

Article

Modal Regression Estimation by Local Linear Approach in High-Dimensional Data Case

Fatimah A. Almulhim ¹, Mohammed B. Alamari ^{2,*}, Ali Laksaci ² and Zoulikha Kaid ²

¹ Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; faalmulhim@pnu.edu.sa

² Department of Mathematics, College of Science, King Khalid University, Abha 62223, Saudi Arabia; alikfa@kku.edu.sa (A.L.); zqayd@kku.edu.sa (Z.K.)

* Correspondence: malamari@kku.edu.sa

Abstract

This paper introduces a new nonparametric estimator for detecting the conditional mode in the functional input variable setting. The estimator integrates a local linear approach with an L^1 -robust algorithm and treats the modal regression as the minimizer of the quantile derivative. As an asymptotic result, we derive the theoretical properties of the estimator by analyzing its convergence rate under the almost complete consistency framework. The result is stated under standard conditions, characterizing both the functional structure of the data and the local linear approximation properties of the model. Moreover, the expression of the convergence rate retains the usual form of the stochastic convergence rate in functional statistics. Simulations and real-data applications demonstrate the algorithm's effectiveness, showing its advantage over existing methods in high-dimensional prediction tasks.

Keywords: functional statistics; conditional mode; smoothing approach; L^1 method; spectrometry processing; robust estimator; quantile regression

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



Academic Editors: Francisco Novoa-Muñoz and Bernardo M. Lagos-Álvarez

Received: 24 May 2025

Revised: 30 June 2025

Accepted: 11 July 2025

Published: 16 July 2025

Citation: Almulhim, F.A.; Alamari, M.B.; Laksaci, A.; Kaid, Z. Modal Regression Estimation by Local Linear Approach in High-Dimensional Data Case. *Axioms* **2025**, *14*, 537. <https://doi.org/10.3390/axioms14070537>

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

In the last decade, interest in the statistical analysis of functional data has grown significantly, driven by its wide range of applications in various domains. The foundational work in this field was first established by [1,2], as pioneer works, and [3] or [4] for recent results and references. Within this field, nonparametric statistics have proven to be highly effective in modeling functional data, particularly when it comes to predicting using conditional mode and conditional quantiles. However, most of the existing nonparametric functional quantiles or mode estimators heavily depend on traditional kernel methods. In this paper, we improve the accuracy and robustness of the previous estimators using the L^1 -local linear estimation (LLE) method.

Modal and quantile regression are popular tools in nonparametric statistics for making predictions. They present an alternative approach to traditional regression, which is based on conditional expectation. Unlike the conventional approaches, the proposed predictors often impart deeper insights, especially when the data have asymmetric or multimodal conditional densities, or when the white noise shows a heavy-tailed distribution (refer to [5] for the limitation of the conditional expectation). The latter demonstrated that the conditional mode outperforms the conditional mean. Based on this, Ref. [6] created a prediction model that uses the derivative of the conditional density for multivariate data. Theoretical

progress was made by [7], who established the asymptotic distribution of modal regression estimators for independent data. Ref. [8] extended the last result to the dependent cases. Further advancements in the conditional mode prediction are comprehensively covered in [9,10]. While the cited works concern the kernel estimator of the conditional mode (CM) in finite-dimensional cases, the Nadaraya–Watson (NW) estimator was extended for functional data by [11]. They estimated this regression model by maximizing the conditional density and proved its almost complete consistency. This result has been generalized to the functional time series case by [12]. The asymptotic distribution of the NW estimator of the conditional mode was explored for the independent case in [13] and for the strongly mixing case by [14]. There were also significant theoretical advancements made by [15], who determined L^p -convergence rates for NW-based functional CM estimators. For an alternative functional time series, Ref. [16] established stochastic consistency for ergodic functional time series data containing missing at random. Further theoretical and practical contributions in the field of functional CM estimation can be found in [17] and the references therein.

On the other hand, quantile regression has gained increasing interest in nonparametric data analysis. It is worth noting that the quantile regression function is also one of the most well-studied models in nonparametric vector statistics. Earlier foundational works in multivariate statistics go back to [18], and were popularized by [19,20]. Once again in functional statistics, quantile regression has received considerable attention. This high level of flexibility justifies the approach, as it can be treated as a linear, semiparametric, or nonparametric model. The functional linear modeling of the conditional quantile was considered by [21,22], which provides comprehensive overviews of the latest developments. While the semiparametric estimation was developed by [23]. Concerning nonparametric estimation, quantile regression is estimated using the functional version of the NW estimator by [11]. They proved the Borel–Cantelli of the constructed estimator. Ref. [24] has stated the moment integrability of the functional NW estimator. We return to [25,26] for recent results.

The main novelty of this work is the estimation by the LLE method. This approach provides an important advantage compared to the NW estimator. In particular, it permits increasing the efficiency of the estimation by reducing bias. This statement has been observed by [27] in the multivariate case and expanded into functional statistics by [28]. From a bibliographical perspective, the LLE technique has been thoroughly explored in functional statistics by many authors. For example, Ref. [29] demonstrated L_2 consistency of the LLE estimator when estimating the regression operator in Hilbert spaces. Ref. [30] proposes a more general estimator allowing coverage of a wider range of functional covariates, while [31] developed an alternative estimator based on the inverse local covariance operator. Ref. [32] introduced functional LLE-fitting for nonparametric conditional models, establishing pointwise and uniform almost-complete (a.co.) consistency for the conditional density mode. The first results on the LLE-estimation of the quantile function were stated by [33]. Their estimation is obtained by inverting the cumulative distribution function (CDF) estimator introduced by [32]. An alternative was recently considered by [34] based on the M -estimation of the quantile regression.

While most current methods focus on estimating the conditional mode by optimizing a kernel-based conditional density estimator, this paper introduces a new LLE approach based on quantile regression. Unlike traditional methods, our estimator accumulates the robustness of L_1 -regression and the performance of the LLE-estimation, which enhances its resistance to outliers and extends its accuracy as a predictor or classifier in functional statistics. Theoretically, we demonstrate the almost complete convergence of the estimator under standard conditions, with precision in the convergence rate. This new estimator's

benefits are also emphasized by comparing it to the standard one using a simulated and real data example. Our empirical analysis shows that the new estimator not only outperforms the accuracy of the previous estimators but also remains robust across a wide range of scenarios. Additionally, to demonstrate its practical value, we apply our method to predict the chemical composition of cocoa using spectrometry data. Typically, we have focused on the prediction of fat and moisture content using near-infrared spectroscopy data as functional regressors. It is well documented that near-infrared spectroscopy is a powerful, non-destructive tool for rapidly quantifying cocoa quality, allowing real-time quality control in production. Thus, combining this spectrometry processing with the proposed functional model can help chocolate producers to maintain good flavor and texture in their products, ensuring quality at every step.

The paper is organized as follows. In the next Section, we introduce the conditional mode based on the L^1 rule and explain its LLE estimator. Section 3 covers the principal assumptions and the main theoretical findings. In Section 4 We outline the principal features of our approach and provide a comparative analysis with existing methods. The finite-sample performance of our proposed estimator is examined in Section 5, using both simulation studies and real-world data applications. The detailed proofs of the intermediate results are presented in the Appendix Section A.

2. The L^1 Conditional Mode and Its Local Linear Estimator

Let $\{(F_i, R_i)\}_{i=1}^n$ be an independent and identically distributed (i.i.d.) sample from the random pair (F, R) , where F is functional input variable taking values in a semi-metric space $(\mathcal{H}, m)$ (m being the semi-metric) and R is the output real variable. Next, fix $\mathfrak{F} \in \mathcal{H}$ and consider a neighborhood $\mathcal{N}_{\mathfrak{F}}$ of $\mathfrak{F}$. We assume the existence of the regular conditional distribution function $D(\cdot|\mathfrak{F})$ of R given $F = \mathfrak{F}$, which is strictly increasing and has a continuous density $d(\cdot|\mathfrak{F})$ with respect to Lebesgue measure on $\mathbb{R}$. The standard conditional mode is usually defined as the maximizer of this conditional density over a compact set S :

$$MR(\mathfrak{F}) = \arg \max_{y \in S} d(y|\mathfrak{F}). \quad (1)$$

However, in order to introduce a more robust definition of the conditional mode, we reformulate this definition and we relate the modal regression to the quantile derivatives. Indeed, for $q \in D^{-1}(S|\mathfrak{F}) = [a_{\mathfrak{F}}, b_{\mathfrak{F}}] \subset (0, 1)$, let $QR_q(\mathfrak{F})$ denote the conditional quantile of order q given $F = \mathfrak{F}$. Following similar ideas to [35] in the linear case, we assume $QR(\cdot|\mathfrak{F})$ is continuously differentiable (C^1) with $D(QR_q(\mathfrak{F})|\mathfrak{F}) > 0$, yielding:

$$QR'_q(\mathfrak{F}) = \frac{\partial QR_q(\mathfrak{F})}{\partial q} = \frac{1}{D(QR_q(\mathfrak{F})|\mathfrak{F})}. \quad (2)$$

It follows that the conditional mode can be expressed as:

$$MR(\mathfrak{F}) = QR_{q_{MR}}(\mathfrak{F}), \quad \text{where } q_{MR} = \arg \min_{q \in [a_{\mathfrak{F}}, b_{\mathfrak{F}}]} QR'_q(\mathfrak{F}). \quad (3)$$

