Nonparametric Functional Least Absolute Relative Error Regression: Application to Econophysics

Ali Laksaci , Ibrahim M. Almanjahi , Mustapha Rachdi *

Posted Date: 10 December 2025

doi: 10.20944/preprints202512.0947.v1

Keywords: functional data; nonparametric regression; stochastic consistency; least absolute deviation; relative error; kernel method; Bandwidth parameter

 doi:10.20944/preprints202512.0947.v1

Article

Nonparametric Functional Least Absolute Relative Error Regression: Application to Econophysics

Ali Laksaci [1], Ibrahim M. Almanjahi [1] and Mustapha Rachdi [2,*]

[1] Department of Mathematics, College of Science, King Khalid University, Abha, 62223, Saudi Arabia
[2] Université Grenoble Alpes (France), Laboratoire AGEIS, EA 7407, AGIM Team, UFR SHS, BP. 47, F38040 Grenoble Cedex 09, France
* Correspondence: mustapha.rachdi@univ-grenoble-alpes.fr; Tel.: +33-4-76-82-58-53
† Current address: Université Grenoble Alpes (France), Laboratoire AGEIS, EA 7407, AGIM Team, UFR SHS, BP. 47, 38040 Grenoble Cedex 09, France.
‡ These authors contributed equally to this work.

Abstract

In this paper, we propose an alternative kernel estimator for the regression operator of a scalar response variable S given a functional random variable $\mathcal{T}$ that takes values in a semi-metric space. The new estimator is constructed through the minimization of the least absolute relative error (LARE). The latter is characterized by its ability to provide a more balanced and scale-invariant measure of prediction accuracy compared to traditional standard absolute or squared error criterion. The LARE is an appropriate tool for reducing the influence of extremely large or small response values, enhancing robustness against heteroscedasticity or/and outliers. This feature makes LARE suitable for functional or high-dimensional data, where variations in scale are common. The high feasibility and strong performance of the proposed estimator is theoretically supported by establishing its stochastic consistency. The latter is derived with precision of the converge rate under mild regularity conditions. The ease implementation and the stability of the estimator are justified by simulation studies and an empirical application to near-infrared (NIR) spectrometry data. Of course the to explore the functional architecture of this data, we employ random matrix theory (RMT) which is a principal analytical tool of econophysics.

Keywords: functional data; nonparametric regression; stochastic consistency; least absolute deviation; relative error; kernel method; Bandwidth parameter

1. Introduction

Examining how a functional predictor co-varies with a scalar response variable is a fundamental issue in functional statistics. This relationship is commonly evaluated through regression models, which is estimated using the least squares loss function. In this paper, we adopt another smoothing approach based on the Least Absolute Relative Error (LARE). Undoubtedly, the LARE criterion improves the robustness and the accuracy of the standard regression approaches.

The general framework of this paper is the nonparametric prediction in functional statistics, which has attracted considerable interest due to numerous valuable contributions in recent years. The foundational work in this field was introduced by [1], who established the almost complete pointwise consistency of the kernel-type regression estimator under the assumption of independent and identically distributed (i.i.d.) observations. [2] proved the convergence of this estimator in the L^p norm. [3] derived the exact asymptotic expression for the corresponding L^p errors. The asymptotic normality of the same estimator in the strong mixing case was later demonstrated by [4]. In [5], the authors established a uniform version of the almost complete convergence rate in the i.i.d. framework. In addition to these classical kernel-based approaches, recent research has explored M-estimation

 doi:10.20944/preprints202512.0947.v1

techniques and/or local linear smoothing methods. Significant results in this field of nonparametric functional prediction were presented by [6–9]., which provide comprehensive references on this topics.

Alternatively, in this paper, we focus on relative error regression (RE-regression), a methodology in which the the relative error is used. The latter constitutes a performance measure in many practical applications. Despite its utility, the literature on nonparametric RE-regression remains limited. The most existing studies adopting parametric approaches. In this context, [9] were among the first to investigate the use of relative squared error in estimation. From a practical point of view, RE-regression has been applied in diverse fields, including medicine [10] and finance [11]. In [12], authors studied estimation via RE-regression in multiplicative regression models, while [13] extended this framework to nonparametric estimation, establishing the convergence of the local linear estimator derived from relative error loss.

In last decade, RE-regression has been explored extensively for dependent data. For instance, [14] considered quasi-associated time series, and [15] studied spatial processes. The functional nonparametric RE-regression was first introduced by [16], who established strong consistency and derived the asymptotic distribution of the RE-regression estimator. [17] extended the method to incomplete functional data, examining kernel smoothing techniques for truncated observations. The quasi-associated functional time series case was investigated by[18]. For a comprehensive overview of the latest progress on this topic, see [19,20].

While all the cited studies estimate the regression operator using the mean least squares relative error loss function, in the present contribution, we employ the absolute relative error loss function. This approach is combined with kernel-based local weighting techniques to develop a robust estimator for the nonparametric functional regression model. This alternative estimator offers substantial gains in robustness and accuracy compared to the earlier approach. Such an improvement is especially important in functional statistics, where data are typically noisy and high-dimensional. Specifically, in functional statistics, extreme values and rapid fluctuations are common, which can disproportionately influence least squares-based estimators. Unlike the Least Squares Relative Error (LSRE), the Absolute Relative Error (ARE) treats deviations proportionally, making it more robust to outliers. This robustness ensures a more reliable estimation of the underlying functional relationship, making ARE particularly suitable for high-frequency applications in many applied areas such as economics, finance, signal processing, and environmental monitoring. From a theoretical point of view, the manipulation of the LARE-regression model is more challenging than LSRE-regression. This difficulty is because the LSRE-regression can be explicitly expressed as the ratio of the first and second inverted conditional moments, whereas the LARE-regression lacks such an explicit form. Its asymptotic properties must be derived through its Bahadur representation. Despite this additional complexity, we have established the almost complete convergence of the proposed estimator with precision of the convergence rate. The latter is stated under standard assumptions in nonparametric functional statistics. The assumed conditions explore the functional nature of the model and the structure of the data. An additional advantage of the LARE-regression model is its broader scope of applicability compared to the LSRE-regression. Indeed, the LSRE-regression is limited to strictly positive response variables; in contrast, the LARE regression overcomes this restriction allowing it to extend its potential applicability. Furthermre, the practical usefulness of the proposed estimator is demonstrated through data-driven selection procedures for tuning parameters, as well as through simulation studies and real data applications, which confirm its robustness and superior performance in practice.

