symmetry

MDPI

Article

Functional Time Series Analysis Using Single-Index L^1-Modal Regression

Mohammed B. Alamari [1], Fatimah A. Almulhim [2,*], Zoulikha Kaid [1] and Ali Laksaci [1]

[1] Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia; malamari@kku.edu.sa (M.B.A.); zqayd@kku.edu.sa (Z.K.); alikfa@kku.edu.sa (A.L.)
[2] Department of Mathematical Sciences, College of Science, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
* Correspondence: faalmulhim@pnu.edu.sa

Abstract: A new predictor in functional time series (FTS) is considered. It is based on the asymmetric weighting function of quantile regression. More precisely, we assume that FTS is generated from a single-index model that permits the observation of endogenous–exogenous variables by combining the nonparametric model with a linear one. In parallel, the L^1-modal predictor is estimated using the M-estimation of the derivative of the conditional quantile of the generated FTS. In the mathematical part, we prove the complete convergence of the constructed estimator, and we determine its convergence rate. An empirical analysis is performed to prove the applicability of the estimator and to evaluate the impact of different structures involved in the smoothing approach. This analysis is carried out using simulated and real data. Finally, the regressive nature of the constructed predictor allows it to provide a robust instantaneous predictor for environmental data.

Keywords: L^1-modal regression; robust estimation; quantile regression; functional data; kernel smoothing; single index; additive model

MSC: 62G08; 62G10; 62G35; 62G07; 62G32; 62G30; 62H12

check for updates

Academic Editor: Junesang Choi

Received: 17 February 2025
Revised: 12 March 2025
Accepted: 14 March 2025
Published: 19 March 2025

Citation: Alamari, M.B.; Almulhim, F.A.; Kaid, Z.; Laksaci, A. Functional Time Series Analysis Using Single-Index L^1-Modal Regression. *Symmetry* **2025**, *17*, 460. https://doi.org/10.3390/sym17030460

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Forecasting the interaction between dependent functional random variables is an important topic of applied data analysis. This issue has been studied by many authors in the past while considering the conditional mean. Alternatively, in this paper, we use modal regression, which is more informative than the conditional mean. The importance of modal regression is justified by its strong link to the conditional distribution (see [1] for more motivation on the conditional mode prediction). Thus, prediction by combining the conditional mode (CM) with the functional single index algorithm (FSI) is a prominent topic in functional data analysis (FDA).

It is well known that the FSI structure is crucial for increasing the performance of nonparametric prediction by using the projection of the exogenous variable over a given functional. The popularity of the FSI structure comes from the simplicity of its linear part and the flexibility of the nonparametric techniques. The first investigations on this topic were developed by [2,3]. We refer also to [4–6] for a list of references. While the cited references consider the vectorial case, we focus in this paper on the functional case. In fact, single-index regression was considered in functional statistics by [7,8]. They studied the Nadaraya–Watson (NW) estimator of the conditional mean under the FSI setting. We return to [9] for the conditional median using the FSI approach. They treated the estimation

of the slope component using a B-spline approach. Moreover, we found the nonparametric estimation of the unknown link function in [9]. The authors of [10] estimate the compact single functional index with a specific coefficient function. The nonparametric estimation of the conditional distribution under the FSI assumption was studied by [6]. The multi-index case was introduced by [11]. They demonstrated the asymptotic property of the Nadaraya–Watson estimator of the unknown parameters in the functional model. Readers interested in the FSI topic may refer to [12,13] for conditional density estimation when the input variables belong to a Hilbert space. We cite [14–16] for an overview of functional semi-parametric modeling. The second component of this work is the L^1 estimation of modal regression. Furthermore, the functional CM was considered by [17]. Since the publication of the cited monograph, a great number of papers have been developed to explore the relationship between a functional variable and a scalar output one. For example, ref. [18] treats the M-regression, ref. [19] considers linear local regression, and [20] introduces the functional relative error. In this context, nonparametric CM estimation has been studied by many authors. For instance, [21] stated the asymptotic normality of the NW estimator of the CM function. The result was obtained in the case of identically distributed and independent observations. In [22], the authors extended this asymptotic normality to the dependent case. We refer to [23] for the complete consistency of the NW-estimator of the spatio-functional mode. The L_p consistency of this estimator was obtained by [24]. The NW-estimator of the mode function in a functional ergodic time series case was introduced by [25]. They treated the sure consistency when the data exhibited some missingness at random (MAR) in the response variable. Ref. [26] provided the local linear estimator of the CM model. We found in this cited paper the asymptotic distribution of the constructed estimator.

Alternatively, in contrast to the previously cited works, we consider in this paper a new estimation method based on L^1 techniques. The latter defines the modal predictor as the minimizer of the derivative of the quantile function. This approach enhances the robustness properties of modal regression. The asymptotic property of the new predictor is developed under a functional time series drawn by the FSI structure. The obtained results confirm the main advantages of the constructed estimator in both robustness and accuracy. In particular, it accumulates the advantages of different parameters involved in the estimators. Indeed, from the single-index modeling, we improve the accuracy of the prediction, and from the L^1-approach, we enhance its robustness. It should be noted that the novelty of this work is the treatment of the dependent case.

We have modeled the dependency aspect using the strong mixing condition. From a practical point of view, the development of the proposed model over the strong mixing FTS allows us to extend its scope of application. Indeed, it is well known that the strong mixing assumption implies the ϕ-mixing and the ρ-mixing as well as the standard time series models (such as AR, ARIMA and ARCH), which satisfy, under certain additional conditions, the strong mixing assumption (see [27]). Furthermore, the computational ability of the estimator is highlighted using real and simulated data. Specifically, the usefulness of the new predictor has been examined using air quality data. More precisely, we compare it to the FSI-standard CM model. Finally, let us point out that the robustness property is fundamental for functional time series prediction, as it permits controlling the deviations of the complicated structure of the data linked to its functional nature and/or its strong correlation, hence the importance of developing a robust version of the functional regression model in FTS data analysis.

The paper is organized as follows. We introduce the model in the following section. Section 3 is devoted to the establishment of the main result. The applicability of the estimator over artificial data and real data is treated in Sections 4 and 5. In Section 6, we

summarize some concluding remarks. The proof of the technical results is delegated to Section 7.

2. Robust Estimator of the Modal-Regression in FSI Structure

Let (T, S) be a couple of random variables in $\mathcal{E} \times \mathbb{R}$. The functional space $\mathcal{E}$ is Hilbertian and has an inner product $< \cdot, \cdot >$. Let $(T_i, S_i)_{1 \leq i \leq n}$, be n a sample of stationary vectors that have the same distribution as (T, S). In a functional single index structure (FSI), the input–output S and T are related

$$\exists \tau \in \mathcal{E} \quad \mathbb{E}[S|T] = \mathbb{E}[S| < \tau, T >]. \tag{1}$$

Furthermore, the identifiability of the FSI structure has been studied by [8]. It is ensured by assuming the differentiability of the conditional expectation and the functional index τ such that $< \tau, v_1 > \ = \ 1$, where v_1 is the first vector of an orthonormal basis of $\mathcal{E}$. In fact, the FSI approach is very efficient in reducing the effect of the infinite dimensionality of the estimator. Thus, we aim in this paper to improve the convergence of the standard functional modal regression by considering the FSI structure. For this purpose, we fix a location point v in $\mathcal{E}$, and we set N_v as a neighborhood of v. In addition, the conditional distribution function of S given $< \tau, T > \ = \ < \tau, v >$, denoted by $CD_\tau(\cdot|v)$, is strictly increasing and has a continuous density $CF_\tau(t|v)$ with respect to Lebesgue's measure over $\mathbb{R}$. The conditional mode in FSI is the maximizer of the conditional probability density function (CD_τ) over a compact K

$$Md_\tau(v) \ = \ \arg\max_{t \in K} CD_\tau(t| < \tau, v >),$$

Alternatively, the function of Md_τ can be related to the quantile regression QR by using

$$QR'(q|v) = \frac{\partial QR(q|v)}{\partial q} = \frac{1}{CD_\tau(QR(q|v)|v)} \quad q \in CD_\tau^{-1}(K|v) = [a_v, b_v] \subset (0, 1).$$

where $QR(q|v)$ is the conditional quantile of order q given $< \tau, T > \ = \ < \tau, v >$. Therefore, the function $Md_\tau(u)$ satisfies

$$Md_\tau(v) = QR(q_\tau|v) \tag{2}$$

with

$$q_\tau = \arg \min_{q \in [a_v, b_v]} QR'(q|v).$$

Thus, the robust estimator of the FSI mode is

$$\widehat{Md_\tau}(v) = \widehat{QR}(\widehat{q_\tau}|v) \tag{3}$$

where

$$\widehat{q_\tau} = \arg \min_{q \in [a_v, b_v]} \widehat{QR'}(q|v).$$

with $\widehat{QR}$ and $\widehat{QR'}$ are, respectively, estimators of the conditional quantile and its derivative. Firstly, the derivative $QR'(q|v)$, $q \in (0, 1)$ is estimated by

$$\widehat{QR'}(q|v) = \frac{\widehat{QR}(q + \ell_n|v) - \widehat{QR}(q - \ell_n|v)}{2\ell_n}.$$

where ℓ_n is a sequence of positive numbers in $\mathbb{R}$. Secondly, the q^{th} conditional quantile $QR(q|v)$ is estimated by the solution with respect to t of the optimization problem

$$\min_{t\in\mathbb{R}} \mathbb{E}\big[L_q(S,t) \mid <\tau, T>=<\tau, v>\big] \tag{4}$$

where $L_q(u) = (2q-1)(u) + |u|$. Finally, the robust estimator of the FSI mode is obtained by estimating

$$\widehat{QR}(q|v) = \arg\min_{t\in\mathbb{R}} \widehat{Q}(t,q|v). \tag{5}$$

where

$$\widehat{Q}(t,q|v) = \frac{\sum_{i=1}^{n} \mathbf{H}(\ell_n^{-1} d_\tau(u,T_i)) L_q(S_i - t)}{\sum_{i=1}^{n} \mathbf{H}(\ell_n^{-1} d_\tau(u,T_i))},\ \forall t \in \mathbb{R}$$

with $d_\tau(x,y) = |<\tau, x-y>|$. $\mathbf{H}$ is a kernel map and ℓ_n is a sequence of positive in $\mathbb{R}$ that tends to zero as n tends to infinity. As mentioned previously, the main purpose of this work is to establish the asymptotic property $\widehat{Md_\tau}(u)$ under the strong mixing assumption. Recall that a strictly stationary sequence of random variables $Z = (\{Z_i = (T_i, S_i), i = 1, 2, \ldots\}$ is strongly mixing if

$$\alpha(n) = \sup\Big\{|\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2)| : A_1 \in \mathbf{Z}_1^k(Z) \text{ and } A_2 \in \mathbf{F}_{k+n}^{\infty}(Z),\ k \in \mathbb{N}^*\Big\}. \to 0 \quad \text{as} \quad n \to \infty$$

where $\mathbf{F}_i^k(Z)$ is the $\sigma-$ algebra generated by $\big\{Z_j : i \le j \le k\big\}$.