Recall that the theoretical conditional quantile $QR_q(\mathfrak{F})$ is robustly defined as M -regression with respect to the following score function

$$\min_t \mathbb{E}[L_q(R - t) | F = \mathfrak{F}], \quad (4)$$

where $L_q(y) = y(q - \mathbb{1}_{\{y < 0\}})$ is the quantile loss function. Thus, the LLE estimator of the conditional mode is strongly linked to the LLE estimator of the quantile regression. For this

last point, we adopt the fast LLE approach of [30], approximating $QR_q(\cdot)$ in a neighborhood $\mathcal{N}_{\mathfrak{F}}$ by:

$$QR_q(\mathfrak{Z}) \approx A + Bm(\mathfrak{F}, \mathfrak{Z}) \quad \forall \mathfrak{Z} \in \mathcal{N}_{\mathfrak{F}}, \quad (5)$$

and estimate the coefficients (A, B) as solutions of

$$\min_{(A,B) \in \mathbb{R}^2} \sum_{i=1}^n L_q(R_i - A - Bm(F_i, \mathfrak{F})) \mathbf{F}\left(\frac{m(\mathfrak{F}, F_i)}{f_n}\right), \quad (6)$$

where $\mathbf{F}$ is a kernel function, f_n is a bandwidth sequence. The conditional quantile estimator is then $\widehat{QR}_q(\mathfrak{F}) = \hat{a}$. Consequently, the LLE estimator of the conditional mode is

$$\widehat{MR}(\mathfrak{F}) = \widehat{QR}(\hat{q}_M R | \mathfrak{F}) \quad \text{where} \quad \hat{q}_M R = \arg \min_{q \in [a_{\mathfrak{F}}, b_{\mathfrak{F}}]} \widehat{QR}'(q | \mathfrak{F}), \quad (7)$$

with $\widehat{QR}'$ being the estimator of the derivative of the conditional quantile defined by

$$\widehat{QR}'(q | \mathfrak{F}) = \frac{\widehat{QR}(q + b_n | \mathfrak{F}) - \widehat{QR}(q - b_n | \mathfrak{F})}{2b_n}, \quad (8)$$

where b_n is a positive bandwidth-like sequence.

We point out that the functional NW estimator studied by [17] is a particular case of this study; it corresponds to $B = 0$. Next, let us clarify that the existence and uniqueness of $MR(\mathfrak{F})$ are guaranteed by the continuity and strict monotonicity of $D(\cdot | \mathfrak{F})$, while $\widehat{MR}(\mathfrak{F})$ may not be unique. In what follows, $\widehat{MR}(\mathfrak{F})$ refers to any minimizer in (6). Furthermore, the primary contribution of this work is to establish asymptotic properties of $\widehat{MR}(\mathfrak{F})$ when $\mathfrak{F}$ takes values in the semi-metric space $\mathcal{F}_{\mathcal{N}}$. To our knowledge, this is the first study to employ this estimation strategy for the conditional mode, even in multivariate settings. Note that the finite-dimensional case ($\mathcal{F}_{\mathcal{N}} = \mathbb{R}^p$) emerges as a special case, highlighting the generality of our approach.

3. Main Results

All along the paper, when no confusion is possible, we will denote by C or C' some strictly positive generic constants, and we set $B(\mathfrak{F}, f) = \{\mathfrak{F}' \in \mathcal{F} : m(\mathfrak{F}', \mathfrak{F}) < f\}$. Now, we list some required conditions that are necessary in deriving the almost complete convergence of $\widehat{MR}(\mathfrak{F})$ of $MR(\mathfrak{F})$.

(B1) For any $r > 0$, $\alpha_{\mathfrak{F}}(f) := \alpha_{\mathfrak{F}}(-f, f) > 0$ and there exists a function $\beta_{\mathfrak{F}}(\cdot)$ such that:

$$\text{for all } t \in (-1, 1), \quad \lim_{f \rightarrow 0} \frac{\alpha_{\mathfrak{F}}(tf_n, f_n)}{\alpha_{\mathfrak{F}}(f_n)} = \beta_{\mathfrak{F}}(t)$$

(B2) The functions $QR(\cdot | \mathfrak{F})$ is of class $C^3([a_{\mathfrak{F}}, b_{\mathfrak{F}}])$ and $D(\cdot | \mathfrak{F})$ satisfies the following Lipschitz condition:

$$\text{for all } (\mathfrak{F}_1, \mathfrak{F}_2) \in N_{\mathfrak{F}}, \quad |D(t | \mathfrak{F}_1) - D(t | \mathfrak{F}_2)| \leq Cd^b(\mathfrak{F}_1, \mathfrak{F}_2) \text{ for some } b > 0,$$

where $N_{\mathfrak{F}}$ denotes a neighborhood of $\mathfrak{F}$.

(B3) The kernel $\mathbf{F}$ is a positive and differentiable function which is supported within $(-1, 1)$, and such that:

$$\begin{pmatrix} \mathbf{F}(1) - \int_{-1}^1 \mathbf{F}'(t) \beta_{\mathfrak{F}}(t) dt & \mathbf{F}(1) - \int_{-1}^1 (t\mathbf{F}(t))' \beta_{\mathfrak{F}}(t) dt \\ \mathbf{F}(1) - \int_{-1}^1 (t\mathbf{F}(t))' \beta_{\mathfrak{F}}(t) dt & \mathbf{F}(1) - \int_{-1}^1 (t^2\mathbf{F}(t))' \beta_{\mathfrak{F}}(t) dt \end{pmatrix}$$

is a positive definite matrix.

(B4) The bandwidth f_n satisfies:

$$\lim_{n \rightarrow \infty} \frac{\log n}{nb_n \alpha_{\mathfrak{F}}(f_n)} = 0$$

The assumed conditions are not very restrictive. They allow us to analyze various aspects of our subject, such as model structure, data framework, and convergence rates. Considering the complexity of our proposed local linear algorithm and the strength of the Borel–Cantelli (BC) consistency, these assumptions offer a good compromise between generality and analytical feasibility. Specifically, each element of the subject is explored through a distinct assumption. For instance, (B1) is particularly essential in NFDA. Specifically, Condition (B1) is indispensable in NFDA, and the function $\alpha_{\mathfrak{F}}(\cdot)$ can be explicitly given for several continuous processes (see [11]). While the regularity conditions are necessary to explore the functional space of the nonparametric path under study, this assumption significantly impacts the bias term in the convergence rate of $\widehat{MR}(\mathfrak{F})$. Finally, conditions (B3)–(B4) relate to the kernel F and the bandwidths f_n and b_n , which cover the technical aspects of the estimator $\widehat{MR}(\mathfrak{F})$. These conditions are less restrictive than the usual technical requirements in conditional mode estimation. Unlike the conditional density approach, which relies on two kernels, our method requires only one. It is important to recall that these conditions are necessary to determine the convergence rate of the kernel estimator under the Borel–Cantelli mode. Typically, a weaker form of consistency—distinct from BC consistency—can be established for the estimator under more restrictive assumptions. The latter is obtained by evaluating the expectation and variance of the estimator. Although this probability consistency is weaker than Borel–Cantelli consistency, it remains sufficient to justify the estimator’s practical applicability. Nevertheless, because this paper aims to demonstrate both the theoretical generality of the model and its practical viability, we have chosen to emphasize the stronger consistency result, which implies the weaker.

The following theorem establishes the almost-complete convergence (cf. [2] for details), (a.co.), of $\widehat{MR}(\mathfrak{F})$. This kind of convergence implies both almost sure convergence and convergence in probability.

Theorem 1. Under assumptions (B1)–(B4) and if $\inf_{q \in (0,1)} \frac{\partial^3 QR_q(\mathfrak{F})}{\partial q} > 0$ we have:

$$\widehat{MR}(\mathfrak{F}) - MR(\mathfrak{F}) = O(f_n^{b/2}) + O(b_n^{1/2}) + O\left(\left(\frac{\log n}{nb_n^2 \alpha_{\mathfrak{F}}(f_n)}\right)^{\frac{1}{4}}\right) \quad a.co.$$

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

$$\widehat{MR}(\mathfrak{F}) - MR(\mathfrak{F}) = \widehat{QR}_{\widehat{q}_{MR}}(\mathfrak{F}) - QR_{\widehat{q}_{MR}}(\mathfrak{F}) + QR_{\widehat{q}_{MR}}(\mathfrak{F}) - QR_{q_{MR}}(\mathfrak{F})$$

Trivially,

$$\widehat{MR}(\mathfrak{F}) - MR(\mathfrak{F}) \leq \sup_{q \in [a_{\mathfrak{F}}, b_{\mathfrak{F}}]} |\widehat{QR}(q|\mathfrak{F}) - t(q|\mathfrak{F})| + QR_{\widehat{q}_{MR}}(\mathfrak{F}) - QR_{q_{MR}}(\mathfrak{F})$$

By Taylor expansion,

$$QR_{\widehat{q}_{MR}}(\mathfrak{F}) - QR_{q_{MR}}(\mathfrak{F}) = (\widehat{q}_{MR} - q_{MR}) QR'_{q_{MR}^*}(\mathfrak{F}), \quad q_{MR}^* \text{ being in } (\widehat{q}_{MR}, q_{MR}). \quad (9)$$

Since q_{MR} is minimizer of $QR'(\mathfrak{F})$,

$$QR'_{\widehat{q_{MR}}}(\mathfrak{F}) - QR'_{q_{MR}}(\mathfrak{F}) = (\widehat{q_{MR}} - q_{MR})^2 QR'''_{q_{MR}^{**}}(\mathfrak{F}), \text{ where } q_{MR}^{**} \in (\widehat{q_{MR}}, q_{MR}). \quad (10)$$