The paper is organized as follows. In Section 2, we introduce the proposed estimation algorithms. Section 3 presents the required assumptions and the main asymptotic results. In Section 4, we discuss methods for selecting the smoothing parameter and the metric using the random matrix theory from econophysics. The performance of the constructed estimator on simulated and real data is evaluated in Section 5. Finally, Section 6 is devoted to summarize some concluding remarks. The technical proofs are provided in the appendix's Appendix A.

 doi:10.20944/preprints202512.0947.v1

2. Functional LARE Regression and Its Estimation

As outlined in the introduction, our main objective is to evaluate the relationship between an exogenous functional variable T and a real-valued endogenous variable S. More precisely, the variables under consideration belong to the space $\mathcal{X} \times \mathbb{R}$. The set $\mathcal{X}$ is a functional space endowed with an appropriate semi-norm N. Henceforth. we fix a point $\mathcal{T}$ in $\mathcal{X}$ and consider a neighborhood $\mathcal{N}_{\mathcal{T}}$ of $\mathcal{T}$ within this functional space. In our earlier contributions ([16,18]), we modeled the underlying relationship between T and S using the LSRE regression defined by

$$REG(\mathcal{T}) = \arg\min_{s} \mathbb{E}\left(\left(\frac{S-s}{S}\right)^2 | T = \mathcal{T}\right).$$

The main feature of this loss function is its ability to provide an explicit definition of the predictor. Indeed, by simple differentiation with respect to θ, we can show that

$$REG(\mathcal{T})) = \frac{\mathbb{E}\left[S^{-1}\middle|T=\mathcal{T}\right]}{\mathbb{E}\left[S^{-2}\middle|T=\mathcal{T}\right]}. \tag{1}$$

Clearly, such an explicit definition of the model simplifies the analysis and derivation of its mathematical properties. Nevertheless, this approach is not without limitations. While the least squares relative error method reduces the effect of large outliers values, it can give excessive weight to large deviations which compromised its robustness against the very small values or highly dispersed observations. To address this limitation, we propose to proceed with the least absolute relative error (LARE) approach, which provides a more balanced treatment of both small and large deviations in the data. Formally, the LARE regression is defined as the estimator that minimizes

$$RER(\mathcal{T}) = \arg\min_{s} \mathbb{E}\left(\left|\frac{S-s}{S}\right| | T = \mathcal{T}\right). \tag{2}$$

Evidently, this alternative loss function enhances the robustness of LSRE regression by treating all deviations proportionally in a linear manner, rather than using the traditional quadratic form. The LARE loss function evaluates the differences between predicted and true values more equitably, providing a more balanced, resilient, and reliable assessment of model performance. This advantage is particularly important in the presence of large deviations or extremely small observations, where the LSRE approach fails to handle. Overall, the LARE rule is more robust and effective, especially in datasets with irregular or heavy-tailed distributions.

For the estimation setup, we consider n independent and identically distributed (i.i.d.) pairs $(T_1, S_1), \ldots, (T_n, S_n)$, sampled from the distribution of (T, S). The kernel estimator of $RER(\mathcal{T})$ is then given by

$$\widehat{RER}(\mathcal{T}) = \arg\min_{s} \frac{\sum_{i=1}^{n} M(m^{-1}N(\mathcal{T}, T_i))\left|\frac{S_i - s}{S_i}\right|}{\sum_{i=1}^{n} M(m^{-1}N(\mathcal{T}, T_i))},$$

where M is a kernel function and $m := m_n$ (to simplify the notation) is a sequence of positive real numbers that goes to zero as n goes to infinity.

The primary objective of this paper is to establish the asymptotic properties of the estimator $\widehat{RER}(\mathcal{T})$ of $RER(\mathcal{T})$, where the explanatory variable T takes values in a semi-metric space $\mathcal{X}$. To the best of our knowledge, this work is the first to employ this approach to construct an estimator of the regression operator, even in multivariate settings. Although the finite-dimensional case can be viewed as a special case of this framework, the infinite-dimensional case is especially interesting, both for its theoretical challenges and the variety of practical applications it supports. For further motivation and detailed discussions on this topic, we refer the reader to [21–23] and the references therein.

 doi:10.20944/preprints202512.0947.v1

3. The Strong Consistency of the Kernel Estimator

Now, to investigate the almost complete convergence, we denote by C or C' strictly positive constants and put

$$M_i = M\big(m^{-1}N(\mathcal{T},T_i)\big), \quad i=1,\dots,n.$$

We define the closed ball of radius r centered at $\mathcal{T}$ as

$$B(\mathcal{T},r)=\{\mathcal{Z}\in\mathcal{X}: N(\mathcal{Z},\mathcal{T})<r\},$$

and set

$$U_1(s,\mathcal{T})=\mathbb{E}[|S|^{-1}\mathbb{1}_{S\le s}\mid T=\mathcal{T}], \quad .U_2(s,\mathcal{T})=\mathbb{E}[|S|^{-1}\mathbb{1}_{S\ge s}\mid T=\mathcal{T}].$$

The following assumptions are imposed:

(HM1) The small-ball probability satisfies $\mathbb{P}(X\in B(\mathcal{T},r))=:\mathcal{G}_{\mathcal{T}}(r)>0$ for all $r>0$, with

$$\lim_{r\to 0}\mathcal{G}_{\mathcal{T}}(r)=0.$$

(HM2) For all $\mathcal{T}_1,\mathcal{T}_2\in\mathcal{V}_{\mathcal{T}}$, the functions U_γ are of class C^1 (with respect to s) and satisfy

$$|U_\gamma(s,\mathcal{T}_1)-U_\gamma(s,\mathcal{T}_2)|\le C\,d^{b_\gamma}(\mathcal{T}_1,\mathcal{T}_2), \quad b_\gamma>0.$$

(HM3) The kernel M is measurable, supported on $(0,1)$, and bounded: $0<C_2\le M(\cdot)\le C_3<\infty$.

(HM4) The small-ball probability satisfies

$$\frac{n\,\mathcal{G}_{\mathcal{T}}(m)}{\log n}\to\infty \quad \text{as } n\to\infty.$$

(HM5) The response variable is bounded by inverse moments:

$$\forall p\ge 2, \quad \mathbb{E}[S^{-p}\mid T=\mathcal{T}]<C<\infty.$$

These assumptions are generally mild and consistent with those commonly adopted in the literature on functional mode estimation. The functional component is typically described by two main conditions. The first, (HM1), concerns the small-ball probability function, which captures the concentration of the probability measure of the functional regressor. The second, (HM2), addresses the time-series structure of the functional data, quantifying the local dependence among functional observations. Both of these conditions are standard in functional time series analysis (see [24] for detailed discussions and examples of functional data satisfying these assumptions). The nonparametric component is characterized by (Hy3), which is commonly used in nonparametric functional statistics and has appeared in several previous studies. Finally, assumptions (HM3)–(HM5) impose technical conditions that facilitate the proofs and ensure the validity of the asymptotic results.