The importance of the mixing assumption comes from the fact that it permits many practical applications, especially in the fields of economics and finance. For example, there exist several time series models satisfying the α-mixing condition, such as ARCH, ARMA, threshold models, EXPAR models, ARCH models, GARCH models, and bilinear Markovian models that are geometrically strongly mixing (see [28,29] for more examples). In this paper, we focus on the almost complete convergence, which is defined below.

Definition 1. *Let $(z_n)_{n\in\mathbb{N}^*}$ be a sequence of real variables; we say that z_n converges almost completely (a.co.) to zero if and only if $\forall \epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon) < \infty$. Moreover, let $(u_n)_{n\in\mathbb{N}^*}$ be a sequence of positive real numbers; we say that $z_n = O(u_n)$ a.co. if and only if $\exists \epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon u_n) < \infty$. This kind of convergence implies both almost certain convergence and convergence in probability.*

3. Main Results

Firstly, we start by denoting C or C' as some positive constants, and we set

$$B(v,r) = \big\{u' \in \mathcal{F} : d_\tau(u', v) < r\big\}.$$

Now, we list some required conditions that are necessary in deriving the almost complete convergence of $\widehat{Md_\tau}(v)$ of $Md_\tau(v)$.

(AS1) $\mathbb{P}(T \in B(v,r)) = \zeta_v(r) > 0$. Furthermore, $\zeta_v(r) \to 0$ as $r \to 0$.

(AS2) The functions $QR(\cdot|v)$ is of class $C^3([a_v, b_v])$ and $CD_\tau(\cdot|v)$ such that the following Lipschitz's condition is satisfied:

$$\text{for all } (v_1, v_2) \in N_v,\ |CD_\tau(t|v_1) - CD_\tau(t|v_2)| \le C d_\tau^b(v_1, v_2) \text{ for some } b > 0.$$

(AS3) The sequence $(Z_i = (T_i, S_i))_{i\in\mathbb{N}}$ satisfies $\exists a > 5, \exists c > 0 :\ \forall n \in \mathbb{N},\ \alpha(n) \le c n^{-a}$ and

$$\mathbb{P}\big((T_i, S_j) \in B(v,r) \times B(v,r)\big) = \psi_v(r) > 0.$$

(**AS4**) **H** is a function with support $(0,1)$ such that $0 < C' < \mathbf{H}(t) < C < \infty$.
(**AS5**) There exists $0 < \eta < 2\frac{a-5}{a+1}$ such that

$$Cn^{\frac{2(5-a)}{(a+1)}+\eta} \leq \varsigma_v(\ell_n),$$

where

$$\xi_v(r) = \max(\zeta_v^2(r), \psi_v(r)).$$

The assumed conditions (**AS1**)–(**AS4**) are classic in nonparametric functional statistics. Assumption (**AS1**) is primordial in this kind of data analysis. The function $\zeta_v(.)$ can be explicitly defined for several continuous processes [17]). Furthermore, the regularity postulate (**AS2**) is assumed to identify the functional space of the model. Such an assumption has a great impact on the bias component in the convergence rate of $\widehat{Md_\tau}(u)$. More precisely, we point out that the conditions (**AS4**)–(**AS5**) are linked to the kernel **H** and the bandwidth ℓ_n. Such technical assumptions are used to simplify the demonstration of the consistency of the estimator $\widehat{Md_\tau}(v)$. In fact, the assumptions are less restrictive than those used in the previous studies of modal prediction in functional data analysis. More precisely, the main advantage of the present contribution is the use of one kernel instead of the standard case approach, which requires two kernels.

Theorem 1. *Under assumptions (***AS1***)–(***AS4***), and if* $\inf_{q\in(0,1)} \frac{\partial^3 QR(q|v)}{\partial q} > 0$ *we have:*

$$\widehat{Md_\tau}(v) - Md_\tau(v) = O(\ell_n^{b'/2}) + O\left(\left(\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\ell_n\zeta_v^2(\ell_n)}\right)^{\frac{1}{4}}\right), \qquad a.co.$$

where $b' = max(b, 1)$.

The obtained convergence rate explores the impact of the strong correlation of the data on the estimation quality. In particular, this aspect is controlled through the function ξ_v. Not surprisingly, the level of dependency has a negative effect on estimation quality. In this sense, the convergence rate decreases when the level of dependency increases. Additionally, the obtained convergence rate highlights the importance of the choice of the smoothing parameter ℓ_n and the functional index τ. The expression of the convergence rate provides a preliminary idea of this direction. Indeed, it suffices to select ℓ_n and τ by minimizing the asymptotic quantity $(\ell_n^{b'/2} + (\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\ell_n\zeta_v^2(\ell_n)})^{\frac{1}{4}})$. Of course, the practical use of this rule requires the plug-in estimation of the unknown functions in this expression, which is not trivial. For this reason, the empirical rules based on the cross-validation criterion are more adequate for practical purposes.

4. Computational Study

This section is devoted to analyzing the computational ability of the constructed estimator. First, we aim to check its easy implantation, and the second aim is to evaluate its behavior in terms of accuracy and robustness. Of course, this analysis is performed through a comparison study between the new estimator and the standard one. More precisely, we will control the effect of dependency, as well as the outlier resistance of both estimators. For this purpose, we generate dependent functional regressors as functional ARCH of order p, whose conditional volatility depends on the following kernel:

$$\text{Default kernel} \quad \psi(t,s) = 12\,t(1-t)\,s(1-s), \quad t,s \in [0,1]. \tag{6}$$

Recall that the functional ARCH of order p is defined by

$$T_i(t) = \sigma_i(t)\epsilon_i(t), \quad \sigma_i^2(t) = w(t) + \sum_{k=1}^{p} \int_0^1 \psi(t,s) T_{i-k}^2(s)\, ds, \quad t \in [0,1]$$

where the innovation $\epsilon_i(t)$ is i.i.d. error generated from an Ornstein–Uhlenbeck process and the constant coefficient $w(t) = 0.1t(1-t)$. The kernel Equation (6) is the default choice in the *FTSgof* R-package. Of course, the correlation level is adjusted via the parameter p. In the cited package, the correlation is quantified through the function *fACF*. The different shapes of the generated functional time series are given in Figures 1–3.

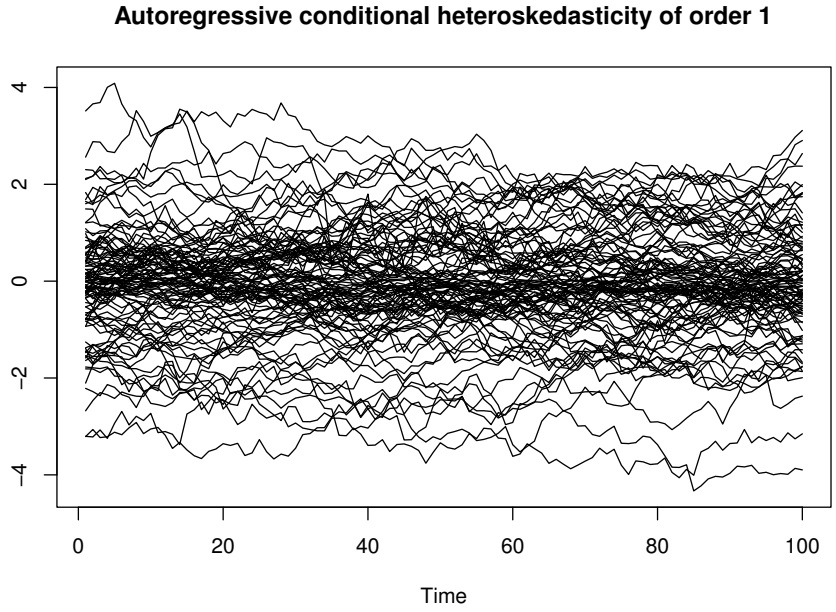


Figure 1. A sample of 100 curves from FARCH(1).

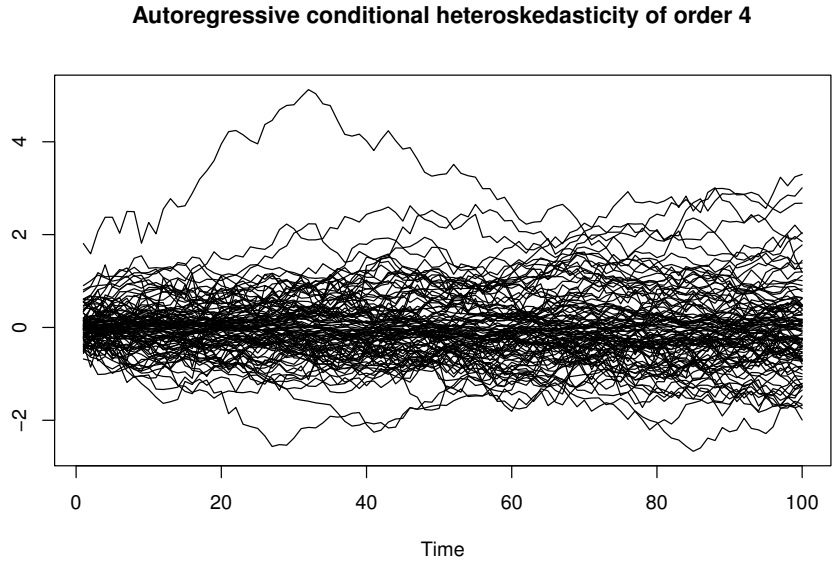


Figure 2. A sample of 100 curves from FARCH(4).

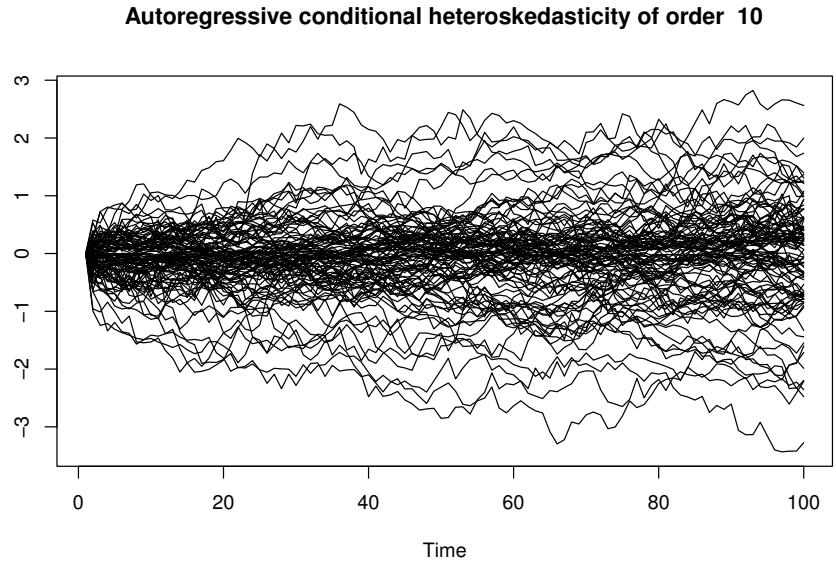


Figure 3. A sample of 100 curves from FARCH(10).