In addition,

$$\begin{aligned} QR'_{\widehat{q_{MR}}}(\mathfrak{F}) - QR'_{q_{MR}}(\mathfrak{F}) &= QR'_{\widehat{q_{MR}}}(\mathfrak{F}) - \widehat{QR'}_{\widehat{q_{MR}}}(\mathfrak{F}) + \widehat{QR'}_{\widehat{q_{MR}}}(\mathfrak{F}) - QR_{q_{MR}}(\mathfrak{F}) \\ &\leq |QR'_{\widehat{q_{MR}}}(\mathfrak{F}) - \widehat{QR'}_{\widehat{q_{MR}}}(\mathfrak{F})| + |\min_q \widehat{QR'}_q(\mathfrak{F}) - \min_q QR'_q(\mathfrak{F})| \\ &\leq 2 \sup_{q \in [a_{\mathfrak{F}}, b_{\mathfrak{F}}]} |\widehat{QR'}_q(\mathfrak{F}) - QR'_q(\mathfrak{F})| \end{aligned} \quad (11)$$

Thus,

$$|\widehat{MR}(\mathfrak{F}) - MR(\mathfrak{F})| \leq C \left(\sup_{q \in [a_{\mathfrak{F}}, b_{\mathfrak{F}}]} |\widehat{QR}_q(\mathfrak{F}) - QR_q(\mathfrak{F})| + \sqrt{\sup_{q \in [a_{\mathfrak{F}}, b_{\mathfrak{F}}]} |\widehat{QR'}_q(\mathfrak{F}) - QR'_q(\mathfrak{F})|} \right).$$

Now, for a large enough n

$$\begin{aligned} \widehat{QR'}_q(\mathfrak{F}) - QR'_q(\mathfrak{F}) &= \frac{\widehat{QR}_{q+b_n}(\mathfrak{F}) - QR_{q+b_n}(\mathfrak{F}) + QR_{q-b_n}(\mathfrak{F}) - \widehat{QR}_{q-b_n}(\mathfrak{F})}{2b_n} \\ &\quad + \frac{QR_{q+b_n}(\mathfrak{F}) - QR_q(\mathfrak{F}) + \widehat{QR}_q(\mathfrak{F}) - QR_{q-b_n}(\mathfrak{F}) - 2b_n QR'_q(\mathfrak{F})}{2b_n} \\ &\leq C b_n^{-1} \sup_{q \in (a_{\mathfrak{F}} - b_n, b_{\mathfrak{F}} + b_n)} |\widehat{QR}_q(\mathfrak{F}) - QR_q(\mathfrak{F})| + O(b_n) \end{aligned}$$

Thus, it suffices to evaluate, for a large enough n

$$\sup_{q \in [0,1]} |\widehat{QR}_q(\mathfrak{F}) - QR_q(\mathfrak{F})|.$$

Thus, Theorem 1 is a consequence of the following proposition. $\square$

Proposition 1. Under assumptions of Theorem 1, we have

$$\sup_{q \in (0,1)} |\widehat{QR}_q(\mathfrak{F}) - QR_q(\mathfrak{F})| = O(f_n^b) + O_{a.co.} \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{\frac{1}{2}} \right)$$

The proof of this proposition is based on the Bahadur representation of the conditional quantile. The latter is a consequence of the following lemmas.

Lemma 1. Let Π_n be a sequence of decreasing real random functions and Y_n be a random real sequence such that

$$Y_n = o_{a.co.}(1) \quad \text{and} \quad \sup_{|v| \leq M} |\Pi_n(v) + \lambda v - Y_n| = o_{a.co.}(1) \quad \text{for certain constants } \lambda, M > 0.$$

Then, for any real sequence v_n such that $\Pi_n(v_n) = o_{a.co.}(1)$, we have:

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

4. On the Potential Impact of the Contribution

- **Comparison with existing approaches**

In an earlier contribution [36], we have investigated robust mode estimation under a functional single-index structure, employing the local constant method. Alternatively, in this study, we examine local linear estimation for the same model under a general functional framework. Firstly, observe that the topic of the present contribution can be viewed as a generalization of [36], in the sense that the local constant constitutes a particular case of the local linear approach, and the functional single index structure is a special case of the general functional structure. Moreover, it is well documented (see, for instance, [27]) that the local linear method has many advantages over the Nadaraya–Watson (the local constant method). Particularly, the local linear estimation is usually used to reduce the bias and the boundary effect of the Nadaraya–Watson method. So, the use of the local linear method instead of the standard kernel method substantially improves the prediction accuracy. On the other hand, it is widely recognized that the single-index model reduces analytical complexity by projecting functional covariates onto a univariate index. Thus, it transforms the functional data analysis problem into a one-dimensional real data issue. This oversimplification is unusual in practice, as it ignores potentially influential higher-dimensional interactions. Our generalized framework of this contribution circumvents this limitation. From a practical point of view, the kernel estimator assumes that the nonparametric model is flat within the neighborhood of the location point, which leads to suboptimal prediction. In contrast, the local linear method assumes that the model has a linear approximation in the neighborhood of the location point, which is more realistic and improves the prediction results.

- **On the bias reduction**

As mentioned in the previous comment, the behavior of the bias term is one of the main reasons for adopting the local linear method. Although the asymptotic behavior of the bias term is usually linked to the smoothness assumptions of the nonparametric model, this term can be significantly improved in the local linear approach. This beneficial characteristic is related to the weighting functions in the local linear estimator (see [27] in the non-functional case). A similar statement can be deduced in the functional case. Specifically, under standard conditions, the local linear approach offers a better bias term compared to the Nadaraya–Watson [36]. Indeed, this improvement is also due to the specific weighting functions implemented in $\widehat{QR}_q(\mathfrak{F})$. It is clear that the leading term of the Bahadur representation of $\widehat{QR}_q(\mathfrak{F})$ such that

$$\mathbb{E} \left[\mathbf{D}_i (A_3 \mathbf{F}_i - A_2 f_n^{-1} \mathbf{D}_i \mathbf{F}_i) \right] = 0$$

where

$$\mathbf{F}_i = \mathbf{F}(f_n^{-1} m(\mathfrak{F}, F_i)), \quad \text{and} \quad \mathbf{D}_i = m(F_i, \mathfrak{F})$$

and

$$A_j = \alpha_{\mathfrak{F}}^{-1}(f_n) f_n^{1-j} \mathbb{E}[\mathbf{D}_i^{j-1} \mathbf{F}_i]; \quad j = 2, 3.$$

Using the same analytical arguments as in [37], we prove that the first part of the bias term can be reduced to $o(f_n^{b/2})$. In conclusion, although the local linear and the Nadaraya–Watson (NW) estimators share similar asymptotic properties, the local linearity of the model and the weighting functions of the LLE method allow us to improve the bias term in certain situations.

- **On the applicability of the estimator**

Of course, the simple use of the estimator greatly depends on choosing its parameters easily. Since the estimator is derived from quantile regression, there are multiple cross-validation methods available for selecting the parameters, particularly the bandwidth parameters associated with the functional component. First, for a given subset H_n of real positive numbers, we consider the cross-validation criterion used by [38].

$$\arg \min_{f_n \in H_n} \frac{1}{n} \sum_{i=1}^n \left(Q_p^L(F_i) - \widehat{Q}_p^{-i}(F_i) \right)^2,$$

where $Q_p^L(X_i)$ denotes a local empirical quantile and $\widehat{Q}_p^{-i}$ is the conditional quantile estimator at F_i after excluding this observation from the sample. Secondly, we apply the cross-validation approach that was utilized by [39]

$$\arg \min_{f_n \in H_n} \frac{1}{n} \sum_{i=1}^n L_p \left(R_i - \widehat{Q}_p^{-i}(F_i) \right).$$

Additionally, various other methods can be utilized, such as the least squares cross-validation technique suggested by [6]:

$$\arg \min_{f_n \in H_n} \frac{1}{n} \sum_{i=1}^n \left[R_i - \widehat{MR}(F_i) \right]^2.$$

Of course, the diversity of selection methods makes the estimator easier to implement in practice.

5. Computational Part

5.1. Simulation Study

In this section, we demonstrate the practical effectiveness of our theoretical development through some simulation experiments. Our principal goal is to show the easy insertion of the estimator in practice and to emphasize clear benefits of the proposed estimator over its competitors, such as the kernel conditional mode and the conditional density-based LLE-mode. In order to cover many practical situations, we generate data from different scenarios. Indeed, we consider two different nonparametric functional models that are:

$$\text{heteroscedastic (Het.) Model : } R_i = 4 \int_0^1 \exp\left(\frac{1}{3 + F_i^3(t)}\right) dt + \cos(3 + F_i^3(t)) \epsilon_i,$$

$$\text{homoscedastic (Hom.) Model : } R_i = 5 \int_0^1 \log\left(\frac{2 + F_i^2(t)}{3 + F_i^3(t)}\right) dt + \epsilon_i,$$

where ϵ_i and F_i are independent. The functional input variable is also generated in two ways: smoothly and roughly. The smooth way is generated according to the following formula:

$$\text{Smooth curves } F_i(t) = a_i \cos(4(b_i - t\pi)) + b_i$$

while the rough way is drawn from the following equation:

$$\text{rough curves } F_i(t) = a_i \cos(4(b_i - t\pi)) + b_i + \eta_{i,t}$$

For both cases, b_i is distributed in $N(0, 0.7)$ and a_i is generated according to a $N(3, 4)$ distribution, while the random variables $\eta_{i,t}$ are drawn from $N(0, 2)$. All the input curves F_i 's are discretized on the same grid represented by 100 equispaced measurements taken in the interval $(0, 1)$. The resulting functional variables are shown in Figures 1 and 2.