Theorem 1. *Given assumptions (HM1)–(HM5), we obtain*

$$|\widehat{RER}(\mathcal{T})-RER(\mathcal{T}))|=O(m^{b_1})+O(m^{b_2})+O_{a.co.}\left(\sqrt{\frac{\log n}{n\,\mathcal{G}_{\mathcal{T}}(m)}}\right). \tag{3}$$

The proof of this Theorem employs the Bahadur representation of the $\widehat{RER}$ estimator. This representation is established through the following auxiliary results.

 doi:10.20944/preprints202512.0947.v1

Lemma 1. *Let $\{\mathcal{U}_n\}$ be a sequence of decreasing real-valued random functions and $\{\mathcal{D}_n\}$ a real random sequence satisfying*

$$\mathcal{D}_n = o_{a.co.}(1),$$
$$\sup_{|d|\leq A} |\mathcal{U}_n(d) + \beta d - \mathcal{D}_n| = o_{a.co.}(1),$$

for some constants $\beta, A > 0$. Then, for any real random sequence d_n such that $\mathcal{U}_n(d_n) = o_{a.co.}(1)$, we have

$$\sum_{n=1}^{\infty} \mathbb{P}(|d_n| \geq A) < \infty.$$

4. Smoothing Parameter Selection and Random Matrix Theory for Metric Approximation

The practical performance of the estimator $\widetilde{R}$ depends crucially on the choice of parameters involved in its construction. In particular, the bandwidth parameter m plays a central role in the implementation of this regression method. Various strategies have been proposed in the nonparametric regression literature to address this issue. In this study, we adopt two widely used approaches from classical regression: the cross-validation rule and the bootstrap algorithm.

4.1. Leave-One-Out Cross-Validation Principle

In regression analysis, the leave-one-out cross-validation (LOOCV) criterion is based on minimizing the mean squared error. This criterion is highly effective for bandwidth selection because it directly optimizes for predictive accuracy. The LOOCV algorithm was successfully extended to the domain of functional statistics by [25] who demonstrated its particular efficacy and suitability for functional data. The main advantage of this approach is its flexibility; for instance, it can be readily adapted to our model for functional LARE regression as

$$m^{opt} = \arg\min_{m\in\mathcal{M}} \sum_{i=1}^{n} \left| \frac{(S_i - \widetilde{RER}^{-i}(T_i))}{S_i} \right|, \tag{4}$$

where $\widetilde{RER}^{-i}(T_i)$ and $\widehat{RER}^{-i}(T_i)$ denote the leave-one-out estimators of RER, respectively, computed without the i-th observation (T_i, S_i). The efficiency of these estimators depends on the choice of the subset $\mathcal{M}$, which determines the optimization of (4).

Concerning the subset $\mathcal{M}$, we point out that there are two primary strategies for choosing it. The local approach and the global one. In the local case, $\mathcal{M}$ is determined by the number of neighborhoods around the location point. In the global case, $\mathcal{M}$ is derived from the quantiles of the distance distribution between the functional regressors.

4.2. Bootstrap Approach

The bootstrap procedure is an alternative popular method to the cross-validation rule. Its principal feature is its data-driven nature, as it empirically approximates the sampling distribution of the estimator without external assumptions. This automatism analysis makes it particularly valuable in complex modeling scenarios, where asymptotic theory may be intractable or slow to converge. Furthermore, the bootstrap provides a direct and computationally feasible means to minimize a targeted loss function. This allows for a more adaptive for many loss functions. For the LARE-regression, we state the following algorithm :

Step 1: Choose an arbitrary bandwidth m_0 and compute $\widehat{RER}_{m_0}(\mathcal{T})$.

Step 2: Compute the residual $\hat{\epsilon} = S - \widehat{RER}_{m_0}(\mathcal{T})$.

Step 3: Generate a bootstrap residual ϵ^* using the distribution. $G^* = ((\sqrt{5}+1)/2\sqrt{5})d_{\hat{\epsilon}(1-\sqrt{5})/2} - ((1-\sqrt{5})/2\sqrt{5})d_{\hat{\epsilon}(\sqrt{5}+1)/2}$, ($d$ is dirac measure).

 doi:10.20944/preprints202512.0947.v1

Step 4: Construct a bootstrap sample $(S_i^*, T_i^*)_i = (\epsilon^* - \widehat{RER}_{m_0}(T_i)$

Step 5: Compute the estimators $\widehat{RER}_m(\mathcal{T})$ using the sample $(S_i^*, T_i^*)_i$.

Step 6: Repeat the steps 3-6 N_B times and we put $\widehat{RER}_m^r(\mathcal{T})$ the estimator at the replication r.

Step 7: Choose the optimal bandwidth m according to the following criterion

$$m_{opt^{Boo}} = \arg\min_m \sum_{r=1}^{N_B} \sum_{i=1}^{n} |\widehat{RER}_{m_0}(T_i) - \widehat{RER}_m(T_i)|.$$

4.3. Random Matrix Metric

The principal task of functional data analysis is to move from comparing two curves directly to comparing them within the context of a larger, noisy system. In this context, Random Matrix Theory (RMT) is used to define a stable distance (without noise) between any two curves by eliminating the noise of the entire system. The main idea is based on the use of RMT to clean the covariance matrix estimated from the data. The eigenvalues and eigenvectors of the noisy covariance matrix are a mixture of true signal and random noise. RMT provides a null model to identify which parts of the spectrum are likely to be signal and which are noise. To do that, we follow the following algorithm:

- *Step 1:*

  Create the sample covariance matrix

  $$\mathbf{C} = \frac{1}{n}\mathbf{T}^{tr}\mathbf{T}$$

  Here $\mathbf{T}$ is assumed to have been column-centered. This matrix $\mathbf{C}$ gives the observed covariances between all pairs of sampling points.

- *Step 2:*

  The Marčenko-Pastur law describes the eigenvalue distribution for a random matrix. We compare the eigenvalue spectrum of our empirical $\mathbf{C}$ to the spectrum predicted by the Marčenko-Pastur distribution for a random matrix with the same dimensions and variance.

  1. Perform the eigen-decomposition: $\mathbf{C} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{tr}$, where $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues λ_i.
  2. Identify the eigenvalues that are outside the support of the random distribution. These are considered as signal eigenvalues. The eigenvalues within the random bounds are considered noise.

- *Step 3:*

  Create a cleaned covariance matrix $\mathbf{C}_{\text{clean}}$ by retaining only the signal components. A common method is to replace the noise eigenvalues with a constant value:

  $$\mathbf{\Lambda}_{\text{clean}} = \text{diag}\big(\max(\lambda_1, \bar{\lambda}_{\text{noise}}), \max(\lambda_2, \bar{\lambda}_{\text{noise}}), \dots\big)$$

  $$\mathbf{C}_{\text{clean}} = \mathbf{V}\mathbf{\Lambda}_{\text{clean}}\mathbf{V}^{tr}$$

  Where $\bar{\lambda}_{\text{noise}}$ is a representative value for the noise eigenvalues .

 doi:10.20944/preprints202512.0947.v1

- *Step 4:*

  With the cleaned covariance matrix, we can define a stable metric. The Mahalanobis distance is a natural choice. For two curves represented as vectors $\vec{\mathcal{T}}$ and $\vec{\mathcal{T}}'$, the **RMT-metric** is:

$$d_{\text{RMT}}(\vec{\mathcal{T}}, \mathcal{T}') = \sqrt{(\mathcal{T} - \mathcal{T}')^{tr}\, \mathbf{C}_{\text{clean}}^{-1}\, (\mathcal{T} - \mathcal{T}')}$$

  This metric is more stable and reliable than one calculated from the noisy covariance matrix $\mathbf{C}^{-1}$, which is highly distorted by spurious correlations.
  In conclusion, RMT acts as a filter to denoise the covariance structure of the dataset. This cleaned structure increases the accurate metric for comparing any two individual curves within that dataset.