For the response variable S, we use the heteroscedastic single-index model is defined as

$$S_i = 2\cos\left(\pi\left(\int_0^1 e_1(t)T_i\right)^2(t)dt\right) + 0.5\sin\left(\frac{1}{1+\left(\int_0^1 e_1(t)T_i(t)\right)^2}\right)\varepsilon_i$$

where e_1 is the first element of the basis function of Karhunen–Loeve, associated with the sample $(T_i)_i$. The variable ε is the white noise independent of T_i. They are generated using three distinct distributions: Laplace, Weibull and log-normal. This choice is motivated by the heavy-tailed property of the considered distributions. On the other hand, the three distributions are unvaried by translation, which permits identifying the true models. It is obtained by shifting the distribution of ε_i. Furthermore, we use the MAD function to detect the outliers in the response variable. We point out that there exist many other alternative outlier detectors (see [30] for more details). However, the MAD function controls the median absolute deviations, which is very simple to execute. The number of detected outliers using the MAD function is displayed in Figures 4–6.

In these simulation experiments, we compare the estimator $\widehat{Md_\tau}$ to the standard one defined by the conditional density as

$$\widetilde{Md}(v) = \arg\max_s \frac{\sum_{i=1}^n \mathbf{H}(\ell_n^{-1}d_\tau(v,T_i))\mathbf{H}(\ell_n^{-1}(s-S_i))}{\sum_{i=1}^n \ell_n \mathbf{H}(\ell_n^{-1}d_\tau(v,T_i))}.$$

Clearly, the performance of the estimators is related to the determination of the different components of the estimation. In this kind of smoothing approach, the bandwidth parameter and the functional index play a crucial role. In this empirical analysis, the sequence ℓ_n and the functional index τ are selected using the leave-one-out mean square cross-validation rule defined by

$$(\tau_{CV}, \ell_{CV}) = \arg\min_{r_n\in H_n}\min_{\tau\in\Theta_n} \frac{1}{n}\sum_{i=1}^n (S_i - \widehat{M}_n^{-i}(T_i))^2,$$

where $\widehat{M}_n^{-i}(T_i)$ is the leave-one-out estimator $\widetilde{Md}$ or $\widehat{Md}$. The subset H_n is the positive real numbers, r_n, such that the ball centered at v with radius r_n contains exactly $k \in= \{5, 15, 25, \ldots, 55\}$ neighbors of v. We point out that we have used the PCA-metric

to construct the previous balls (see [17] for more details about this metric). In addition, the subset $\Theta_n = \left\{\theta \in \mathcal{F}, \theta = \sum_{i=1}^{k_0} c_i e_i,\ \|\theta\| = 1, \text{ and } \exists j \in 1, \ldots, k_0 \text{ such that } N_\theta(e_j) > 0\right\}$, $(c_i)_i$ being in $\{-1, 0, 1\}$ and e_i is where $(e_i)_{i=1,\ldots,k_0=8}$ is finite basis functions of the Hilbert subspace spanned by the covariates $(T_i)_i$. The latter is determined by the k_0 first eigenvectors associated with the k_0 largest eigenvalues of the variance-covariance matrix of the explanatory variables T_i. Also, we specify that we have used the quadratic kernel. Finally, we examine the efficiency of the estimators $\widetilde{Md}$ or $\widehat{Md}$ using the mean absolute error by

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |S_i - \widehat{M}_n(T_i)|.$$

We compute this error for the three functional time series (for $p = 1, 4, 10$), for three values of $n = 50, 100, 250$ and for two cases. In the first case, we consider the data without modification, while in the second case, we remove the outliers observations detected in Figures 4–6. The obtained results are presented in the following Table 1.

Finally, we see clearly that the MAE error is more stable in the $\widehat{Md}$ estimator. In this sense, the variability of the MAE is more important in the standard case $\widetilde{Md}$. However, the accuracy of the estimation is also affected by the degree of dependency and the non-parametric distribution. There is a significant difference between different functional time series cases, and the MAE-error increases with the level of dependency.

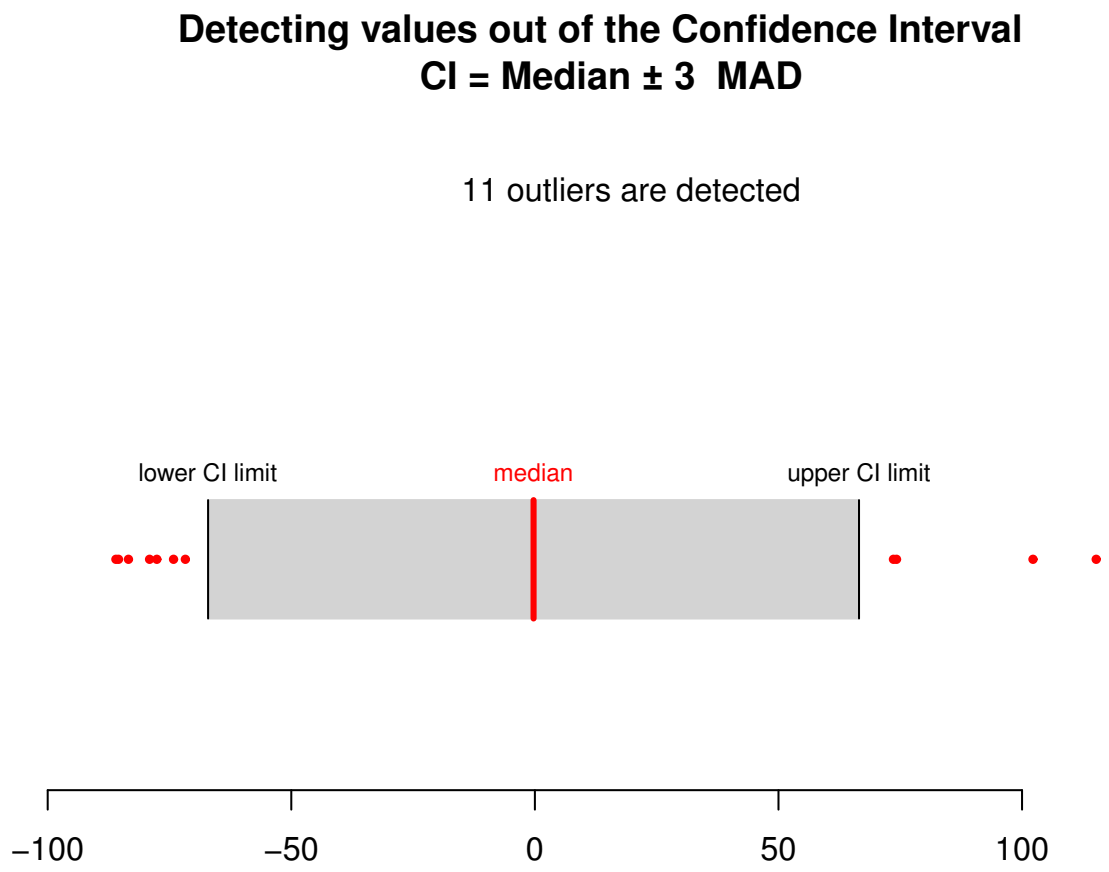


Figure 4. The umber of outliers detected using the MAD-test. The outliers are displayed in red dots.

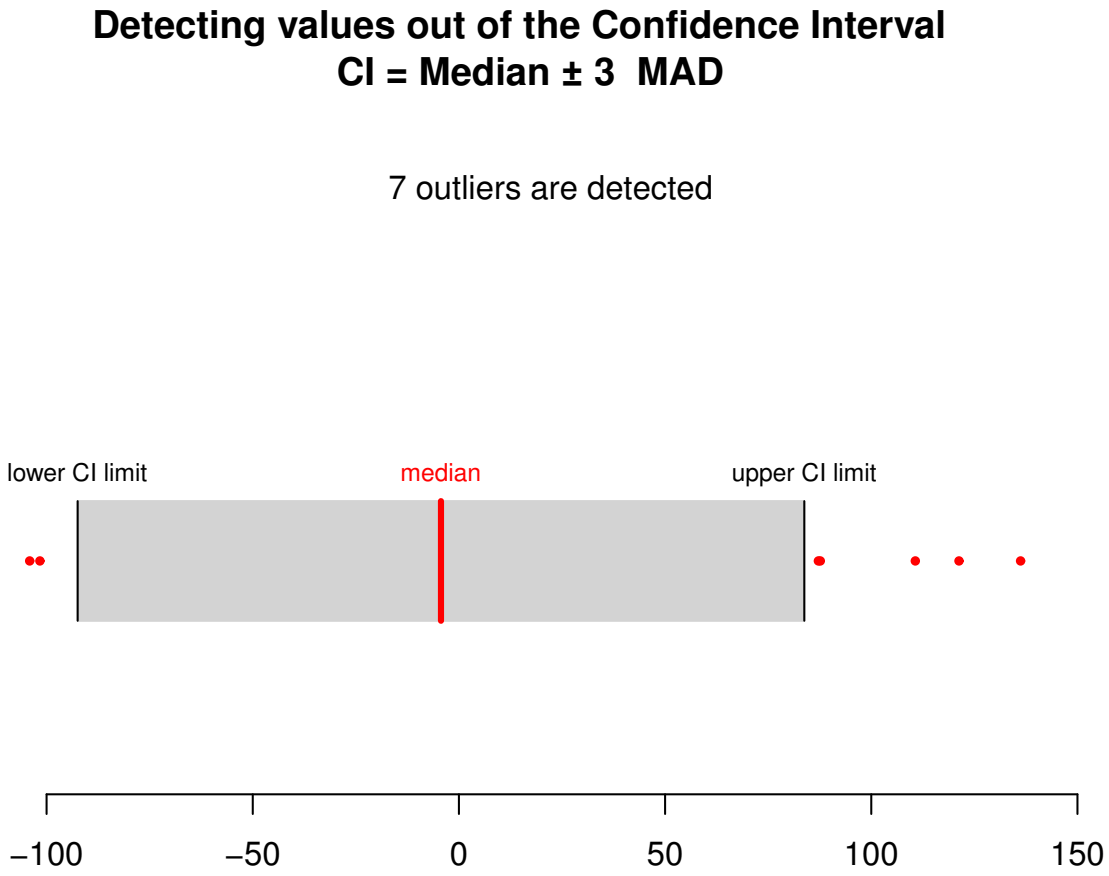


Figure 5. The number of outliers detected using the MAD test. The outliers are displayed in red dots.