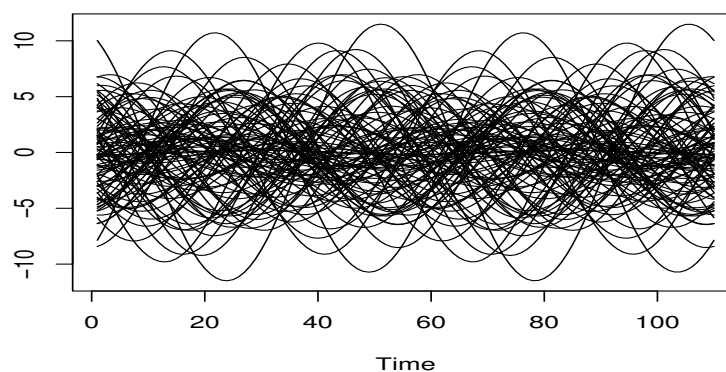


Figure 1. Smooth curves $F_i(t)$ for 100 equispaced measurements t in $(0, 1)$, as displayed on the X-axis.

Obviously, choosing various curve forms is essential for evaluating how the data's functional structure affects the accuracy of estimates. Typically, the spline metric is used for modeling smooth curves, while the PCA metric is used for rough curves.

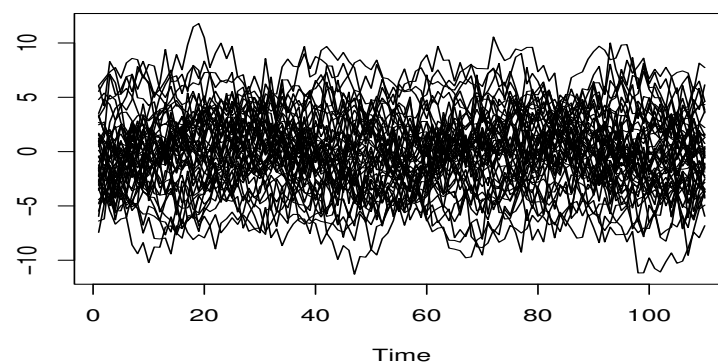


Figure 2. Rough curves $F_i(t)$ for 100 equispaced measurements t in $(0, 1)$, as displayed on the X-axis.

This strategy of data sampling is intentional. It allows us to test the robustness of different models and assess how the data functionality affects the estimation accuracy. We also evaluate the nonparametric model using three distributions: Normal, Weibull, and Laplace. These were chosen because they remain consistent under translation and exhibit different levels of heavy-tailed behavior, which helps us to measure how well the estimators handle outliers. In mathematical terms, the two competitors' modal regression estimators are defined by:

$$\text{The NW estimator: } \widetilde{MR}(\mathfrak{F}) = \arg \max_y \frac{\sum_{i=1}^n \mathbf{F}(f_n^{-1}m(\mathfrak{F}, F_i)) \mathbf{F}(b_n^{-1}(y - R_i))}{\sum_{i=1}^n b_n \mathbf{F}(f_n^{-1}m(\mathfrak{F}, F_i))}, \quad (12)$$

$$\text{and the LLE estimator: } \overline{MR}(\mathfrak{F}) = \arg \max_y \frac{\sum_{i,j=1}^n \mathbf{W}_{ij} \mathbf{F}(b_n^{-1}(y - R_j))}{\sum_{i,j=1}^n b_n \mathbf{W}_{ij}}, \quad (13)$$

where $W_{ij}(x) = m(F_i, \mathfrak{F})(m(F_i, \mathfrak{F}) - m(F_j, \mathfrak{F})) \mathbf{F}(f_n^{-1}m(\mathfrak{F}, F_i)) \mathbf{F}(f_n^{-1}m(\mathfrak{F}, F_j))$.

Of course, the behavior of the three estimators, $\widetilde{MR}$, $\widehat{MR}$, and $\overline{MR}$, as well as their straightforward insertion, is influenced by the choice of parameters, metric m , kernel $\mathbf{F}$, and smoothing parameters (f_n, b_n) . Evidently, the selection of the metric m is intrinsically linked to the smoothness of the functional curves. Smoother curves are better analyzed by

the B -spline metric, which accommodates their regularity, while rougher curves require the PCA-metric to handle their more irregular structure. We refer to [2] for more details on the definitions of these metrics. The choice of kernel is motivated by technical considerations. In order to incorporate the assumption (B3), we have simulated using the quadratic kernel in $(-1, 1)$. For the smoothing parameters (f_n, b_n) , we have used the cross-validation rule based on the mean square error defined by

$$MSE(\check{M}R) = \frac{1}{n} \sum_{i=1}^n (R_i - \check{M}R(F_i))^2, \quad (14)$$

where $\check{M}R$ represents either $\widehat{M}R$, $\overline{M}R$, or $\widetilde{M}R$. It follows that

$$(f^{opt}, b^{opt}) = \arg \min_{f_n \in \mathbf{H}_n, b_n \in \mathbf{B}_n} MSE(\check{M}R)$$

$\mathbf{H}_n$ (respectively, $\mathbf{B}_n$) is the set of positive real numbers, f_n , such that the ball centered at $\mathfrak{F}$ with radius f_n (respectively, the interval $[y - b_n, y + b_n]$) contains exactly k neighbors of (respectively, y). Typically, the number k is selected from the subset $\{5, 10, 20, \dots, 0.5 * n\}$. For our goal, we generate $n = 200$ observations of (F_i, R_i) , and we examine the robustness and the accuracy of different estimators using the MSE defined in (14). The obtained results are reported in the following table.

The performance of these estimators is strongly influenced by various components of the study, particularly the nature of the functional data and the nonparametric approach. Nevertheless, combining L^1 -techniques with the LLE-smoothing method proves highly advantageous in terms of robustness and accuracy. As evidenced by the MSE results in Table 1, the variability in $\widehat{M}R$ is lower than that of the other estimators. This lower variability demonstrates that the L^1 -LLE estimator for the conditional mode offers greater robustness and improved accuracy.

Table 1. MSE for different estimators in different scenarios .

Cond. Dist.	Model	CURVES	$\widehat{M}R$	$\overline{M}R$	$\widetilde{M}R$
Normal	Hom.model	Smooth	0.0258	0.0897	0.1672
		Rough	0.1067	0.4676	0.5679
Normal	Het.model	Smooth	0.2672	0.7426	1.0578
		Rough	0.3617	1.2725	1.8811
Laplace	Hom. model	Smooth	0.4257	0.6150	0.6317
		Rough	0.9804	1.0922	1.1788
Laplace	Het. model	Smooth	0.8967	1.6824	1.7088
		Rough	0.9176	2.4521	2.6588
Weibull	Hom. model	Smooth	0.5179	1.5005	1.5446
		Rough	0.8399	2.4873	2.7098
Weibull	Het. model	Smooth	0.7840	1.6253	1.4102
		Rough	0.9705	3.3567	4.3456

5.2. Real Data Application

Spectrometry data analysis is a powerful tool for generating high-frequency functional data. In this study, we focus on food quality as a critical real-world issue. Clearly, ensuring high-quality food is a top priority for both consumers and producers. Traditionally, quality control relies on chemical testing as an effective but destructive, time-consuming, and expensive approach. Instead, we provide a smart solution that combines spectrometry techniques with advanced statistical tools. It permits us to create an efficient, cost-effective,

and non-invasive approach to ensure food quality. Specifically, we demonstrate how our new predictor can accurately predict fat and moisture contents in chocolate. Therefore, we evaluate our approach against leading alternatives, including the NW estimator, LLE-conditional expectation, and density-based LLE-mode. This computational study allows us to demonstrate the superiority and practical advantages of $\widehat{MR}$ in robustness as well as in accuracy. We compare these three approaches by testing their predictive performance using the near-infrared spectroscopy curve as functional input F , with the fat (or moisture) quantity as response variable R . The near-infrared data set available on the website (<https://data.mendeley.com/datasets/7734j4fd98/1>, (accessed on 8 May 2025)). This issue has been considered by many authors using different algorithms (see [40] for a list of references). Formally, in this real data example, we assume that the spectrum curve gives the absorbance of light across wavelengths ranging from 1000 to 2500 nm. These curves are visualized in Figure 3.

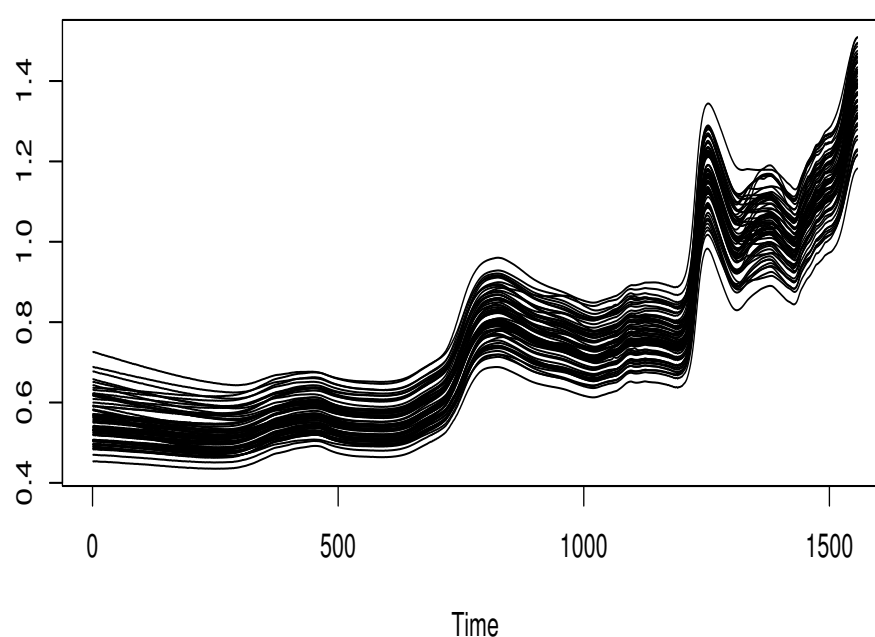


Figure 3. NIR curves.