5. Data-Driven Analysis

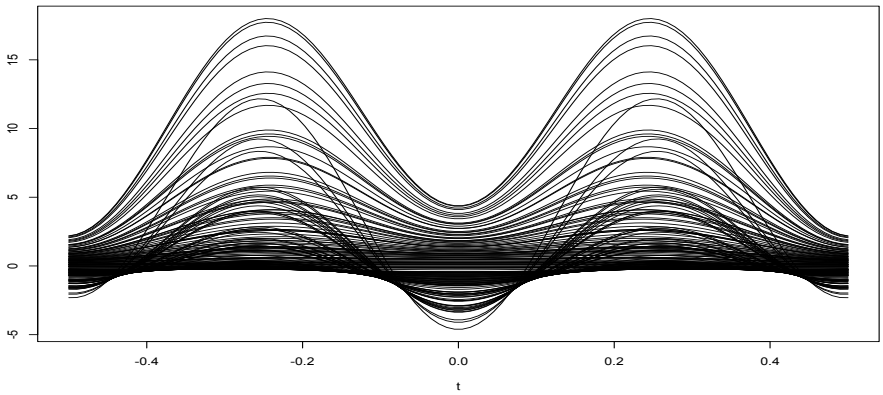
5.1. A Simulated Data Case

This section is devoted to demonstrate the practical implementation of the proposed estimator through two illustrative examples. First, we assess the feasibility of both selection methods (LOOCV and Bootstrap selectors). Second, we illustrate the superior efficiency of the LARE regression over the LSRE approach. For this purpose, we generate two types of functional variables:

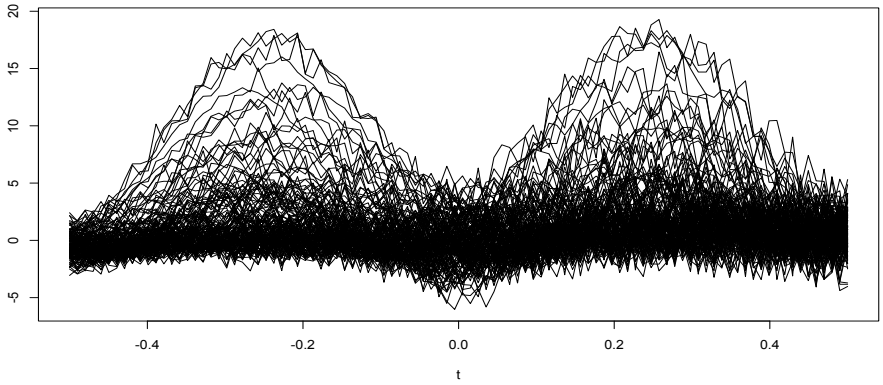
$$\begin{cases} \text{Smooth functional regessor} & T_i(t) = 2b_i \cos(\pi t) + 3b_i \sin^2(2\pi t) + 4b_i t^2 \ \text{ for all } t \in (-1/2, 1/2) \\ \text{Rough functional regessor} & T_i(t) = 2b_i \cos(\pi t) + 3b_i \sin^2(\pi t) + 4b_i t + \eta_i(t) \ \text{ for all } t \in (-1/2, 1/2) \end{cases}$$

where b_i is generated from a $N(0,1)$ distribution and $\eta_i(t)$ is drown from $N(t, 1+t)$ distribution. The curves are then discretized over a grid of points given by 100 equispaced measurements in $(0,1)$. A visual comparison of the two functional samples is provided in (cf. Figure 1).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 December 2025 doi:10.20944/preprints202512.0947.v1



(**a**) A sample of 150 of smooth curves.



(**b**) A sample of 150 of rough curves.

Figure 1. A sample of 150 curves.

We model the scalar response Y using the following non-parametric regression

$$S_i = 3\int_{-0.5}^{0.5} \frac{dt}{1+T_i^2(t)} + \epsilon_i \text{ for } i = 1,\ldots,150, \tag{5}$$

where the errors (ϵ_i) are assumed to be independent of (T_i). For the simulation study, we consider four error types of white noise generated from a normal mixture distribution, as follows:

$$\begin{cases} \text{Standard normal distribution (S.N.D.)} & N(0,1) \\ \text{Skewed unimodal distribution (S.U.D.)} & \frac{1}{5}N(0,1)+\frac{1}{5}N(\frac{1}{2},\frac{4}{9})+\frac{3}{5}N(\frac{13}{12},\frac{25}{81}) \\ \text{Skewed bimodal distribution (C.B.D.)} & \frac{3}{4}N(0,1)+\frac{1}{4}N(\frac{3}{2},\frac{1}{9}) \\ \text{Claw distribution (C.D.)} & \frac{1}{2}N(0,1)+\sum_{i=1}^{4}\frac{1}{10}N(\frac{i}{2}-1,\frac{1}{100}) \end{cases}$$

Further details on normal mixture distributions can be found in [26]. Now, in order to compare the LOOCV to the Bootstrap selector, we optimize both rules over $\mathcal{M}$, the set of m such that the ball centered at $\mathcal{T}$ with radius m contains exactly k neighbors of $\mathcal{T}$. Specifically, the number k is selected from the subset $\{5,15,25,\ldots,75\}$. Observe that the Bootstrap selector requires a pilot bandwidth m_0 which is chosen using the R- routine *h.default* from the R-package *fda.usc* . The two selection procedures performance are tested by comparing their averaged absolute errors ASE, where

$$\text{ASE} = \frac{1}{150}\sum_{i=1}^{150}\left|S_i - \widehat{RER}(T_i)\right|.$$

 doi:10.20944/preprints202512.0947.v1

The model performance also depends critically on using a metric N that is appropriate for the data type. The latter is related to the smoothness of the curves. Here, the shape of the curves (cf. Figures 1) confirms that the metric defined by the derivatives is adequate for the smooth curves,

$$N(\mathcal{T},\mathcal{T}') = \left(\int_0^1 (\mathcal{T}^{(i)}(t) - \mathcal{T}'^{(i)}(t))^2 dt\right)^{1/2},$$

where $\mathcal{T}^{(i)}$ denotes the ith derivative of the curve $\mathcal{T}$. While for the rough curves we employ the PCA semi metric associated to the q-greatest eigenvalues.
While the simulation study investigated numerous values of i, q, and n, we report here only the results for $i = 2$, $q = 1$, and $n = 150$ for brevity. A quadratic kernel on $(0,1)$ was used for this empirical analysis . Table 1 summarizes the obtained ASE-results for both selectors.

Table 1. ASE results.