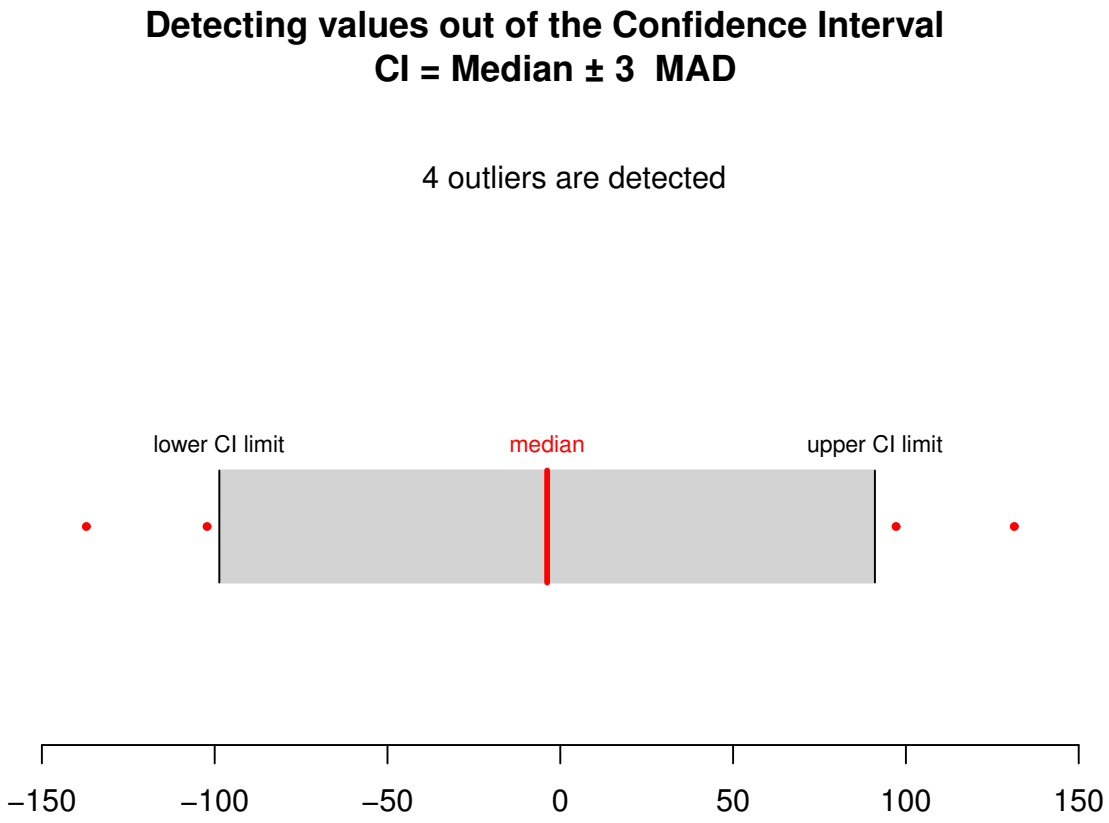


Figure 6. The number of outliers detected using the MAD test. The outliers are displayed in red dots.

Table 1. MAE results.

The White Noise Distribution	n	p	Outliers Case	$MAE\ (\widetilde{Md})$	$MAE\ (\widehat{Md})$
Laplace distribution	50	1	with outliers	2.43	1.37
		1	without outliers	1.06	0.51
		4	with outliers	7.92	2.29
		4	without outliers	1.41	0.39
		10	with outliers	7.82	2.92
		10	without outliers	1.95	0.52
	100	1	with outliers	2.36	1.29
		1	without outliers	0.98	0.43
		4	with outliers	7.61	2.17
		4	without outliers	1.32	0.28
		10	with outliers	7.66	2.84
		10	without outliers	1.89	0.39
	250	1	with outliers	2.17	1.07
		1	without outliers	0.84	0.32
		4	with outliers	7.43	2.06
		4	without outliers	1.11	0.12
		10	with outliers	7.32	2.52
		10	without outliers	1.72	0.25
Weibull distribution	50	1	with outliers	4.54	2.95
		1	without outliers	1.81	0.77
		4	with outliers	4.82	3.23
		4	without outliers	2.43	1.03
		10	with outliers	5.91	3.48
		10	without outliers	2.46	1.05
	100	1	with outliers	4.32	2.88
		1	without outliers	1.63	0.69
		4	with outliers	4.53	3.02
		4	without outliers	2.01	0.85
		10	with outliers	5.78	3.34
		10	without outliers	2.26	0.92
	250	1	with outliers	4.14	2.63
		1	without outliers	1.32	0.43
		4	with outliers	4.25	2.98
		4	without outliers	1.93	0.71
		10	with outliers	5.49	3.11
		10	without outliers	2.01	0.74
Log-normal distribution	50	1	with outliers	6.53	1.32
		1	without outliers	2.50	0.79
		4	with outliers	5.42	2.54
		4	without outliers	2.62	1.12
		10	with outliers	6.61	2.66
		10	without outliers	3.08	1.48
	100	1	with outliers 2	6.36	1.19
		1	without outliers	2.04	0.67
		4	with outliers	5.13	2.33
		4	without outliers	2.39	0.98
		10	with outliers	6.44	2.59
		10	without outliers	2.95	1.23
	250	1	with outliers 2	6.11	1.02
		1	without outliers	1.82	0.51
		4	with outliers	5.02	2.07
		4	without outliers	2.14	0.71
		10	with outliers	6.36	2.28
		10	without outliers	2.77	1.09

5. Real Data Example

In addition to the empirical analysis developed in the last section, we aim in this section to examine the applicability of the proposed model to a real-life example. More precisely, our aim is to show how the robust estimation of the modal regression is useful in the prediction problem. Indeed, we use the predictor $\widehat{Md}$ to forecast the daily temperature in Houston two months ahead using historical data from the past. For this aim, we consider the dataset of this city, which is available on the website "https://kilthub.cmu.edu/articles/dataset/Compiled_daily_temperature_and_precipitation_data_for_the_U_S_cities/7890488" (accessed on 15 February 2025). The station provides daily measurements from 1889 to 2023.

Recall that the main feature of functional data analysis is its ability to model complicated data with high dimensionality. For this reason, it is necessary to examine the reliability of the proposed method over an explanatory variable with a long trajectory. This consideration permits the incorporation of the principle of functional statistics and the evaluation of the role of the functional single-index modeling in this complicated prediction issue. We point out that the choice of an explanatory variable with few historical data is not beneficial for functional data analysis. In this real-data study, we have examined many periods (25 days, 60 days, and 75 days, among others). We have observed that the historical data of 60 days are more informative than the other examined cases. Thus, we will predict the daily temperature of one day given the curves of the last 60 days. For this aim, we construct a functional time series by cutting the whole curve into sub-curves $T_i = (T_i(t))$, $i = 1, \ldots, n+1$, each of which contains the data for 60 days. We compare the robust predictor $\widehat{Md}$ to the standard one $\widetilde{Md}$. In order to highlight the robustness property, we conduct this comparison study in both cases (with outliers and without outliers). Furthermore, the outlier curves are obtained using the routine code *outliers.depth.trim* from the R-package *fda.usc*. We return to [31] for more discussion on the functional outlier detector. The following Figure 7 displays some functional regressors as well as the outlier curves (in red).

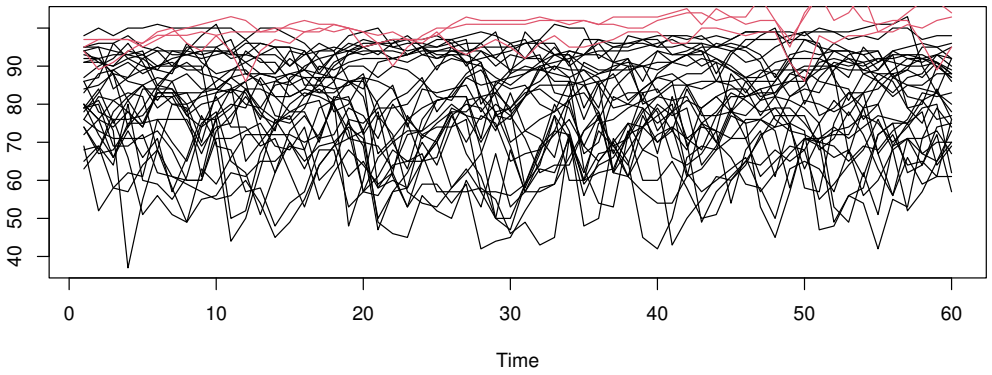


Figure 7. A sample of curves and the number of outlier curves.

In the first case, we execute the function with the initial data without changes, while in the second case, we remove the outlier curves as well as the outliers in the response variable. Finally, we compute the predictors $\widetilde{Md}$ or $\widehat{Md}$ using the same procedure as in Section 4 to choose the bandwidth ℓ_n and the index τ. In addition, we specify that we have used the quadratic kernel function to compute the estimators. The prediction results are plotted in Figures 8–11, where we show the true values of temperature for the last 60 days and the predicted values.

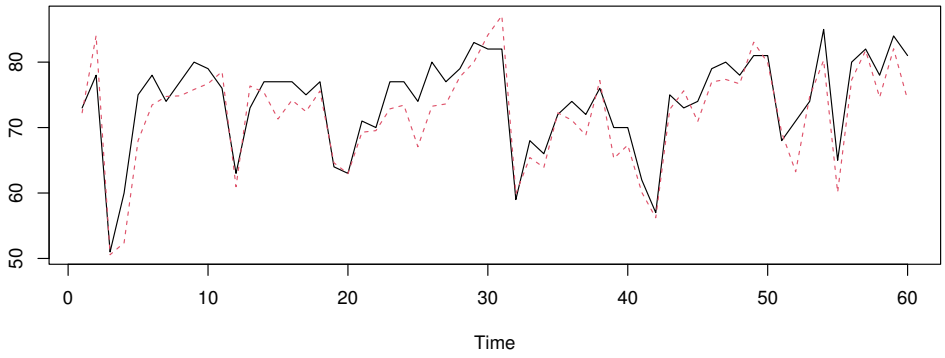


Figure 8. The prediction with the robust mode $\widehat{Md_\tau}$ with outliers. The true values on black line and the predicted on the red line.