These curves represent the absorbance spectrum (NIR) in the wavelength range from 1000 to 2500 nm for a total of 72 bulk intact cocoa bean samples. Each bulk amounted to 50 g of intact beans. Recall that the choice of the response variable R is motivated by the fact that the fat or moisture quantities are determinants of cocoa quality. The quality of cocoa beans is the cornerstone of exceptional chocolate production, profoundly influencing the sensory, nutritional, and commercial value of the final product. Thus, precise determination of cocoa contents is essential for compliance with chocolate quality standards, economic valuation, and research in agriculture and nutrition. Consequently, optimizing fat and moisture percentage is primordial to delivering high-quality chocolate products that meet both consumer and industry demands. Now, to put our predictors into action, we need to carefully select the principal parameters that drive our estimator's performance. This step is crucial to ensuring the potential feasibility of the four predictors $\widehat{MR}$, $\widetilde{MR}$, $\overline{MR}$, and

$$\overline{CE} = \frac{\sum_{i,j=1}^n \mathbf{w}_{ij} R_j}{\sum_{i,j=1}^n b_n \mathbf{w}_{ij}},$$

Based on our earlier discussion, the kernel selection is a direct consequence of assumption (B3), ensuring theoretical consistency. Meanwhile, the metric m plays a crucial role in the quantification of the smoothness of the functional curves F_i , allowing us to better analyze their behavior. Clearly, the curve's shape in Figure 3 shows that the metric of L_2 between the first derivatives of the curves (the curves being regular) is more suitable for this data. Once again, the smoothing parameters (f_n, b_n) according the same strategy as in (14). This latest approach has proven highly effective in our predictive modeling, highlighting the importance of leave-one-out cross-validation data-driven algorithms in functional data analysis. Figures 4 and 5 summarize our statistical predictions for fat and moisture for the 70 observations, comparing each approach's performance by plotting predicted values against the true values.

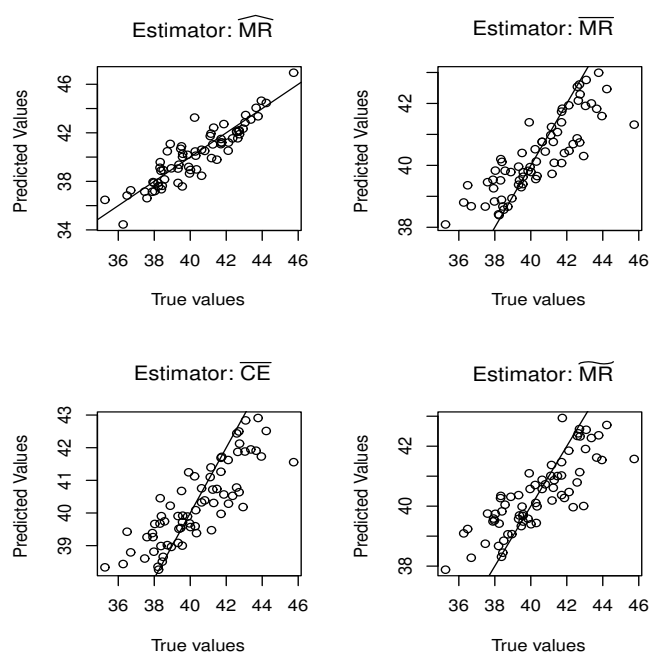


Figure 4. Moisture prediction.

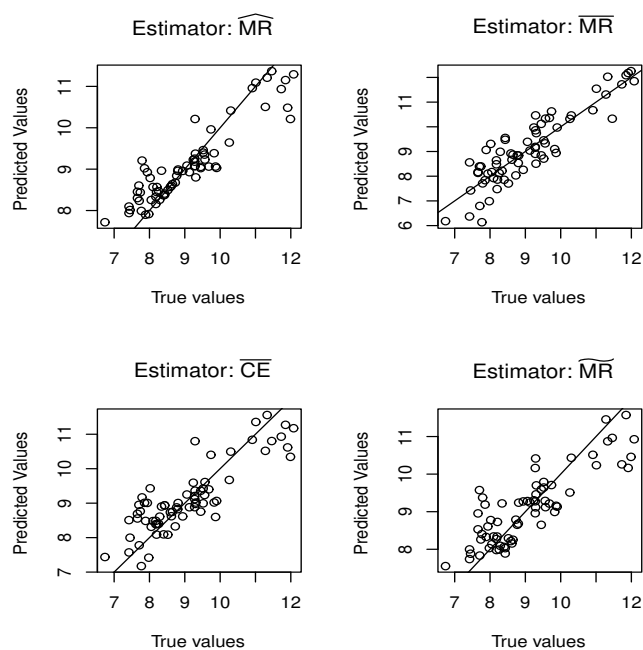


Figure 5. Fat prediction.

We use (MSE) error to assess model performance for the prediction results. It is defined by

$$MSE(\hat{\theta}) = \frac{1}{70} \sum_{i=1}^{70} (R_i - \hat{M}R(F_i))^2,$$

where $\hat{\theta}$ can represent either $\widehat{MR}$, $\overline{MR}$, $\widetilde{MR}$ or $\overline{CE}$. Clearly, the predictor $\widehat{MR}$ outperforms the other predictors in precision. The latter predicts (fat, moisture) with MSE equal, respectively, to (0.97, 1.06). It is the best one compared to the results obtained by the predictors $\overline{MR}$, $\widetilde{MR}$, and $\overline{CE}$. Typically, we record for both components (fat, moisture) (1.70, 2.89) for $\overline{MR}$, (4.89, 7.48) for $\widetilde{MR}$ and (3.45, 6.22) for $\overline{CE}$.

Author Contributions: The authors contributed approximately equally to this work. Formal analysis, F.A.A.; Validation, M.B.A.; Writing—review & editing, A.L. and Z.K. All authors have read and agreed to the published 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 for funding this work through Small Research Project under grant number RGP1/41/46.

Data Availability Statement: The data used in this study are available through the link <https://data.mendeley.com/datasets/7734j4fd98/1> (accessed on 8 May 2025).

Acknowledgments: The authors would like to thank the Editor, the Associate-Editor and two anonymous reviewers for their valuable comments and suggestions which improved substantially the quality of an earlier version of this paper. 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 for funding this work through Small Research Project under grant number RGP1/41/46.

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

Appendix A. Proofs of Intermediate Results

Proof of Proposition 1. We apply Lemma 1 on

$$v_n = \begin{pmatrix} \hat{A} - A \\ f_n(\hat{B} - B) \end{pmatrix}$$

$$\mathcal{X}_n(v) = \frac{1}{n\alpha_{\mathfrak{F}}(\tilde{F}_n)} \sum_{i=1}^n \zeta_q(v) \begin{pmatrix} 1 \\ f_n^{-1} \mathbf{D}_i \end{pmatrix} \mathbf{F}_i, \quad \text{for } v = \begin{pmatrix} c \\ d \end{pmatrix}$$

and

$$Y_n = \mathcal{X}_n(v_0) \quad \text{with } v_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

where

$$\zeta_q(v) = q - \mathbf{1}_{R_i \leq (c + A) + (f_n^{-1}d + B)\mathbf{D}_i}, \quad \mathbf{F}_i = \mathbf{F}(f_n^{-1}m(\mathfrak{F}, F_i)), \quad \text{and } \mathbf{D}_i = m(F_i, \mathfrak{F}).$$

So, the main result is a consequence of the following lemmas. $\square$

Lemma A1. Under assumptions (B1)–(B5), we have

$$\|Y_n\| = O(f_n^{\min(k_1, k_2)}) + O_{a.co.} \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{1/2} \right).$$

Proof of Lemma A1. Firstly, we set

$$\mathbf{Z}_i^1 = (q - \mathbf{1}_{[R_i \leq A + B\mathbf{D}_i]}) \mathbf{F}_i - \mathbb{E} \left[(q - \mathbf{1}_{[R_i \leq A + B\mathbf{D}_i]}) \mathbf{F}_i \right]$$

and

$$\mathbf{Z}_i^2 = (q - \mathbf{1}_{[R_i \leq A + B\mathbf{D}_i]}) \mathbf{D}_i \mathbf{F}_i - \mathbb{E} \left[(q - \mathbf{1}_{[R_i \leq A + B\mathbf{D}_i]}) \mathbf{F}_i \right].$$

Then

$$Y_n - \mathbb{E}[Y_n] = \begin{pmatrix} Y_n^1 \\ Y_n^2 \end{pmatrix}$$

where

$$\begin{cases} Y_n^1 = \frac{1}{n \alpha_{\mathfrak{F}}(\tilde{F}_n)} \sum_{i=1}^n \mathbf{Z}_i^1 \\ Y_n^2 = \frac{1}{n h \alpha_{\mathfrak{F}}(\tilde{F}_n)} \sum_{i=1}^n \mathbf{Z}_i^2. \end{cases}$$

Obviously

$$|\mathbf{Z}_i^1| \leq C \quad \text{and} \quad |\mathbf{Z}_i^2| \leq C' h.$$

Moreover,

$$\mathbb{E}[\mathbf{Z}_i^1]^2 \leq C \alpha_{\mathfrak{F}}(\tilde{F}_n) \quad \text{and} \quad \mathbb{E}[\mathbf{Z}_i^2]^2 \leq C' f_n^2 \alpha_{\mathfrak{F}}(\tilde{F}_n).$$