Method	Distribution	LOOCV	Bootstrap
Smooth curves selector	S.N.D.	0.83	0.97
	S.U.D.	0.98	1.23
	S.B.D.	1.12	1.37
	C.D.	1.18	1.23
Rough curves	S.N.D.	1.17	1.84
	S.U.D.	1.26	1.74
	S.B.D.	1.29	2.18
	C.D.	1.52	1.57

We observe that the LOOCV strategy yields superior results compared to the Bootstrap method. This advantage is pronounced in the case of Skewed Bimodal Distributions (S.B.D.), whereas for Contaminated Distributions (C.D.), the two methods perform nearly equivalently. Furthermore, both bandwidth selectors demonstrate relatively strong efficiency across the tested scenarios.

In the second part of this illustration, we examine the robustness of the proposed estimator by comparing it to the LSER-regression defined by

$$\widetilde{RER} = \frac{\sum_{i=1}^n S_i^{-1} M(m^{-1}N(\mathcal{T},T_i))|}{\sum_{i=1}^n M(m^{-1}N(\mathcal{T},T_i))S_i^{-2}}.$$

It is important to note that robustness can be assessed through various metrics, including the influence function, gross-error sensitivity, breakdown point, local shift sensitivity or bias/variance under contamination, among others (see [27] for a comprehensive overview). While some of these metrics are challenging to apply in the context of functional statistics, they all share the common goal of evaluating the stability of a statistical method and quantifying its sensitivity to deviations from model assumptions. In this experiment, we assess robustness by analyzing the variance under the contaminated data-generating process. Specifically, we construct a contamination model by taking it as a white noise, where

$$\varepsilon_i = t\varepsilon_{1_i} + (1-t)\varepsilon_{2_i}, \quad t \in (0,1), \quad i = 1,2,$$

where ε_{1_i} is generated from S.N.D and ε_{2_i} is drown from C.D. Under this contamination, the empirical variance is defined by

$$V(t) = Var(\widehat{RER}(T_i))_{i=1}^n.$$

In Figures 2 and 3, we plot the function $V(t)$ for both estimators in both cases (smooth and rough case).

 doi:10.20944/preprints202512.0947.v1

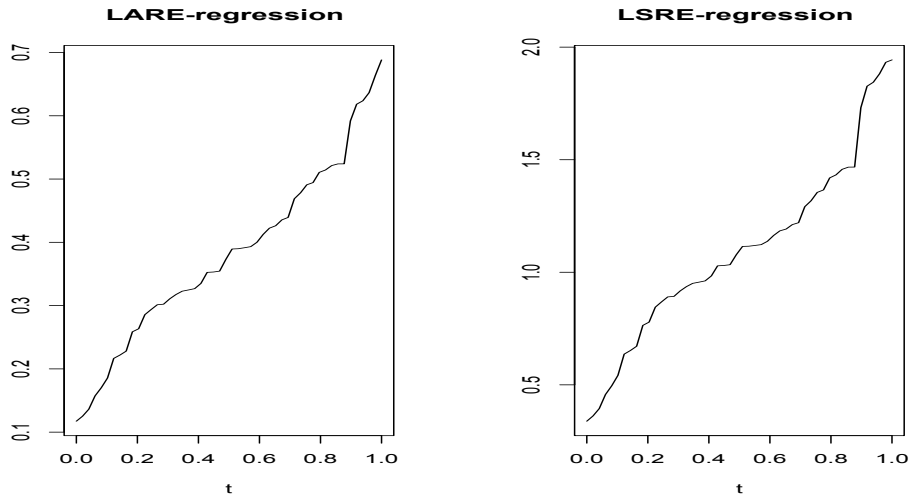


Figure 2. *Smooth case: The variability of the function $V(t)$ over 50 point in $(0,1)$.*

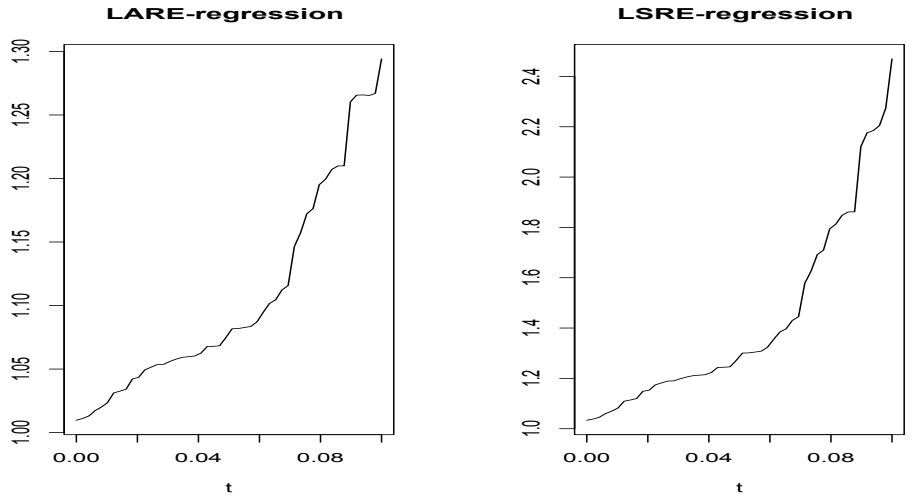


Figure 3. *Rough case: The variability of the function $V(t)$ over 50 point in $(0,1)$.*

Without surprising Figures 2 and 3 confirm the robustness of $\widehat{RER}$ compared to $\widetilde{RER}$. It shows that the function $V(t)$ of $\widehat{RER}$ has a small variation compared to the variation in $\widetilde{RER}$. This conclusion is confirmed by computing the range of $V(t)$, defined as

$$RG = \max V(t) - \min V(t),$$

which serves as a measure of the function's variability. This benchmark is computed over 50 discretized points in the interval (0,1). Specifically, in the smooth case, we obtained 0.6 for the LARE-regression compared to 1.9 for the LSRE-regression. The same statement for the rough case was observed, and RG =0.3 for the LARE-regressions against $RG = 1.45$ for the LSRE-regression case.

5.2. Real Data Application

This section presents a predictive analysis using the LARE-model. Specifically, we conduct a comparative study of several regression predictors. More precisely, we focus in this empirical analysis on prediction of the sulfone content in onion diets using Near-Infrared (NIR) spectral curves. Sulfones is organic compounds highly beneficial for human health. Their biomedical significance is well-documented in the literature, with studies indicating efficacy in treating conditions such as mucous-membrane inflammation and gastrointestinal disorders. Furthermore, research has highlighted their

 doi:10.20944/preprints202512.0947.v1

potential to suppress tumor initiation (see [28]). Our objective is to use the NIR spectrometry to quantify the relationship between onion consumption in a rodent model and subsequent sulfone levels in urine. The functional NIR data utilized in this analysis are available in the website http://www.models.life.ku.dk (accessed on 20 October 2025). The sample preparation involved a two-step chemical process: initial spectrometry processing, followed by spectral acquisition in the wavelength region of 9.6 to 0.3 ppm. The resulting dataset comprises 32 observations, which were recolored with sufficient replication. The spectral curves for these observations are plotted in Figure 4. Further details on this dataset are available in [28].

