric Regression. Ann. Statist. *5* (1977), 595–645. DOI:10.1214/aos/1176343886

[27] W. Stute: Conditional Empirical Processes. Ann. Statist. *14(2)* (1986), 638–647. DOI:10.1214/aos/1176349943

[28] M. A. Tanner and W. H. Wong: The estimation of the hazard function from randomly censored data by the kernel method. Ann. Statist. *11* (1983), 989–993. DOI:10.1214/aos/1176346265

[29] J. V. Uspensky: Introduction to Mathematical Probability. McGraw-Hill Book Company 1937.

[30] G. S. Watson and M. R. Leadbetter: Hazard Analysis I. Biometrika *51* (1964), 175–184. DOI:10.1093/biomet/51.1-2.175

[31] X. Xiong, P. Zhou, and Ch. Ailian: Asymptotic normality of the local linear estimation of the conditional density for functional time series data. Comm. Statist. Theory Methods *47* (2017), 3418–3440. DOI:10.1080/03610926.2017.1359292

[32] E. Youndjé, P. Sarda, and P. Vieu: Optimal smooth hazard estimates. Test *5* (1996), 379–394. DOI:10.1007/BF02562624

[33] K. Yu, Z. Lu, and J. Stander: Quantile regression : applications and current research areas. The Statistician *52* (2003), 331–350. DOI:10.1111/1467-9884.00363

Chaima Hebchi, Lab. of Statistics and Process stochastic (LSPS), University Djillali liabès, BP 89, Sidi bel Abbes 22000. Algeria.
e-mail: chaima.hebchi@gmail.com

Abdelhak Chouaf, Lab. of Statistics and Process stochastic (LSPS), University Djillali liabès, BP 89, Sidi bel Abbes 22000. Algeria.
e-mail: abdo_stat@yahoo.fr