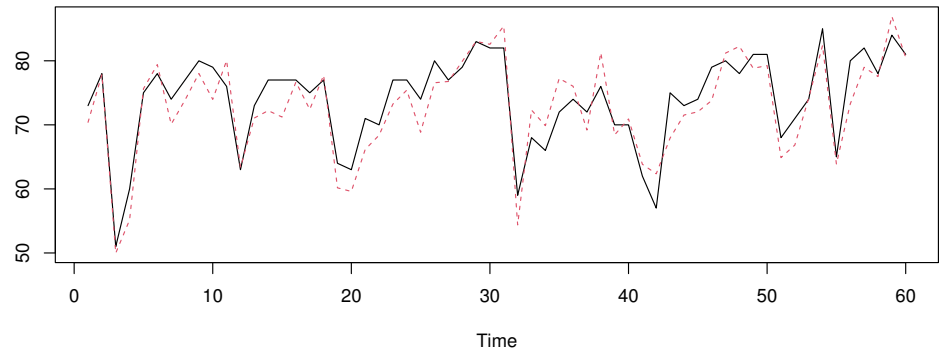


Figure 9. The prediction with the standard mode $\widetilde{Md_\tau}$ with outliers.The true values on black line and the predicted on the red line.

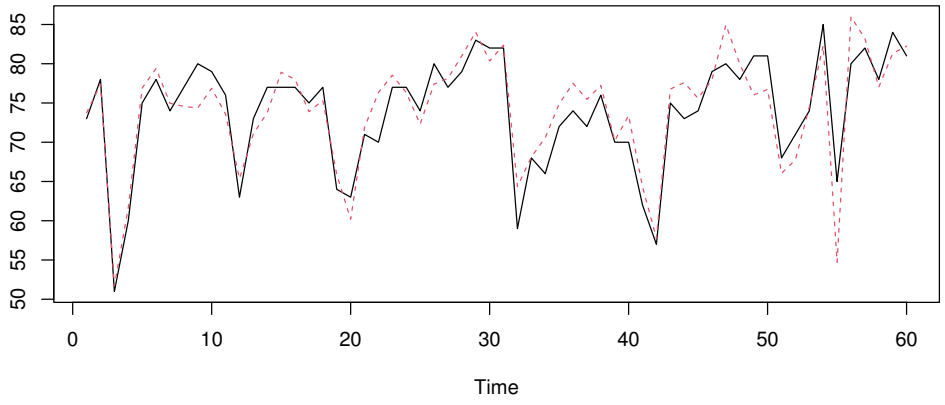


Figure 10. The prediction with the robust mode $\widehat{Md_\tau}$ without outliers. The true values on black line and the predicted on the red line.

We can see that the predictor $\widehat{Md}$ is more appropriate for the first case, when the data contain some outliers. Meanwhile, the estimator $\widetilde{Md}$ performs well when the outliers are removed. This conclusion can be confirmed using the mean absolute error (MAE) used in the last section. We have obtained, for the first situation, an MAE of 3.75 for $\widehat{Md}$ against $MAE = 7.18$ for $\widetilde{Md}$. The MAE values of the second situation are $MAE = 3.82$ for $\widetilde{Md}$ versus 4.51 for $\widehat{Md}$.

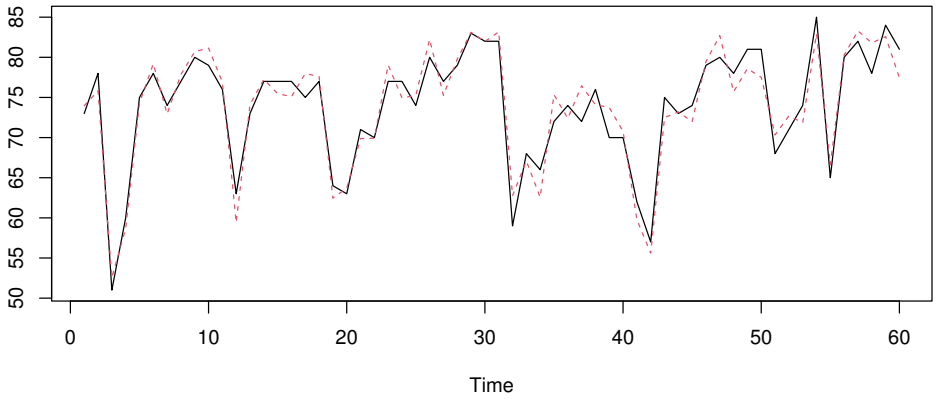


Figure 11. The prediction with the standard mode $\widetilde{Md_\tau}$ without outliers. The true values on black line and the predicted on the red line.

6. Conclusions and Prospects

In this paper, we have investigated the robust estimation of the conditional mode in a functional time series case. The estimator is derived from the robust estimation of the quantile regression. The constructed estimator constitutes an alternative estimator to the standard one based on the conditional density estimation. Moreover, it aims to improve the accuracy as well as the robustness of the conditional mode prediction. This statement is confirmed by the obtained asymptotic result. Indeed, we show that the new estimator can be defined by one kernel and one bandwidth parameter, which allows reducing the almost complete convergence rate.

The second component of this contribution is the specific functional structure, which is modeled using the single-index smoothing algorithm combined with the strong mixing assumption. We have explored this feature in the expression of the convergence. In this sense, the consistency of the estimator is strongly affected by the level of correlation, as well as the choice of the functional index. This aspect is explored in the empirical analysis. Furthermore, the simulation study proves the easy implementation of the constructed estimator. Moreover, it confirms the robustness and the accuracy of the predictor compared to other alternative estimators. From this empirical analysis, we conclude that the performance of the predictor is impacted by the selection of the different parameters involved in the estimator. The arbitrary choice of these parameters can negatively impact the quality of the prediction. For this reason, we can say that the nonexistence of an automatic procedure to select the best bandwidth and the best functional form is the principal concern that negatively affects the applicability of the constructed estimator. In fact, such an issue constitutes an important prospect for the future.

Additionally, for the last open question, we also list the local linear estimator of the modal regression. In this sense, we will combine the robust estimation with the local linear approaches. Recall that local linear modeling is an important algorithm for improving the bias term in the convergence rate. Finally, it is well known that the asymptotic normality of the robust estimator is interesting for the estimation of the confidence interval. Thus, the statement of this asymptotic result is also important for practical purposes.

7. The Mathematical Development

This section is dedicated to proving the asymptotic results of the paper.

The proof of Theorem 1 is based on some standard analytical arguments. Indeed, we write

$$\begin{aligned}\widehat{Md_\tau}(v) - Md_\tau(v) &= \widehat{QR}(\widehat{q_\tau}|v) - QR(q_\tau|v)\\ &= \widehat{QR}(\widehat{q_\tau}|v) - QR(\widehat{q_\tau}|v) + QR(\widehat{q_\tau}|v) - QR(q_\tau|v)\\ &\leq \max_{q\in[a_v,b_v]} |\widehat{QR}(q|v) - QR(q|v)| + QR(\widehat{q_\tau}|v) - QR(q_\tau|v)\end{aligned} \tag{7}$$

From Taylor expansion,

$$QR(\widehat{q_\tau}|v) - QR(q_\tau|v) = (\widehat{q_\tau} - q_\tau)QR'(q_\tau^*|v), \quad q_\tau^* \text{ being in } (\widehat{q_\tau}, \quad q_\tau). \tag{8}$$

Because q_τ is a minimizer of $QR'(\cdot|v)$,

$$QR'(\widehat{q_\tau}|v) - QR'(q_\tau|v) = (\widehat{q_\tau} - q_\tau)^2 t'''(q_\tau^{**}|v), \text{ where } q_\tau^{**} \in (\widehat{q_\tau}, \quad q_\tau). \tag{9}$$

Moreover,

$$\begin{aligned}QR'(\widehat{q_\tau}|v) - QR'(q_\tau|v) &= QR'(\widehat{q_\tau}|v) - \widehat{QR'}(\widehat{q_\tau}|v) + \widehat{QR'}(\widehat{q_\tau}|v) - QR(q_\tau|v)\\ &\leq |QR'(\widehat{q_\tau}|v) - \widehat{QR'}(\widehat{q_\tau}|v)| +\\ &\quad |\min_q \widehat{QR'}(q|v) - \min_q QR'(q|v)|\\ &\leq 2\max_{q\in[a_v,b_v]} |\widehat{QR'}(q|v) - QR'(q|v)|\end{aligned} \tag{10}$$

Combining Equations (7)–(10),

$$|(\widehat{Md_\tau}(v) - Md_\tau(v)| \leq C\left(\max_{q\in[a_v,b_v]} |\widehat{QR}(q|v) - QR(q|v)| + \sqrt{\max_{q\in[a_v,b_v]} |\widehat{QR'}(q|v) - QR'(q|v)|}\right).$$

The definition $\widehat{QR'}(q|v)$ gives

$$\begin{aligned}\widehat{QR'}(q|v) - QR'(q|v) &= \frac{\widehat{QR}(q+\ell_n|v) - QR(q+\ell_n|v) + QR(q-\ell_n|v) - \widehat{QR}(q-\ell_n|v)}{2\ell_n}\\ &\quad + \frac{QR(q+\ell_n|v) - QR(q|v) + QR(q|v) - QR(q-\ell_n|v) - 2\ell_n QR'(q|v)}{2\ell_n}\\ &\leq C\ell_n^{-1} \max_{q\in(a_v-\ell_n,b_v+\ell_n)} |\widehat{QR}(q|v) - QR(q|v)| + O(\ell_n)\end{aligned}$$

So,

$$\max_{q\in[0,1]} |\widehat{QR}(q|v) - QR(q|v)|.$$

Thus, Theorem 1 is a consequence of Proposition 1.