It follows that

$$Y_n^1 - \mathbb{E}[Y_n^1] = O_{a.co.} \left(\sqrt{\frac{\log n}{n \alpha_{\mathfrak{F}}(\tilde{F}_n)}} \right)$$

and

$$Y_n^2 - \mathbb{E}[Y_n^2] = O_{a.co.} \left(\sqrt{\frac{\log n}{n \alpha_{\mathfrak{F}}(\tilde{F}_n)}} \right).$$

Next,

$$\begin{aligned} \mathbb{E}[Y_n^1] &= \frac{1}{\alpha_{\mathfrak{F}}(\tilde{F}_n)} \mathbb{E} \left[(q - \mathbf{1}_{[Y_1 \leq A + B\mathbf{D}_1]}) \mathbf{F}_1 \right] \\ &\leq \frac{1}{\alpha_{\mathfrak{F}}(f_n)} \mathbb{E} |D(QR_q(\mathfrak{F})|\mathfrak{F}) - D((A + B\mathbf{D}_1)\mathbf{F}_1|F)| \\ &= O(f_n^{\min(k_1, k_2)}). \end{aligned}$$

Similarly,

$$\begin{aligned} \mathbb{E}[Y_n^2] &= \frac{1}{h \alpha_{\mathfrak{F}}(\tilde{F}_n)} \mathbb{E} \left[(q - \mathbf{1}_{[Y_1 \leq A + B\mathbf{D}_1]}) \mathbf{D}_1 \mathbf{F}_1 \right] \\ &\leq \frac{1}{h \alpha_{\mathfrak{F}}(f_n)} \mathbb{E} |D(QR_q(\mathfrak{F})|\mathfrak{F}) - D((A + B\mathbf{D}_1)\mathbf{D}_1 \mathbf{F}_1|F)| \\ &= O(f_n^{\min(k_1, k_2)}). \end{aligned}$$

Therefore,

$$\|Y_n\| = O(f_n^{\min(k_1, k_2)}) + O_{a.co.} \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{1/2} \right)$$

□

Lemma A2. Under assumptions (B1)–(B5), we have

$$\sup_{\|v\| \leq M} \|\mathcal{X}_n(v) + \lambda_0 Dv - Y_n\| = o_{a.co.}(1)$$

with

$$D = \begin{pmatrix} \mathbf{F}(1) - \int_{-1}^1 K'(t) \beta_{\mathfrak{F}}(t) dt & \mathbf{F}(1) - \int_{-1}^1 (t\mathbf{F}(t))' \beta_{\mathfrak{F}}(t) dt \\ \mathbf{F}(1) - \int_{-1}^1 (t\mathbf{F}(t))' \beta_{\mathfrak{F}}(t) dt & \mathbf{F}(1) - \int_{-1}^1 (t^2\mathbf{F}(t))' \beta_{\mathfrak{F}}(t) dt \end{pmatrix}$$

and

$$\lambda_0 = d(QR_q(\mathfrak{F})|\mathfrak{F})$$

Proof of Lemma A2. We prove

$$\sup_{\|v\| \leq M} \|\mathcal{X}_n(v) - Y_n - \mathbb{E}[\mathcal{X}_n(v) - Y_n]\| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\alpha_{\mathfrak{F}}(\tilde{F}_n)}}\right) \quad (\text{A1})$$

and

$$\sup_{\|v\| \leq M} \|\mathbb{E}[\mathcal{X}_n(v) - Y_n] + d(QR_q(\mathfrak{F})|\mathfrak{F})Dv\| = O(f_n^{\min(k_1, k_2)}). \quad (\text{A2})$$

Next, use the compactness of the ball $B(0, M)$ in $\mathbb{R}^2$ and we write

$$B(0, M) \subset \bigcup_{j=1}^{d_n} B(v_j, l_n), \quad v_j = \begin{pmatrix} c_j \\ d_j \end{pmatrix} \text{ and } l_n = d_n^{-1} = 1/\sqrt{n}.$$

Taking $j(v) = \arg \min_j |v - v_j|$ and we use the fact that

$$\begin{aligned} & \sup_{\|v\| \leq M} \|\mathcal{X}_n(v) - Y_n - \mathbb{E}[\mathcal{X}_n(v) - Y_n]\| \\ & \leq \sup_{\|v\| \leq M} \|\mathcal{X}_n(v) - \mathcal{X}_n(v_j)\| \\ & \quad + \sup_{\|v\| \leq M} \|\mathcal{X}_n(v_j) - Y_n - \mathbb{E}[\mathcal{X}_n(v_j) - Y_n]\| \\ & \quad + \sup_{\|v\| \leq M} \|\mathbb{E}[\mathcal{X}_n(v) - \mathcal{X}_n(v_j)]\|. \end{aligned}$$

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

$$\sup_{\|v\| \leq M} \|\mathcal{X}_n(v) - \mathcal{X}_n(v_j)\| \leq \frac{1}{n\alpha_{\mathfrak{F}}(\tilde{F}_n)} \sum_i Z_i$$

where

$$\Omega_i = \sup_{\|v\| \leq M} \mathbf{1}_{[|R_i - (c_j + A) - (f_n^{-1}d_j + d)\mathbf{D}_i| \leq Cl_n]} \left\| \begin{pmatrix} 1 \\ f_n^{-1}\mathbf{D}_i \end{pmatrix} \right\| \mathbf{F}_i.$$

Clearly

$$|\Omega_i| \leq C, \quad \mathbb{E}[\Omega_i] = O(l_n \alpha_{\mathfrak{F}}(\tilde{F}_n)) \quad \text{and} \quad \mathbb{E}[\Omega_i^2] = O(l_n \alpha_{\mathfrak{F}}(\tilde{F}_n)).$$

By using the fact that

$$l_n = o\left(\left(\frac{\log n}{n\alpha_{\mathfrak{F}}(\tilde{F}_n)}\right)^{1/2}\right) \quad (\text{A3})$$

we get

$$\sup_{\|v\| \leq M} \|\mathcal{X}_n(v) - \mathcal{X}_n(v_j)\| = O_{a.co.} \left(\left(\frac{\log n}{n\alpha_{\mathfrak{F}}(\tilde{F}_n)} \right)^{1/2} \right).$$

For the last term,

$$\sup_{\|v\| \leq M} \|\mathbb{E}[\mathcal{X}_n(v) - \mathcal{X}_n(v_j)]\| \leq \frac{1}{\alpha_{\mathfrak{F}}(\tilde{F}_n)} \mathbb{E}[\Omega_1] \leq Cl_n$$

then, by (A3) we obtain

$$\sup_{\|v\| \leq M} \|\mathbb{E}[\mathcal{X}_n(v) - \mathcal{X}_n(v_j)]\| = o_{a.co.} \left(\left(\frac{\log n}{n\alpha_{\mathfrak{F}}(\tilde{F}_n)} \right)^{1/2} \right).$$

Now

$$\sup_{\|v\| \leq M} \|\mathcal{X}_n(v_j) - Y_n - \mathbb{E}[\mathcal{X}_n(v_j) - Y_n]\|.$$

For this

$$\mathcal{X}_n(v_j) - Y_n - \mathbb{E}[\mathcal{X}_n(v_j) - Y_n] = \begin{pmatrix} \Theta_n^1(v_j) \\ \Theta_n^2(v_j) \end{pmatrix}$$

where

$$\begin{cases} \Theta_n^1(v_j) = \frac{1}{n\alpha_{\mathfrak{F}}(f_n)} \sum_{i=1}^n \Lambda_i^1 \\ \Theta_n^2(v_j) = \frac{1}{nf_n\alpha_{\mathfrak{F}}(f_n)} \sum_{i=1}^n \Lambda_i^2 \end{cases}$$

with

$$\Lambda_i^1 = (\zeta_q(v_j) - \zeta_q(v_0)) \mathbf{F}_i - \mathbb{E}[(\zeta_q(v_j) - \zeta_q(v_0)) \mathbf{F}_i]$$

and

$$\Lambda_i^2 = (\zeta_q(v_j) - \zeta_q(v_0)) \mathbf{D}_i \mathbf{F}_i - \mathbb{E}[(\zeta_q(v_j) - \zeta_q(v_0)) \mathbf{D}_i \mathbf{F}_i].$$

clearly

$$|\Lambda_i^1| \leq C \quad \text{and} \quad |\Lambda_i^2| \leq C' f_n.$$

Moreover,

$$\mathbb{E}[\Lambda_i^1]^2 \leq C\alpha_{\mathfrak{F}}(f_n) \quad \text{and} \quad \mathbb{E}[\Lambda_i^2]^2 \leq C' f_n^2 \alpha_{\mathfrak{F}}(f_n).$$

Therefore, there exists $\eta > 0$ such that

$$\begin{aligned} \sum_n \mathbb{P} \left(\sup_{\|v\| \leq M} \|\mathcal{X}_n(v_j) - Y_n - \mathbb{E}[\mathcal{X}_n(v_j) - Y_n]\| \geq \eta \sqrt{\frac{\log n}{n\alpha_{\mathfrak{F}}(f_n)}} \right) \\ \leq \sum_n d_n \max_j \mathbb{P} \left(\|\mathcal{X}_n(v_j) - Y_n - \mathbb{E}[\mathcal{X}_n(v_j) - Y_n]\| \geq \eta \sqrt{\frac{\log n}{n\alpha_{\mathfrak{F}}(f_n)}} \right) < \infty \end{aligned}$$

which complete the proof of (A1).