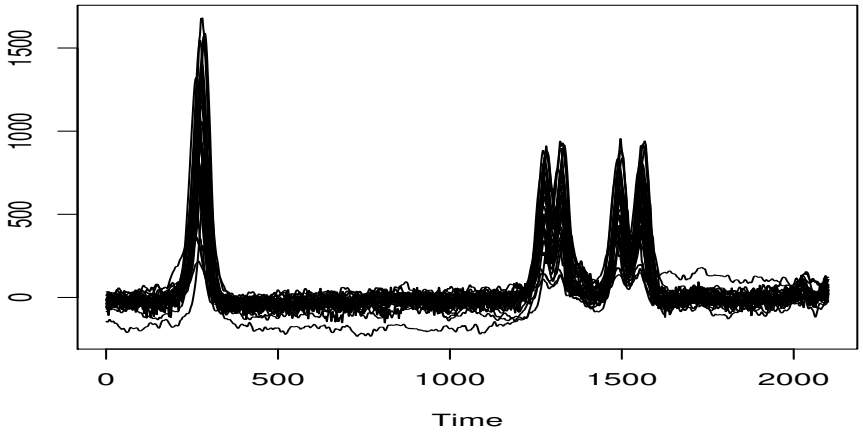


Figure 4. The spectral curves of the NIR data.

For the prediction task, we compare the performance of three regression estimators:

$$\widehat{RER}(\mathcal{T}) = \arg\min_{s} \frac{\sum_{i=1}^{n} M(m^{-1}N(\mathcal{T},T_i))\left|\frac{S_i - s}{S_i}\right|}{\sum_{i=1}^{n} M(m^{-1}N(\mathcal{T},T_i))} \quad \text{(LARE-Regression)},$$

$$\widetilde{RER}(\mathcal{T}) = \frac{\sum_{i=1}^{n} S_i^{-1} M(m^{-1}N(\mathcal{T},T_i))|}{\sum_{i=1}^{n} M(m^{-1}N(\mathcal{T},T_i))S_i^{-2}} \quad \text{(LSRE-Regression)},$$

$$\overline{RE}(\mathcal{T}) = \frac{\sum_{i=1}^{n} S_i M(m^{-1}N(\mathcal{T},T_i))|}{\sum_{i=1}^{n} M(m^{-1}N(\mathcal{T},T_i))} \quad \text{(NW-Regression)}.$$

To ensure a fair comparison, common computational strategies were employed for the three estimators. This includes the use of the RMT-metric, an identical kernel function M and the application of the rule specified in Equation (4) for selecting the bandwidth parameter for the three methods.

To evaluate the robustness of these estimators, we introduced contamination by multiplying a subset of observations by a factor of 10. The impact of this contamination was assessed by controlling the variation in the Mean Absolute Error (MAE). Specifically, for a given contamination proportion $t \in [0,1]$, we multiplied $[tn]$ observations (where $[.]$ denotes the integer part) and computed the variability of the prediction error function using

$$\text{AME}(t) = \frac{1}{35}\sum_{k=1}^{35}(S_k - \dot{Pred}(T_k)),$$

where $\dot{Pred}$ denotes one of the three compared estimators. The resulting prediction errors are displayed in Figure 5.

 doi:10.20944/preprints202512.0947.v1

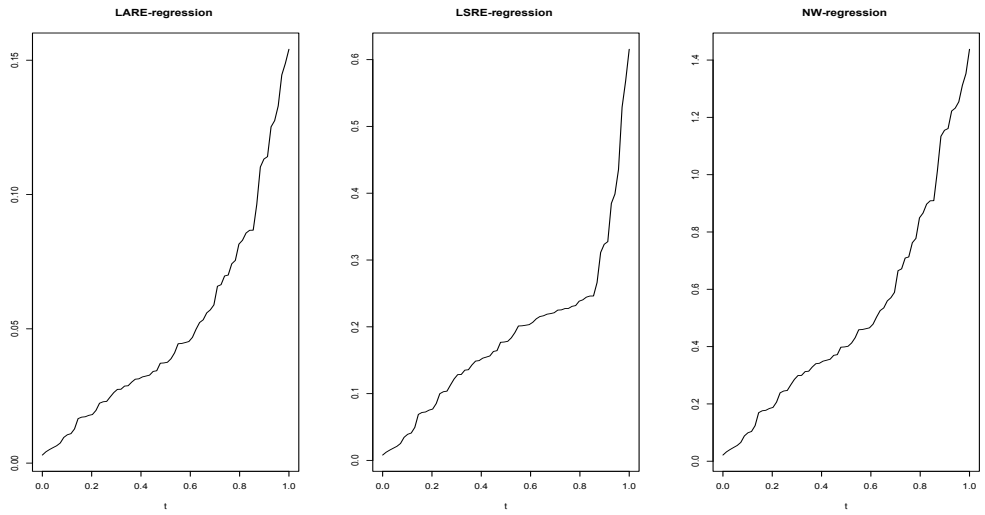


Figure 5. Comparison of prediction errors across the three regression estimators under increasing data contamination.

The results indicate that while all three predictors have satisfactory performance under uncontaminated conditions($t = 0$), the LARE approaches demonstrates a significant advantage. This advantage is particularly evident in the stability of the LARE estimator ($\widehat{RER}$), which exhibits lower variability in its MAE under contamination. The superior stability of the LARE method is quantitatively confirmed by examining the range of its error function. The range of the LARE-regression was observed to be between 0.02 and 0.15, compared to a range of $[0.02,\ 0.6]$ for the LSRE regression and $[0.05,\ 1.4]$ for the standard nonparametric regression. This empirical finding aligns with the theoretical expectation that estimators based on absolute relative error (LARE) are more robust than those based on squared error (LSRE), due to the reduced influence of large residuals.

6. Conclusions and Prospects

This work introduces a robust framework for functional regression estimation by employing the least absolute relative error loss function. We construct a novel estimator and establish its asymptotic properties, specifically, proving its complete consistency. The functional structure is modeled using the small ball probability function that relates the topological structure of the probability measure of the functional explanatory variable. Furthermore, the employment of the LARE rule significantly increases the robustness of the model. The derived asymptotic results are obtained under general assumptions, ensuring their applicability to a broad class of functional data and continuous-time processes. Empirical analyses demonstrate the straightforward implementation of our estimator and confirm its superior robustness and predictive accuracy compared to alternative methods.
This contribution opens several promising questions for future research. A natural extension would be to develop a local linear version of this robust modal regression estimator. It is well-known that local linear modeling can improve the asymptotic convergence rate by reducing boundary bias. Establishing the asymptotic normality of the proposed robust estimator is another critical open question, as it is a prerequisite for constructing confidence intervals and conducting hypothesis tests. Other important directions include exploring linear or partially linear formulations of this regression. A particularly challenging task would be to extend these results to more complex data structures, such as censored data or functional time series. Clearly, the treatment of these complex structures requires new mathematical tools and developments that extend beyond the scope of the present work.