Proposition 1. *Under the postulates of Theorem 1, we have*

$$\max_{q\in(0,1)} |\widehat{QR}(q|v) - QR(q|v)| = O(\ell_n^b) + O_{a.co.}\left(\left(\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\,\zeta_v^2(\ell_n)}\right)^{\frac{1}{2}}\right).$$

The proof is based on

Lemma 1 (see [32]). *Let $\mathbf{Z}_n$ be a sequence of decreasing real random functions, and let L_n be a random real sequence such that*

$$L_n = o_{a.co.}(1) \quad \text{and} \quad \sup_{|\varsigma|\leq M} |\mathbf{Z}_n(\varsigma) + \lambda\varsigma_n - L_n| = o_{a.co.}(1) \quad \text{for certain constants} \quad \lambda, M > 0.$$

Then, for any real random sequence δ_n such that $\mathbf{Z}_n(\varsigma_n) = o_{a.co.}(1)$, we have

$$\sum_{n=1}^{\infty} \mathbb{P}\{|\varsigma_n| \geq M\} < \infty.$$

Proof. Apply Lemma 1 on

$$\mathbf{Z}_n(\varsigma) = \frac{1}{n\mathbb{E}[\mathbf{H}_1]} \sum_{i=1}^{n} (q - \mathbb{1}_{S_i} \leq (\varsigma + QR(q|v)))\mathbf{H}_i,$$

$$\varsigma_n = \widehat{QR}(q|v) - QR(q|v) \quad \text{and} \quad L_n = \mathbf{Z}_n(0).$$

where $\mathbf{H}(\ell_n^{-1} d_\tau(u, T_i))$. Observe that $\mathbf{Z}_n(\varsigma_n) = 0$. So, it suffices

$$L_n = o_{a.co.}(1) \tag{11}$$

$\exists M, \chi > 0$ such that

$$\max_{|\varsigma|\leq M} |\mathbf{Z}_n(\varsigma) + \chi\varsigma - L_n| = o_{a.co.}(1). \tag{12}$$

For Equation (11), we put

$$\xi_i(z) = \Big(q - \mathbb{1}_{[S_i \leq QR(q|v)]}\Big)\mathbf{H}_i - \mathbb{E}\Big[\Big(q - \mathbb{1}_{[S_i \leq QR(q|v)]}\Big)\mathbf{H}_i\Big].$$

Then,

$$L_n - \mathbb{E}[L_n] = \frac{1}{n\mathbb{E}[\mathbf{H}_1]} \sum_{i=1}^{n} \xi_i(z).$$

implies

$$\begin{aligned} \forall \epsilon > 0 \qquad \mathbb{P}\{|L_n - \mathbb{E}[L_n]| > \varepsilon\} &= \mathbb{P}\left\{\left|\frac{1}{n\mathbb{E}[\mathbf{H}_1]} \sum_{i=1}^{n} \xi_i(z)\right| > \varepsilon\right\} \\ &\leq \mathbb{P}\left\{\left|\sum_{i=1}^{n} \xi_i(z)\right| > \varepsilon n\mathbb{E}[\mathbf{H}_1]\right\}. \end{aligned}$$

We evaluate asymptotically

$$\begin{aligned} \mathcal{S}_n^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(\xi_i(z), \xi_j(z)) \\ &= \sum_{i=1}^{n}\sum_{i\neq j} \mathrm{Cov}(\xi_i(z), \xi_j(z)) + nVar[\xi_1(z)]. \end{aligned}$$

For the covariance part, we split the sum into two sets defined by

$$\mathcal{S}_1 = \{(i, j) \text{ such that } 1 \leq i - j \leq u_n\}$$

and

$$\mathcal{S}_2 = \{(i,j) \text{ such that } u_n + 1 \leq i - j \leq n - 1\}.$$

Let $J_{1,n}$ and $J_{2,n}$ be the sum of covariance over S_1 and S_2, respectively. Obviously,

$$\begin{aligned} J_{1,n} &= \sum_{S_1} \left|\mathrm{Cov}(\xi_i(z), \xi_j(z))\right| \\ &\leq C \sum_{S_1} \left|\mathbb{E}\left[\mathbf{H}_i \mathbf{H}_j\right]\right| + \left|\mathbb{E}[\mathbf{H}_i]\mathbb{E}\left[\mathbf{H}_j\right]\right|. \end{aligned}$$

From (**AS**1), (**AS**3) and (**AS**4), we have

$$J_{1,n} \leq C n u_n \xi_v(\ell_n).$$

Next, for the quantity $J_{2,n}$ we use Davydov–Rio's inequality, to show that

$$|\mathrm{Cov}(\xi_i(z), \xi_j(z))| \leq C\alpha(|i-j|).$$

We deduce

$$J_{2,n} = \sum_{S_2} |\mathrm{Cov}(\xi_i(z), \xi_j(z))| \leq \frac{n u_n^{-a+1}}{a-1}.$$

Taking

$$u_n = \left(\frac{1}{\xi_v(\ell_n)}\right)^{1/a},$$

to

$$\sum_{i=1}^{n} \sum_{i \neq j} \mathrm{Cov}(\xi_i(z), \xi_j(z)) = O\Big(n \xi_v(\ell_n)^{(a-1)/a}\Big).$$

Next, from (**AS**1) to (**AS**4) we obtain

$$\mathbb{V}ar(\xi_1(z)) \leq \mathbb{E}\Big[\mathbf{H}_i^2\Big] = O(\zeta_v(\ell_n)).$$

We conclude

$$\mathcal{S}_n^2 = O\Big(n \xi_v^{(1/2)}(\ell_n)\Big). \tag{13}$$

Finally, from Fuk–Nagaev's inequality on $\xi_i(z)$

$$\begin{aligned} \mathbb{P}\{L_n - \mathbb{E}[L_n] > \varepsilon\} &\leq \mathbb{P}\left\{\left|\sum_{i=1}^{n} \xi_i(z)\right| > \varepsilon n \mathbb{E}[\mathbf{H}_1]\right\} \\ &\leq C(C_1 + C_2), \end{aligned}$$

where for all $\zeta > 0$ and $\varepsilon > 0$,

$$C_1 = \left(1 + \frac{\varepsilon^2 n^2 (\mathbb{E}[\mathbf{H}_1])^2}{\mathcal{S}_n^2 \zeta}\right)^{-\zeta/2} \text{ and } C_2 = n\zeta^{-1}\left(\frac{\zeta}{\varepsilon n \mathbb{E}[\mathbf{H}_1]}\right)^{a+1}.$$

For

$$\varepsilon = \epsilon_0 \frac{\sqrt{n \ln n \xi_v^{1/2}(\ell_n)}}{n\mathbb{E}[\mathbf{H}_1]} \text{ and } \zeta = C(\ln n)^2,$$

to obtain,

$$\begin{aligned} C_2 &\leq C n^{1-(a+1)/2} \xi_v(\ell_n)^{-(a+1)/4} (\ln n)^{(3a-1)/2} \\ &\leq C n^{-1-c_1}, \end{aligned} \tag{14}$$

and

$$C_1 \leq C\left(1+\frac{\lambda^2 \ln n}{\zeta}\right)^{-\zeta/2} \leq Cn^{-1-c_2}, \tag{15}$$

for $c_1 > 0$ and $c_2 > 0$. Combining Equations (14) and (15) to conclude that

$$L_n - \mathbb{E}[L_n] = O_{a.co.}\left(\sqrt{\frac{\ln n \xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right). \tag{16}$$

Further,

$$\begin{aligned}\mathbb{E}[L_n] &= \frac{1}{\mathbb{E}[\mathbf{H}_1]}\mathbb{E}\Big[\Big(q - \mathbb{1}_{[S_1 \leq QR(q|v)]}\Big)\mathbf{H}_1\Big] \\ &\leq \frac{1}{\mathbb{E}[\mathbf{H}_1]}\mathbb{E}|CF_\tau(QR(q|v)|v) - CF_\tau(QR(q|v)|v)\mathbf{H}_1| \\ &= O(\ell_n^b).\end{aligned}$$

Therefore,

$$|L_n| = O(\ell_n^b) + O_{a.co.}\left(\sqrt{\frac{\ln n \xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)$$

For Equation (12), we have

$$\max_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma) - L_n - \mathbb{E}[\mathbf{Z}_n(\varsigma) - L_n]| = O_{a.co.}\left(\sqrt{\frac{\ln n \xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right) \tag{17}$$

and the bias

$$\max_{|\varsigma|\leq M}|\mathbb{E}[\mathbf{Z}_n(\varsigma) - L_n] + CF_\tau(t_q(u)|v)\varsigma| = O(\ell_n^b). \tag{18}$$

For Equation (17), we write

$$[-M, M] \subset \bigcup_{j=1}^{d_n}[\varsigma_j - l_n, \varsigma_j + l_n], \; for \; \varsigma_j \in [-M, M] \text{ and } l_n = d_n^{-1} = 1/\sqrt{n}.$$

So, $\forall \varsigma \in [-M, M]$ and we put $j(\varsigma) = \arg\min_j |\varsigma - \varsigma_j|$. Thus

$$\begin{aligned}\max_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma) - L_n - \mathbb{E}[\mathbf{Z}_n(\varsigma) - L_n]| &\leq \max_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma) - \mathbf{Z}_n(\varsigma_{j(\varsigma)})| \\ &+ \max_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma_{j(\varsigma)}) - L_n - \mathbb{E}\Big[\mathbf{Z}_n(\varsigma_{j(\varsigma)}) - L_n\Big]| + \max_{|\varsigma|\leq M}|\mathbb{E}\Big[\mathbf{Z}_n(\varsigma) - \mathbf{Z}_n(\varsigma_{j(\varsigma)})\Big]|.\end{aligned}$$

Since $|\mathbb{1}_{[Y<a]} - \mathbb{1}_{[Y<b]}| \leq \mathbb{1}_{[|Y-b|\leq|a-b|]}$ then

$$\max_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma) - \mathbf{Z}_n(\varsigma_j)| \leq \frac{1}{n\mathbb{E}[\mathbf{H}_1]}\sum_i \mathcal{X}_i$$

where

$$\mathcal{X}_i = \max_{|\varsigma|\leq M}\mathbb{1}_{\left[|S_i - \varsigma_{j(\varsigma)} - QR(q|v)|\leq Cl_n\right]}\mathbf{H}_i.$$

Once again, we determine the asymptotic behavior of the variance term to apply Fuk–Nagaev's inequality. Indeed

$$\mathcal{S}_n^2 = Var\left[\sum_{i=1}^{n} \mathcal{X}_i\right].$$

So,

$$\begin{aligned}\mathcal{S}_n^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(\mathcal{X}_i, \mathcal{X}_j) \\ &= \sum_{i=1}^{n}\sum_{i\neq j} \mathrm{Cov}(\mathcal{X}_i, \mathcal{X}_j) + nVar[\mathcal{X}_1].\end{aligned}$$

We split the sum into two sets defined by

$$E_1 = \{(i,j) \text{ such that } 1 \leq i-j \leq u_n\},$$

$$E_2 = \{(i,j) \text{ such that } u_n + 1 \leq i-j \leq n-1\}.$$

Put $I_{1,n}$ and $I_{2,n}$, the sum of covariance, over E_1 and E_2, respectively. We have

$$\begin{aligned}I_{1,n} &= \sum_{E_1} \left|\mathrm{Cov}(\mathcal{X}_i, \mathcal{X}_j)\right| \\ &\leq Cl_n^2 \sum_{E_1} \left|\mathbb{E}\left[\mathbf{H}_i\mathbf{H}_j\right]\right| + \left|\mathbb{E}[\mathbf{H}_i]\mathbb{E}\left[\mathbf{H}_j\right]\right|.\end{aligned}$$