Concerning (A2) we write

$$\mathcal{X}_n(v) - Y_n = \begin{pmatrix} \Theta_n^1(v) \\ \Theta_n^2(v) \end{pmatrix}$$

where

$$\Theta_n^1(v) = \frac{1}{n\alpha_{\mathfrak{F}}(f_n)} \sum_{i=1}^n (\zeta_q(v) - \zeta_q(v_0)) \mathbf{F}_i$$

and

$$\Theta_n^2(\nu) = \frac{1}{nh\alpha_{\mathfrak{F}}(f_n)} \sum_{i=1}^n (\zeta_q(\nu) - \zeta_q(\nu_0)) \mathbf{D}_i \mathbf{F}_i$$

On the other hand, we have

$$\begin{aligned} \mathbb{E}[\Theta_n^1(\nu)] &= -\frac{1}{\alpha_{\mathfrak{F}}(f_n)} \mathbb{E}\left[\left(\mathbf{1}_{[Y_1 \leq (c+A) + (f_n^{-1}d+B)\mathbf{D}_1]} - \mathbf{1}_{[Y_1 \leq A+B\mathbf{D}_1]}\right) \mathbf{F}_1\right] \\ &= -\frac{1}{\alpha_{\mathfrak{F}}(f_n)} \mathbb{E}\left[\left(D((c+A) + (f_n^{-1}d+B)\mathbf{D}_1) \mathbf{F}_1 - D((A+B\mathbf{D}_1) \mathbf{F}_1)\right) \mathbf{F}_1\right] \\ &= -\frac{1}{\alpha_{\mathfrak{F}}(f_n)} \mathbb{E}\left[\left(D((c+A) + (f_n^{-1}d+B)\mathbf{D}_1) \mathbf{F}_1 - D(A+B\mathbf{D}_1) \mathbf{F}_1\right) \mathbf{F}_1\right] + O(f_n^{k_1}) \\ &= -\frac{1}{\alpha_{\mathfrak{F}}(f_n)} \mathbb{E}\left[d(A+B\mathbf{D}_1) \mathbf{F}_1 \mathbf{F}_1' + O(f_n^{k_1}) + o(\|\nu\|)\right] \\ &= -d(QR_q(\mathfrak{F})|\mathfrak{F}) \frac{1}{\alpha_{\mathfrak{F}}(f_n)} \left(\mathbb{E}\mathbf{F}_1, f_n^{-1}\mathbb{E}[\mathbf{D}_1 \mathbf{F}_1]\right) \nu + O(f_n^{\min(k_1, k_2)}) + o(\|\nu\|) \end{aligned}$$

Similarly,

$$\mathbb{E}[\Theta_n^2(\nu)] = -d(QR_q(\mathfrak{F})|\mathfrak{F}) \frac{1}{h\alpha_{\mathfrak{F}}(f_n)} \left(\mathbb{E}\mathbf{D}_1 \mathbf{F}_1, f_n^{-1}\mathbb{E}[\mathbf{D}_1^2 \mathbf{F}_1]\right) \nu + O(f_n^{\min(k_1, k_2)}) + o(\|\nu\|).$$

Thus

$$\begin{aligned} \mathbb{E}[\mathcal{X}_n(\nu) - Y_n] &= -d(QR_q(\mathfrak{F})|\mathfrak{F}) \frac{1}{\alpha_{\mathfrak{F}}(f_n)} \begin{pmatrix} \mathbb{E}[\mathbf{F}_i] & \mathbb{E}[\mathbf{F}_i f_n^{-1} \mathbf{D}_i] \\ \mathbb{E}[\mathbf{F}_i f_n^{-1} \mathbf{D}_i] & \mathbb{E}[f_n^{-2} \mathbf{D}_i^2 \mathbf{F}_i] \end{pmatrix} \nu \\ &\quad + O(f_n^{\min(k_1, k_2)}) + o(\|\nu\|). \end{aligned}$$

Obviously

$$f_n^{-a} \mathbb{E}[\mathbf{D}_i^{-a} \mathbf{F}_i^b] = \alpha_{\mathfrak{F}}(f_n) \left(\mathbf{F}^b(1) - \int_{-1}^1 (u^a \mathbf{F}^b(u))' \mathbf{D}(u) du \right) + o(\alpha_{\mathfrak{F}}(f_n)).$$

Hence,

$$\sup_{\|\nu\| \leq M} \|\mathbb{E}[\mathcal{X}_n(\nu) - Y_n] + d(QR_q(\mathfrak{F})|\mathfrak{F}) D\nu + o(\|\nu\|)\| = O(f_n^{\min(k_1, k_2)})$$

implying the result (A2). $\square$

Lemma A3. Under assumptions (B1)–(B5), we have

$$\sup_{|\nu| \leq M} \sup_{q \in [0,1]} |\Pi_n(\nu, q) - \mathbb{E}[\Pi_n(\nu, q)]| = O_{a.co.} \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{1/2} \right).$$

where

$$\Pi_n(\nu, q) = \frac{1}{n\alpha_{\mathfrak{F}}(\tilde{F}_n)} \sum_{i=1}^n \zeta_q(\nu) \begin{pmatrix} 1 \\ f_n^{-1} \mathbf{D}_i \end{pmatrix} \mathbf{F}_i, \quad \text{for } \nu = \begin{pmatrix} c \\ d \end{pmatrix}$$

Proof of Lemma A3. The compactness of $[0, 1]$ implies

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

Next, for all $q \in [0, 1]$ we put $F_q = \arg \min_k |q - q_k|$ and we evaluate the term as function of v and q . We have,

$$\begin{aligned} \sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v, q) - \mathbb{E}[\Pi_n(v, q)]| &\leq \sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v, q) - \Pi_n(v_{j(v)}, q)| \\ &+ \sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v_{j(v)}, q) - \Pi_n(v_{j(v)}, q_{F_q})| \\ &+ \sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v_{j(v)}, q_{F_q}) - \mathbb{E}[\Pi_n(v_{j(v)}, q_{F_q})]| \\ &+ \sup_{|v| \leq M} \sup_{q \in [0,1]} |\mathbb{E}[\Pi_n(v_{j(v)}, q_{F_q})] - \mathbb{E}[\Pi_n(v, q_{F_q})]| \\ &+ \sup_{|v| \leq M} \sup_{q \in [0,1]} |\mathbb{E}[\Pi_n(v, q_{F_q})] - \mathbb{E}[\Pi_n(v, q)]|. \end{aligned}$$

Firstly, we write

$$\sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v, q) - \Pi_n(v_{j(v)}, q)| \leq \frac{1}{n\mathbb{E}[\mathbf{F}_1]} \sum_i \Omega_i^0$$

with

$$\Omega_i^0 = \sup_{|v| \leq M} \sup_{q \in [0,1]} \mathbf{1}_{[|R_i - v_{j(v)} - QR(q|\mathfrak{F})| \leq Cl_n]} \mathbf{F}_i.$$

As

$$|\Omega_i^0| \leq C, \quad \mathbb{E}[\Omega_i^0] = O(l_n \alpha_{\mathfrak{F}}(f_n)) \quad \text{and} \quad \mathbb{E}[\Omega_i^{02}] = O(l_n \alpha_{\mathfrak{F}}(f_n))$$

we get

$$\sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v, q) - \Pi_n(v_{j(v)}, q)| = O_{a.co.} \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{1/2} \right)$$

and

$$\sup_{|v| \leq M} \sup_{q \in [0,1]} |\mathbb{E}[\Pi_n(v, q) - \Pi_n(v_{j(v)}, q)]| = o \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{1/2} \right).$$

Similarly,

$$\sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v_{j(v)}, q) - \Pi_n(v_{j(v)}, q_{F_q})| \leq \frac{1}{n\mathbb{E}[\mathbf{F}_1]} \sum_i \Omega_i^1$$

with

$$\Omega_i^1 = \sup_{|v| \leq M} \sup_{q \in [0,1]} \left(\mathbf{1}_{[|R_i - v_{j(v)} - t(q|\mathfrak{F})| \leq Cl_n]} + \mathbf{1}_{[|R_i - t(q_{F_q}|\mathfrak{F})| \leq Cl_n]} \right) \mathbf{F}_i.$$

Since

$$|\Omega_i^1| \leq C, \quad \mathbb{E}[\Omega_i^1] = O(l_n \alpha_{\mathfrak{F}}(f_n)) \quad \text{and} \quad \mathbb{E}[\Omega_i^{01}] = O(l_n \alpha_{\mathfrak{F}}(f_n))$$

we obtain

$$\sup_{|v| \leq M} \sup_{q \in [0,1]} |\Pi_n(v, q) - \Pi_n(v_{j(v)}, q)| = O_{a.co.} \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{1/2} \right)$$

and

$$\sup_{|v| \leq M} \sup_{q \in [0,1]} |\mathbb{E}[\Pi_n(v, q) - \Pi_n(v_{j(v)}, q)]| = o \left(\left(\frac{\log n}{n \alpha_{\mathfrak{F}}(f_n)} \right)^{1/2} \right).$$

By the same manner

$$\Pi_n(\nu_j, q_k) - \mathbb{E}[\Pi_n(\nu_j, q_k)] = \frac{1}{n\mathbb{E}[\mathbf{F}_1]} \sum_{i=1}^n \Lambda'_i$$

where

$$\Lambda'_i = \left(\mathbf{1}_{R_i \leq t(q_k|\mathfrak{F})} - \mathbf{1}_{R_i \leq \nu_j + t(q_k|\mathfrak{F})} \right) \mathbf{F}_i - \mathbb{E} \left[\left(\mathbf{1}_{R_i \leq t(q_k|\mathfrak{F})} - \mathbf{1}_{R_i \leq \nu_j + t(q_k|\mathfrak{F})} \right) \mathbf{F}_i \right].$$