 doi:10.20944/preprints202512.0947.v1

Appendix A

Proof of Lemma 1. The proof of this lemma adapts techniques employed by [29]. For an arbitrary positive constant $\eta > 0$, we decompose the probability by

$$\mathbb{P}(|d_n| \geq A) = \mathbb{P}(|d_n| \geq A,\ |\mathcal{U}_n(d_n)| < \eta) + \mathbb{P}(|\mathcal{U}_n(d_n)| \geq \eta),$$

and implying

$$\mathbb{P}(|d_n| \geq A) \leq \mathbb{P}\left(\inf_{|d|\geq A} |\mathcal{U}_n(d)| < \eta\right) + \mathbb{P}(|\mathcal{U}_n(d_n)| \geq \eta).$$

Since $\mathcal{U}_n(d_n) = o_{a.co.}(1)$, it remains to establish that

$$\sum_n \mathbb{P}\left(\inf_{|d|\geq A} |\mathcal{U}_n(d)| < \eta\right) < \infty.$$

Utilizing the monotonicity of $\mathcal{U}_n$, we observe that

$$\inf_{|d|\geq A} |\mathcal{U}_n(d)| = \min(-\mathcal{U}_n(A),\ \mathcal{U}_n(-A)).$$

Thus, it suffices to demonstrate that

$$\sum_n \mathbb{P}(-\mathcal{U}_n(A) \leq \eta) < \infty \quad \text{and} \quad \sum_n \mathbb{P}(\mathcal{U}_n(-A) \leq \eta) < \infty.$$

For the first term, we proceed with the following estimation,

$$\begin{aligned}
\sum_n \mathbb{P}(-\mathcal{U}_n(A) \leq \eta) &\leq \sum_n \mathbb{P}(-\mathcal{U}_n(A) \leq \eta,\ A\beta - 2\eta \geq \mathcal{D}_n) \\
&\quad + \sum_n \mathbb{P}(A\beta - 2\eta \leq \mathcal{D}_n) \\
&\leq \sum_n \mathbb{P}(\mathcal{U}_n(M) \geq -\eta,\ A\beta - \mathcal{D}_n \geq 2\eta) \\
&\quad + \sum_n \mathbb{P}(A\beta - 2\eta \leq \mathcal{D}_n) \\
&\leq \sum_n \mathbb{P}(|\mathcal{U}_n(M) + \beta M - \mathcal{D}_n| \geq \eta) \\
&\quad + \sum_n \mathbb{P}(|\mathcal{D}_n| \geq \beta M - 2\eta) \\
&\leq \sum_n \mathbb{P}\left(\sup_{|d_1|\leq A} |\mathcal{U}_n(d_1) + \beta d_1 - \mathcal{D}_n| \geq \eta\right) \\
&\quad + \sum_n \mathbb{P}(|\mathcal{D}_n| \geq \beta M - 2\eta).
\end{aligned}$$

 doi:10.20944/preprints202512.0947.v1

The second term is bounded analogously by

$$\begin{aligned}\sum_n \mathbb{P}(\mathcal{U}_n(-A) \le \eta) &\le \sum_n \mathbb{P}(\mathcal{U}_n(-A) \le \eta,\ 2\eta - A\beta \le \mathcal{D}_n) \\ &\quad + \sum_n \mathbb{P}(2\eta - A\beta \ge \mathcal{D}_n) \\ &\le \sum_n \mathbb{P}(-\mathcal{U}_n(-A) \ge -\eta,\ \mathcal{D}_n + A\beta \ge 2\eta) \\ &\quad + \sum_n \mathbb{P}(A\beta - 2\eta \le -\mathcal{D}_n) \\ &\le \sum_n \mathbb{P}(|\mathcal{U}_n(-A) + \beta(-A) - \mathcal{D}_n| \ge \eta) \\ &\quad + \sum_n \mathbb{P}(|\mathcal{D}_n| \ge \beta M - 2\eta) \\ &\le \sum_n \mathbb{P}\left(\sup_{|d_1| \le A} |\mathcal{U}_n(d_1) + \beta d_1 - \mathcal{D}_n| \ge \eta\right) \\ &\quad + \sum_n \mathbb{P}(|\mathcal{D}_n| \ge \beta A - 2\eta).\end{aligned}$$

Finally, selecting $\eta > \frac{\beta A}{2}$ and invoking condition (ii) yields

$$\sum_n \mathbb{P}(|d_n| \ge A) < \infty.$$

Proof of Theorem 1. First, observe that through straightforward differentiation, we establish that $\widehat{RER}$ satisfies the estimating equation,

$$\sum_{i=1}^{n} |S_i|^{-1}(0.5 - \mathbb{1}_{S_i \le \widehat{RER}})M_i = 0. \tag{A1}$$

The desired result follows from an application of Lemma 1 with the following identifications.

$$\mathcal{U}_n(d) = \frac{1}{n\mathbb{E}[M_1]} \sum_{i=1}^{n} |S_i^{-1}|\Big(0.5 - \mathbb{1}_{S_i \le (d + RER(\mathcal{T}))}\Big)M_i,$$

$$d_n = \widehat{RER}(\mathcal{T}) - RER(\mathcal{T}), \quad \text{and} \quad \mathcal{D}_n = \mathcal{U}_n(0).$$

This gives the required Bahadur representation. From Equation (A1), we have $\mathcal{U}_n(d_n) = 0$. Thus, it remains to verify the conditions of Lemma. 1:

$$\mathcal{D}_n = o_{a.co.}(1), \tag{A2}$$

$$\sup_{|d| \le A} |\mathcal{U}_n(d) + \beta d - \mathcal{D}_n| = o_{a.co.}(1), \quad \text{for some } M, \beta > 0. \tag{A3}$$

The claimed results are presented in the following Lemmas

Lemma A1. *Given assumptions (HM1)–(HM5), we obtain*

$$|\mathcal{D}_n| = O(m^{b_1}) + O(m^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\,\mathcal{G}_{\mathcal{T}}(m)}}\right).$$

Proof. Define the centered random variables,

$$\mathcal{V}_i = |S_i^{-1}|\Big(0.5 - \mathbb{1}_{[S_i \le RER(\mathcal{T})]}\Big)M_i - \mathbb{E}\Big[|S_i^{-1}|\Big(0.5 - \mathbb{1}_{[S_i \le RER(\mathcal{T})]}\Big)M_i\Big].$$

 doi:10.20944/preprints202512.0947.v1

Then, the centered version of $\mathcal{D}_n$ becomes

$$\mathcal{D}_n - \mathbb{E}[\mathcal{D}_n] = \frac{1}{n\mathbb{E}[M_1]} \sum_{i=1}^{n} \mathcal{V}_i.$$

The proof of this lemma relies on the exponential inequality established in Corollary A.8(ii) of [24], which requires a bounding of the quantity $E|\mathcal{V}_i^p|$. First, for any $j \leq m$, we observe that

$$\begin{aligned} E\left|S_1^{-j} M_1^j\right| &= E\left[M_1^j E\left[\left|S_1^{-j\gamma}\right| \mid T_1\right]\right] \\ &= CE\left[M_1^j\right] \\ &\leq C'\mathcal{G}_{\mathcal{T}}(m), \end{aligned}$$

which implies

$$\frac{1}{E^j[M_1]} E\left|S_1^{-j} M_1^j\right| = O\left(\mathcal{G}_{\mathcal{T}}(m)^{-j+1}\right). \tag{A4}$$