By (**AS**3) and (**AS**5)

$$I_{1,n} \leq Cnl_n^2 u_n \xi_v(\ell_n).$$

Next, we use Davydov–Rio's inequality for

$$q > \frac{2a-2}{a-2}$$

to show

$$|\mathrm{Cov}(\mathcal{X}_i, \mathcal{X}_j)| \leq Cl_n^{2/q} \xi_v^{1/q}(\ell_n) \alpha^{1-2/q}(|i-j|).$$

Hence,

$$I_{2,n} = \sum_{E_2} |\mathrm{Cov}(\mathcal{X}_i, \mathcal{X}_j)| \leq Cn(\xi_v(\ell_n) l_n^2)^{1/q} u_n^{-a(q-2)/q+1}.$$

Choose

$$u_n = \left(\xi_v(\ell_n) l_n^2\right)^{-(q-1)/a(q-2)}.$$

Thus,

$$\sum_{i=1}^{n}\sum_{i\neq j} \mathrm{Cov}(\mathcal{X}_i, \mathcal{X}_j) = o\left(nl_n \xi_v^{1/2}(\ell_n)\right).$$

From (**AS**1) and (**AS**4), we obtain

$$\mathbb{V}ar(\mathcal{X}_1) = O\left(l_n \xi_v^{1/2}(\ell_n)\right).$$

The Fu–Nagaev inequality implies

$$l_n = o\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n) \ln n}{n\zeta_v^2(\ell_n)}}\right) \tag{19}$$

to obtain

$$\sup_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma)-\mathbf{Z}_n(\varsigma_j)|=O_{a.co.}\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\zeta_v^2(\ell_n)}}\right). \tag{20}$$

For the last term,

$$\sup_{|\varsigma|\leq M}|\mathbb{E}\big[\mathbf{Z}_n(\varsigma)-\mathbf{Z}_n(\varsigma_j)\big]|\leq\frac{1}{n\mathbb{E}[\mathbf{H}_1]}\mathbb{E}[\mathcal{X}_1]\leq Cl_n$$

By Equation (19) we obtain

$$\sup_{|\varsigma|\leq M}|\mathbb{E}\big[\mathbf{Z}_n(\varsigma)-\mathbf{Z}_n(\varsigma_j)\big]|=o\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\zeta_v^2(\ell_n)}}\right).$$

Therefore,

$$\sup_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma_j)-L_n-\mathbb{E}\big[\mathbf{Z}_n(\varsigma_j)-L_n\big]|.$$

implies

$$\mathbf{Z}_n(\varsigma_j)-L_n-\mathbb{E}\big[\mathbf{Z}_n(\varsigma_j)-L_n\big]=\frac{1}{n\mathbb{E}[\mathbf{H}_1]}\sum_{i=1}^{n}\Gamma_i$$

with

$$\Gamma_i=\Big(\mathbb{1}_{S_i\leq t(p|x)}-\mathbb{1}_{S_i\leq\xi_j+t(p|x)}\Big)\mathbf{H}_i-\mathbb{E}\left[\Big(\mathbb{1}_{S_i\leq t(p|x)}-\mathbb{1}_{S_i\leq\xi_j+t(p|x)}\Big)\mathbf{H}_i\right].$$

Reasoning by the same manner as in L_n, we show that

$$Var\left[\sum_{i=1}^{n}\Gamma_i\right]=O\Big(n\xi_v^{(1/2)}(\ell_n)\Big).$$

Consequently, $\exists\eta>0$ such that

$$\begin{aligned}\sum_n\mathbb{P}&\left(\max_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma_{j(\varsigma)})-L_n-\mathbb{E}\Big[\mathbf{Z}_n(\varsigma_{j(\varsigma)})-L_n\Big]|\geq\eta\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)\right)\\&\leq\sum_n d_n\max_j\mathbb{P}\left(|\mathbf{Z}_n(\varsigma_j)-L_n-\mathbb{E}\big[\mathbf{Z}_n(\varsigma_j)-L_n\big]|\geq\eta\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)\right)<\infty\end{aligned}$$

which proves Equation (17).
Concerning Equation (18),

$$\begin{aligned}\mathbb{E}[\mathbf{Z}_n(\varsigma)-L_n]&=-\frac{1}{\mathbb{E}[\mathbf{H}_1]}\mathbb{E}\Big[\Big(\mathbb{1}_{[S_1\leq\varsigma+QR(q|v)]}-\mathbb{1}_{[S_1\leq QR(q|v)]}\Big)\mathbf{H}_i\Big]\\&=-\frac{1}{\mathbb{E}[\mathbf{H}_1]}\mathbb{E}[(CF_\tau(\varsigma+QR(q|v)|v)-CF_\tau(QR(q|v)|v))\mathbf{H}_1]\\&=-\frac{1}{\mathbb{E}[\mathbf{H}_1]}\mathbb{E}[(CF_\tau(\varsigma+QR(q|v)|v)-CF_\tau(QR(q|v)|v))\mathbf{H}_1]+O(\ell_n^b)\\&=-\frac{\varsigma CF_\tau(QR(q|v)|v)}{\mathbb{E}[\mathbf{H}_1]}\mathbb{E}[\mathbf{H}_1]+O(\ell_n^b)+o(\varsigma).\end{aligned}$$

implying

$$\mathbb{E}[\mathbf{Z}_n(\varsigma) - L_n] \quad = \quad -CF_\tau(t_q(u)|v)\varsigma + O(\ell_n^b) + o(\varsigma)$$

Therefore,

$$\widehat{QR}(q|v) - QR(q|v) = \varsigma_n = \frac{1}{CF_\tau(t_q(u)|v)} L_n + O\Big(\max_{|\varsigma|\le M} |\mathbf{Z}_n(\varsigma) + \chi\varsigma - L_n|\Big). \tag{21}$$

Now, from the Bahadur representation, we demonstrate the uniform consistency of $\widehat{QR}(q|v) - QR(q|v)$. We prove

$$\max_{q\in[0,1]} |L_n - \mathbb{E}[L_n]| \quad \text{and} \quad \max_{q\in[0,1]} \max_{|\xi|\le M} |\mathbf{Z}_n(\varsigma) - L_n - \mathbb{E}[\mathbf{Z}_n(\varsigma) - L_n]|.$$

We demonstrate and prove only the second one. From the compactness $[0,1]$,

$$[0,1] \subset \bigcup_{k=1}^{d_n} [q_k - l_n, q_k + l_n], \; for \; q_k \in [0,1].$$

So, $\forall q \in [0,1]$ and we put $H(q) = \arg\min_k |q - q_k|$ and

$$\mathbf{U}_n(\varsigma, q) = \mathbf{Z}_n(\varsigma) - L_n.$$

Therefore,

$$\begin{aligned}
&\max\nolimits_{|\varsigma|\le M} \max\nolimits_{q\in[0,1]} |\mathbf{U}_n(\varsigma,q) - \mathbb{E}[\mathbf{U}_n(\varsigma,q)]| \\
&+ \le \max_{|\varsigma|\le M} \max_{q\in[0,1]} |\mathbf{U}_n(\varsigma,q) - \mathbf{U}_n(\varsigma_{j(\varsigma)},q)| \\
&+ \max_{|\varsigma|\le M} \max_{q\in[0,1]} |\mathbf{U}_n(\varsigma_{j(\varsigma)},q) - \mathbf{U}_n(\varsigma_{j(\varsigma)},q_{H(q)})| \\
&+ \max_{|\varsigma|\le M} \max_{q\in[0,1]} |\mathbf{U}_n(\varsigma_{j(\varsigma)},q_{H(q)}) - \mathbb{E}[\mathbf{U}_n(\varsigma_{j(\varsigma)},q_{H(q)})]| \\
&+ \max_{|\varsigma|\le M} \max_{q\in[0,1]} |\mathbb{E}[\mathbf{U}_n(\varsigma_{j(\varsigma)},q_{H(q)})] - \mathbb{E}[\mathbf{U}_n(\varsigma,q_{H(q)})] \\
&+ \max_{|\varsigma|\le M} \max_{q\in[0,1]} |\mathbb{E}[\mathbf{U}_n(\varsigma,q_{H(q)})] - \mathbb{E}[\mathbf{U}_n(\varsigma,q)].
\end{aligned}$$

Thus

$$\max_{|\varsigma|\le M} \max_{q\in[0,1]} |\mathbf{U}_n(\varsigma,q) - \mathbf{U}_n(\varsigma_{j(\varsigma)},q)| \le \frac{1}{n\mathbb{E}[\mathbf{H}_1]} \sum_i \mathcal{X}_i^0$$

with

$$\mathcal{X}_i^0 = \max_{|\varsigma|\le M} \max_{q\in[0,1]} \mathbb{1}_{[|S_i - \varsigma_{j(\varsigma)} - QR(q|v)|\le Cl_n]} \mathbf{H}_i.$$

$$Var\left[\sum_{i=1}^n \mathcal{X}_i^0\right] = O\Big(nl_n\xi_v^{1/2}(\ell_n)\Big) \quad \text{and} \quad \mathbb{E}[\mathcal{X}_i^0] = O\Big(nl_n\xi_v^{1/2}(\ell_n)\Big)$$

we obtain

$$\max_{|\varsigma|\le M} \max_{q\in[0,1]} |\mathbf{U}_n(\varsigma,q) - \mathbf{U}_n(\varsigma_{j(\varsigma)},q) = O_{a.co.}\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)$$

and

$$\max_{|\varsigma|\leq M}\max_{q\in[0,1]}|\mathbb{E}\Big[\mathbf{U}_n(\varsigma,q)-\mathbf{U}_n(\varsigma_{j(\varsigma)},q)\Big]=o_{a.co.}\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\zeta_v^2(\ell_n)}}\right).$$

Similarly,

$$\max_{|\varsigma|\leq M}\max_{q\in[0,1]}|\mathbf{U}_n(\varsigma_{j(\varsigma)},q)-\mathbf{U}_n(\varsigma_{j(\varsigma)},q_{\mathbf{H}(q)})|\leq\frac{1}{n\mathbb{E}[\mathbf{H}_1]}\sum_i\mathcal{X}_i^1$$

with

$$\mathcal{X}_i^1=\max_{|\varsigma|\leq M}\max_{q\in[0,1]}\Big(\mathbb{1}_{\left[|S_i-\varsigma_{j(\varsigma)}-QR(q|v)|\leq Cl_n\right]}+\mathbb{1}_{\left[|S_i-QR(q_{\mathbf{H}(q)}|v)|\leq Cl_n\right]}\Big)\mathbf{H}_i.$$