By Bernstein inequality

$$|\Lambda_i| \leq C \quad \text{and} \quad \mathbb{E}[\Lambda_i]^2 \leq C\alpha_{\mathfrak{F}}(f_n)$$

to show that there exists $\eta > 0$ such that

$$\begin{aligned} \sum_n \mathbb{P} \left(\sup_{|\nu| \leq M} \sup_{q \in [0,1]} |\Pi_n(\nu_j(\nu), q_{F_q}) - \mathbb{E}[\Pi_n(\nu_j(\nu), q_{F_q})]| \geq \eta \sqrt{\frac{\log n}{n\alpha_{\mathfrak{F}}(f_n)}} \right) \\ \leq \sum_n d_n^2 \max_j \max_k \mathbb{P} \left(|\Pi_n(\nu_j, q_k) - \mathbb{E}[\Pi_n(\nu_j, q_k)]| \geq \eta \sqrt{\frac{\log n}{n\alpha_{\mathfrak{F}}(f_n)}} \right) < \infty. \end{aligned}$$

Therefore,

$$\sup_{q \in [0,1]} \sup_{|\nu| \leq M} |\Pi_n(\nu, q) - Y_n - \mathbb{E}[\Pi_n(\nu, q) - Y_n]| = O_{a.co.} \left(\sqrt{\frac{\log n}{n\alpha_{\mathfrak{F}}(f_n)}} \right)$$

Observe that by the same techniques we prove that

$$\sup_{q \in [0,1]} |Y_n - \mathbb{E}[Y_n]| = O_{a.co.} \left(\sqrt{\frac{\log n}{n\alpha_{\mathfrak{F}}(f_n)}} \right)$$

□

References

1. Ramsay, J.O.; Silverman, B.W. *Functional Data Analysis*, 2nd ed.; Springer: New York, NY, USA, 2005.
2. Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis. Theory and Practice*; Springer Series in Statistics; Springer: New York, NY, USA, 2006.
3. Xu, Y. Functional Data Analysis. In *Springer Handbook of Engineering Statistics*; Springer: London, UK, 2023; pp. 67–85.
4. Ahmed, M.S.; Frévent, C.; Génin, M. Spatial Scan Statistics for Functional Data. In *Handbook of Scan Statistics*; Springer: New York, NY, USA, 2024; pp. 629–645.
5. Collomb, G.; Härdle, W.; Hassani, S. A note on prediction via estimation of the conditional mode function. *J. Stat. Plan. Inference* **1986**, *15*, 227–236. [\[CrossRef\]](#)
6. Quintela-Del-Rio, A.; Vieu, P. A nonparametric conditional mode estimate. *J. Nonparametr. Stat.* **1997**, *8*, 253–266. [\[CrossRef\]](#)
7. Ioannides, D.; Matzner-Løber, E. A note on asymptotic normality of convergent estimates of the conditional mode with errors-in-variables. *J. Nonparametr. Stat.* **2004**, *16*, 515–524. [\[CrossRef\]](#)
8. Louani, D.; Ould-Saïd, E.L.I.A.S. Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametr. Stat.* **1999**, *11*, 413–442. [\[CrossRef\]](#)
9. Bouzebda, S.; Khadani, S.; Slaoui, Y. Asymptotic normality of the regression mode in the nonparametric random design model for censored data. *Commun. Stat. Theory Methods* **2023**, *52*, 7069–7093. [\[CrossRef\]](#)
10. Bouzebda, S.; Didi, S. Some results about kernel estimators for function derivatives based on stationary and ergodic continuous time processes with applications. *Communications in Statistics. Theory Methods* **2022**, *51*, 3886–3933. [\[CrossRef\]](#)

11. Ferraty, F.; Laksaci, A.; Vieu, P. Estimating some characteristics of the conditional distribution in nonparametric functional models. *Stat. Inference Stoch. Process.* **2006**, *9*, 47–76. [\[CrossRef\]](#)
12. Bouzebda, S.; Chaouch, M.; Laïb, N. Limiting law results for a class of conditional mode estimates for functional stationary ergodic data. *Math. Methods Stat.* **2016**, *25*, 1066–1307. [\[CrossRef\]](#)
13. Ezzahrioui, M.H.; Ould-Saïd, E. Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametr. Stat.* **2008**, *20*, 3–18. [\[CrossRef\]](#)
14. Ezzahrioui, M.H.; Saïd, E.O. Some asymptotic results of a non-parametric conditional mode estimator for functional time-series data. *Stat. Neerl.* **2010**, *64*, 171–201. [\[CrossRef\]](#)
15. 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. [\[CrossRef\]](#)
16. Bouanani, O.; Laksaci, A.; Rachdi, M.; Rahmani, S. Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika* **2019**, *46*, 199–233. [\[CrossRef\]](#)
17. Azzi, A.; Belguerna, A.; Laksaci, A.; Rachdi, M. The scalar-on-function modal regression for functional time series data. *J. Nonparametr. Stat.* **2024**, *36*, 503–526. [\[CrossRef\]](#)
18. Stone, C.J. Consistent nonparametric regression. *Ann. Stat.* **1977**, *5*, 595–620. [\[CrossRef\]](#)
19. Koenker, R.; Zhao, Q. Conditional quantile estimation and inference for ARCH models. *Econom. Theory* **1996**, *12*, 793–813. [\[CrossRef\]](#)
20. Hallin, M.; Lu, Z.; Yu, K. Local linear spatial quantile regression. *Bernoulli* **2009**, *15*, 659–686. [\[CrossRef\]](#)
21. Cardot, H.; Crambes, C.; Sarda, P. Quantile regression when the covariates are functions. *J. Nonparametr. Stat.* **2005**, *17*, 841–856. [\[CrossRef\]](#)
22. Wang, H.; Ma, Y. Optimal subsampling for quantile regression in big data. *Biometrika* **2021**, *108*, 99–112. [\[CrossRef\]](#)
23. Jiang, Z.; Huang, Z. Single-index partially functional linear quantile regression. *Commun. Stat.-Theory Methods* **2024**, *53*, 1838–1850. [\[CrossRef\]](#)
24. Dabo-Niang, S.; Kaid, Z.; Laksaci, A. Spatial conditional quantile regression: Weak consistency of a kernel estimate. *Rev. Roum. Math. Pures Appl.* **2012**, *57*, 311–339.
25. Chowdhury, J.; Chaudhuri, P. Nonparametric depth and quantile regression for functional data. *Bernoulli* **2019**, *25*, 395–423. [\[CrossRef\]](#)
26. Mutis, M.; Beyaztas, U.; Karaman, F.; Lin Shang, H. On function-on-function linear quantile regression. *J. Appl. Stat.* **2025**, *52*, 814–840. [\[CrossRef\]](#)
27. Fan, J. *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability* 66; Routledge: Abingdon-on-Thames, UK, **2018**.
28. Rachdi, M.; Laksaci, A.; Demongeot, J.; Abdali, A.; Madani, F. Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Comput. Stat. Data Anal.* **2014**, *73*, 53–68. [\[CrossRef\]](#)
29. Baïllo, A.; Grané, A. Local linear regression for functional predictor and scalar response. *J. Multivar. Anal.* **2009**, *100*, 102–111. [\[CrossRef\]](#)
30. Barrientos-Marin, J.; Ferraty, F.; Vieu, P. Locally modelled regression and functional data. *J. Nonparametr. Stat.* **2010**, *22*, 617–632. [\[CrossRef\]](#)
31. Berline, A.; Elamine, A.; Mas, A. Local linear regression for functional data. *Ann. Inst. Stat. Math.* **2011**, *63*, 1047–1075. [\[CrossRef\]](#)
32. Demongeot, J.; Laksaci, A.; Madani, F.; Rachdi, M. Functional data: Local linear estimation of the conditional density and its application. *Statistics* **2013**, *47*, 26–44. [\[CrossRef\]](#)
33. Messaci, F.; Nemouchi, N.; Ouassou, I.; Rachdi, M. Local polynomial modelling of the conditional quantile for functional data. *Stat. Methods Appl.* **2015**, *24*, 597–622. [\[CrossRef\]](#)
34. Laksaci, A.; Ould Saïd, E.; Rachdi, M. Uniform consistency in number of neighbors of the k NN estimator of the conditional quantile model. *Metrika* **2021**, *84*, 895–911. [\[CrossRef\]](#)
35. Ota, H.; Kato, K.; Hara, S. Quantile regression approach to conditional mode estimation. *Electron. J. Stat.* **2019**, *13*, 3120–3160. [\[CrossRef\]](#)
36. Almulhim, F.A.; Alamari, M.B.; Bouzebda, S.; Kaid, Z.; Laksaci, A. Robust Estimation of L 1-Modal Regression Under Functional Single-Index Models for Practical Applications. *Mathematics* **2025**, *13*, 602. [\[CrossRef\]](#)
37. Belarbi, F.; Chemikh, S.; Laksaci, A. Local linear estimate of the nonparametric robust regression in functional data. *Stat. Probab. Lett.* **2018**, *134*, 128–133. [\[CrossRef\]](#)
38. Iglesias Pérez, M.D.C. Estimación de la función de distribución condicional en presencia de censura y truncamiento: Una aplicación al estudio de la mortalidad en pacientes diabéticos. *Estad. Esp.* **2003**, *45*, 275–302.

39. Yuan, M. GACV for quantile smoothing splines. *Comput. Stat. Data Anal.* **2006**, *50*, 813–829. [[CrossRef](#)]
40. Panchbhai, K.G.; Lanjewar, M.G. Portable system for cocoa bean quality assessment using multi-output learning and augmentation. *Food Control* **2025**, *174*, 111234. [[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.