Since

$$Var\left[\sum_{i=1}^n\mathcal{X}_i^1\right]=O\Big(nl_n\xi_v^{1/2}(\ell_n)\Big)\quad\text{and}\quad\mathbb{E}[\mathcal{X}_i^1]=O\Big(nl_n\xi_v^{1/2}(\ell_n)\Big)$$

we obtain

$$\max_{|\varsigma|\leq M}\max_{q\in[0,1]}|\mathbf{U}_n(\varsigma,q)-\mathbf{U}_n(\varsigma_{j(\varsigma)},q)=O_{a.co.}\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\zeta_v^2(\ell_n)}}\right)$$

and

$$\max_{|\varsigma|\leq M}\max_{q\in[0,1]}|\mathbb{E}\Big[\mathbf{U}_n(\varsigma,q)-\mathbf{U}_n(\varsigma_{j(\varsigma)},q)\Big]=o\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\zeta_v^2(\ell_n)}}\right).$$

Putting

$$\mathbf{U}_n(\varsigma_j,q_k)-\mathbb{E}[\mathbf{U}_n(\varsigma_j,q_k)]=\frac{1}{n\mathbb{E}[\mathbf{H}_1]}\sum_{i=1}^n\mathcal{Y}_i'$$

where

$$\mathcal{Y}_i'=\Big(\mathbb{1}_{S_i\leq QR(q_k|v)}-\mathbb{1}_{S_i\leq\varsigma_j+QR(q_k|v)}\Big)\mathbf{H}_i$$

$$-\mathbb{E}\Big[\Big(\mathbb{1}_{S_i\leq QR(q_k|v)}-\mathbb{1}_{S_i\leq\varsigma_j+QR(q_k|v)}\Big)\mathbf{H}_i\Big].$$

Reasoning by the same manner as in L_n, we prove that

$$Var\left[\sum_{i=1}^n\mathcal{Y}_i'\right]=O\Big(n\xi_v^{(1/2)}(\ell_n)\Big).$$

Consequently, $\exists\eta>0$ such that

$$\sum_n\mathbb{P}\left(\sup_{|\delta|\leq M}\Big|F_n(\delta_{j(\delta)})-L_n-\mathbb{E}\Big[F_n(\delta_{j(\delta)})-L_n\Big]\Big|\geq\eta\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)\right)$$

$$\leq\sum_n d_n\max_j\mathbb{P}\left(\big|F_n(\delta_j)-L_n-\mathbb{E}\big[F_n(\delta_j)-L_n\big]\big|\geq\eta\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\zeta_v^2(\ell_n)}}\right)\right)<\infty$$

Then, $\eta>0$ such that

$$\sum_n\mathbb{P}\left(\max_{|\xi|\leq M}\max_{q\in[0,1]}|\mathbf{U}_n(\varsigma_{j(\varsigma)},q_{\mathbf{H}(q)})-\mathbb{E}[\mathbf{U}_n(\varsigma_{j(\varsigma)},q_{\mathbf{H}(q)})]|\geq\eta\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)\right)$$

$$\leq\sum_n d_n^2\max_j\max_k\mathbb{P}\left(|\mathbf{U}_n(\varsigma_j,q_k)-\mathbb{E}[\mathbf{U}_n(\varsigma_j,q_k)]|\geq\eta\left(\sqrt{\frac{\xi_v^{1/2}(\ell_n)\ln n}{n\zeta_v^2(\ell_n)}}\right)\right)<\infty.$$

Hence,

$$\max_{q\in[0,1]}\max_{|\varsigma|\leq M}|\mathbf{Z}_n(\varsigma)-L_n-\mathbb{E}[\mathbf{Z}_n(\varsigma)-L_n]|=O_{a.co.}\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)$$

We prove

$$\max_{q\in[0,1]}|L_n-\mathbb{E}[L_n]|=O_{a.co.}\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right)$$

We conclude

$$\max_{q\in[0,1]}|\widehat{QR}(q|v)-QR(q|v)|=O(\ell_n^b)+O_{a.co.}\left(\sqrt{\frac{\ln n\xi_v^{1/2}(\ell_n)}{n\zeta_v^2(\ell_n)}}\right).$$

□

Author Contributions: Conceptualization, A.L.; Methodology, M.B.A. and A.L.; Investigation, F.A.A.; Data curation, Z.K. and A.L.; Writing—original draft, A.L.; Writing—review & editing, Z.K.; Project administration, Z.K.; Funding acquisition, F.A.A. All authors have read and agreed to the final version of the manuscript.

Funding: This research was funded by Princess Nourah bint Abdulrahman University Researchers; Supporting Project number (PNURSP2025R515); Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia; and the Deanship of Scientific Research and Graduate Studies at King Khalid University through the Research Groups Program under grant number R.G.P. 1/163/45.

Data Availability Statement: The data used in this study are available through the link https://kilthub.cmu.edu (accessed on 2 February 2025).

Acknowledgments: The authors would like to thank the referees for their very valuable comments and suggestions, which led to a considerable improvement of the manuscript. The authors thank and extend their appreciation to the funders of this work. This work was supported by Princess Nourah bint Abdulrahman University Researchers; Supporting Project number (PNURSP2025R515); Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, and the Deanship of Scientific Research and Graduate Studies at King Khalid University through the Research Groups Program under grant number R.G.P. 1/163/45.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1 Collomb, G.; Härdle, W.; Hassani, S. A note on prediction via estimation of conditional mode function. *J. Statist. Plan. Inf.* **1987**, *15*, 227–236. [CrossRef]
2 Härdle, W.; Hall, P.; Ichimura, H. Optimal smoothing in single-index models. *Ann. Statist.* **1993**, *21*, 157–178. [CrossRef]
3 Stute, W.; Zhu, L.-X. Nonparametric checks for single-index models. *Ann. Statist.* **2005**, *33*, 1048–1083. [CrossRef]
4 Tang, Q.; Kong, L.; Rupper, D.; Karunamuni, R.J. Partial functional partially linear single-index models. *Statist. Sin.* **2021**, *31*, 107–133. [CrossRef]
5 Zhou, W.; Gao, J.; Harris, D.; Kew, H. Semi-parametric single-index predictive regression models with cointegrated regressors. *J. Econom.* **2024**, *238*, 105577. [CrossRef]
6 Zhu, H.; Zhang, R.; Liu, Y.; Ding, H. Robust estimation for a general functional single index model via quantile regression. *J. Korean Stat. Soc.* **2022**, *51*, 1041–1070. [CrossRef]
7 Ferraty, F.; Peuch, A.; Vieu, P. Modèle à indice fonctionnel simple. *Comptes Rendus Math.* **2003**, *336*, 1025–1028. [CrossRef]
8 Ait-Saïdi, A.; Ferraty, F.; Kassa, R.; Vieu, P. Cross-validated estimations in the single-functional index model. *Statistics* **2008**, *42*, 475–494. [CrossRef]
9 Jiang, Z.; Huang, Z.; Zhang, J. Functional single-index composite quantile regression. *Metrika* **2023**, *86*, 595–603. [CrossRef]

10 Nie, Y.; Wang, L.; Cao, J. Estimating functional single index models with compact support. *Environmetrics* **2023**, *34*, e2784. [CrossRef]
11 Chen, D.; Hall, P.; Müller, H.-G. Single and multiple index functional regression models with nonparametric link. *Ann. Statist.* **2011**, *39*, 1720–1747. [CrossRef]
12 Ling, N.; Xu, Q. Asymptotic normality of conditional density estimation in the single index model for functional time series data. *Statist. Probab. Lett.* **2012**, *82*, 2235–2243. [CrossRef]
13 Attaoui, S. On the nonparametric conditional density and mode estimates in the single functional index model with strongly mixing data. *Sankhya A* **2014**, *76*, 356–378. [CrossRef]
14 Han, Z.-C.; Lin, J.-G.; Zhao, Y.-Y. Adaptive semiparametric estimation for single index models with jumps. *Comput. Stat. Data Anal.* **2020**, *151*, 107013. [CrossRef]
15 Hao, M.; Liu, K.Y.; Su, W.; Zhao, X. Semiparametric estimation for the functional additive hazards model. *Can. J. Stat.* **2024**, *52*, 755–782. [CrossRef]
16 Kowal, D.R.; Canale, A. Semiparametric Functional Factor Models with Bayesian Rank Selection. *Bayesian Anal.* **2023**, *18*, 1161–1189. [CrossRef]
17 Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis*; Springer: New York, NY, USA, 2006.
18 Azzedine, N.; Laksaci, A.; Ould-Saïd, E. On robust nonparametric regression estimation for functional regressor. *Stat. Probab. Lett.* **2008**, *78*, 3216–3221. [CrossRef]
19 Barrientos-Marin, J.; Ferraty, F.; Vieu, P. Locally modelled regression and functional data. *J. Nonparametric Stat.* **2010**, *22*, 617–632. [CrossRef]
20 Demongeot, J.; Hamie, A.; Laksaci, A.; Rachdi, M. Relative-error prediction in nonparametric functional statistics: Theory and practice. *J. Multivar. Anal.* **2016**, *146*, 261–268. [CrossRef]
21 Ezzahrioui, M.; Ouldsaïd, E. Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametric Stat.* **2008**, *20*, 3–18. [CrossRef]
22 Ezzahrioui, M.; Ould-Said, E. Some asymptotic results of a non-parametric conditional mode estimator for functional time-series data. *Stat. Neerl.* **2010**, *64*, 171–201. [CrossRef]
23 Dabo-Niang, S.; Laksaci, A. Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *Pub. Inst. Stat. Univ. Paris* **2007**, *3*, 27–42. [CrossRef]
24 Dabo-Niang, S.; Kaid, Z.; Laksaci, A. Asymptotic properties of the kernel estimate of spatial conditional mode when the regressor is functional. *AStA Adv. Stat. Anal.* **2015**, *99*, 131–160. [CrossRef]
25 Ling, N.; Liu, Y.; Vieu, P. Conditional mode estimation for functional stationary ergodic data with responses missing at random. *Statistics* **2016**, *50*, 991–1013. [CrossRef]
26 Bouanani, O.; Rahmani, S.; Laksaci, A.; Rachdi, M. Asymptotic normality of conditional mode estimation for functional dependent data. *Indian J. Pure Appl. Math.* **2020**, *51*, 465–481. [CrossRef]
27 Engle, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **1982**, *50*, 987–1007. [CrossRef]
28 Jones, D.A. Nonlinear autoregressive processes. *Proc. Roy. Soc. A* **1978**, *360*, 71–95.
29 Bollerslev, T. General autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [CrossRef]
30 Leys, C.; Ley, C.; Klein, O.; Bernard, P.; Licata, L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **2013**, *49*, 764–766. [CrossRef]
31 Febrero, M.; Galeano, P.; González-Manteiga, W. Outlier detection in functional data by depth measures with application to identify abnormal NO_x levels. *Environmetrics* **2008**, *19*, 331–345. [CrossRef]
32 Azzi, A.; Belguerna, A.; Laksaci, A.; Rachdi, M. The scalar-on-function modal regression for functional time series data. *J. Nonparametric Stat.* **2024**, *36*, 503–526. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.