Moreover, we have

$$\frac{1}{E[M_1]} E\left[|S_1^{-1}| M_1\right] \leq C.$$

Applying the binomial theorem yields

$$\begin{aligned} E|\mathcal{V}_i^p| &\leq C \sum_{j=0}^{p} \frac{1}{(E[M_1])^j} E\left|S_1^{-j} M_1^j\right| \\ &\leq C \max_{j=0,\ldots,p} \mathcal{G}_{\mathcal{T}}(m)^{-j+1} \\ &\leq C\mathcal{G}_{\mathcal{T}}(m)^{-p+1}. \end{aligned}$$

Consequently,

$$E|\mathcal{V}_i^p| = O(\mathcal{G}_{\mathcal{T}}(m)^{-p+1}). \tag{A5}$$

We now apply the aforementioned exponential inequality with $a = \mathcal{G}_{\mathcal{T}}(m)^{-1/2}$. For all $\eta > 0$,

$$\mathbb{P}\left(|\mathcal{D}_n - E\mathcal{D}_n| > \eta \sqrt{\frac{\log n}{n\,\mathcal{G}_{\mathcal{T}}(m)}}\right) \leq C' n^{-C\eta^2}.$$

Selecting η sufficiently large yields

$$\sum_n \mathbb{P}\left(|\mathcal{D}_n - E\mathcal{D}_n| > \eta \sqrt{\frac{\log n}{n\,\mathcal{G}_{\mathcal{T}}(m)}}\right) < \infty.$$

Finally, we obtain

$$\mathcal{D}_n - \mathbb{E}[\mathcal{D}_n] = O_{a.co.}\left(\sqrt{\frac{\log n}{n\mathcal{G}_{\mathcal{T}}(\mathcal{D}_n)}}\right).$$

Further, we have

$$\begin{aligned} \mathbb{E}[|\mathcal{D}_n] &= \frac{1}{\mathbb{E}[M_1]} \mathbb{E}\left[\left|S_1^{-1}\right| \left(0.5 - \mathbb{1}_{[S_1 \leq RER(\mathcal{T})]}\right) M_1\right] \\ &\leq \frac{1}{\mathbb{E}[M_1]} \mathbb{E}[M_1 | U_1(RER(\mathcal{T}), \mathcal{T}) - U_1(RER(\mathcal{T}), T_1)|] \\ &+ \frac{1}{\mathbb{E}[M_1]} \mathbb{E}[M_1 | U_2(RER(\mathcal{T}), \mathcal{T}) - U_2(RER(\mathcal{T}), T_1)|] \\ &= O(m^{b_1}) + O(m^{b_2}). \end{aligned}$$

 doi:10.20944/preprints202512.0947.v1

It follows that

$$|\mathcal{D}_n| = O(m^{b_1}) + O(m^{b_2}) + O_{a.co.}\left(\left(\frac{\log n}{n\,\mathcal{G}_{\mathcal{T}}(m)}\right)^{1/2}\right)$$

□

Lemma A2. *Given assumptions (HM1)–(HM5), we obtain*

$$\sup_{|d|\le A} |\mathcal{U}_n(d) + \beta d - \mathcal{D}_n| = O(m^{b_1}) + O(m^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\,\mathcal{G}_{\mathcal{T}}(m)}}\right).$$

Proof. We split the proof in two parts: the dispersion term

$$\sup_{|d|\le A} |\mathcal{U}_n(d) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n]| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\,\mathcal{G}_{\mathcal{T}}(m)}}\right) \tag{A6}$$

and the bias term

$$\sup_{|d|\le A} |\mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n] + U((RER(\mathcal{T}), \mathcal{T})d| = O(m^{b_1}) + O(m^{b_2}), \tag{A7}$$

where $U(s, calT) = \frac{\partial U_2(s,\mathcal{T}) - U_1(s,\mathcal{T})}{\partial s}$.

For (A6), we exploit the compactness of the interval $[-A, A]$, and construct a finite covering

$$[-A, A] \subset \bigcup_{j=1}^{d_n} [d_j - l_n, d_j + l_n], \quad \text{with } d_j \in [-A, A] \text{ and } l_n = d_n^{-1} = 1/\sqrt{n}.$$

For any $d \in [-A, A]$, define $j(d) = \arg\min_j |d - d_j|$. Then, we decompose the supremum as

$$\begin{aligned}\sup_{|d|\le A} |\mathcal{U}_n(d) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n]| \le & \sup_{|d|\le A} \left|\mathcal{U}_n(d) - \mathcal{U}_n(d_{j(d)})\right| \\ & + \sup_{|d|\le A} \left|\mathcal{U}_n(d_{j(d)}) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d_{j(d)}) - \mathcal{D}_n]\right| \\ & + \sup_{|d|\le M} \left|\mathbb{E}[\mathcal{U}_n(d) - \mathcal{U}_n(d_{j(d)})]\right|.\end{aligned}$$

Using the inequality $|\mathbb{1}_{[S<a]} - \mathbb{1}_{[S<b]}| \le \mathbb{1}_{[|S-b|\le|a-b|]}$, we bound the first term as

$$\sup_{|d|\le A} |\mathcal{U}_n(d) - \mathcal{U}_n(d_j)| \le \frac{1}{n\mathbb{E}[M_1]} \sum_{i=1}^{n} Z_i,$$

where

$$Z_i = \sup_{|d|\le A} |S_i^{-1}| \mathbb{1}_{[|S_i - d_{j(d)} - RER(\mathcal{T})| \le C l_n]} M_i.$$

The random variables Z_i satisfy

$$\mathbb{E}[Z_i] = O(l_n \mathcal{G}_{\mathcal{T}}(m)), \quad \mathbb{E}[Z_i^p] = O(l_n \mathcal{G}_{\mathcal{T}}^{-(p+1)}(m)).$$

Given the condition

$$l_n = o\left(\left(\frac{\log n}{n\mathcal{G}_{\mathcal{T}}(m)}\right)^{1/2}\right), \tag{A8}$$

 doi:10.20944/preprints202512.0947.v1

we obtain

$$\sup_{|d|\le A} |\mathcal{U}_n(d) - \mathcal{U}_n(d_j)| = O_{a.co.}\left(\left(\frac{\log n}{n\mathcal{G}_{\mathcal{T}}(m)}\right)^{1/2}\right).$$

For the third term, we have

$$\sup_{|d|\le M} |\mathbb{E}[\mathcal{U}_n(d) - \mathcal{U}_n(d_j)]| \le \frac{1}{n\mathbb{E}[M_1]}\mathbb{E}[Z_1] \le Cl_n,$$

and by (A8),

$$\sup_{|d|\le A} |\mathbb{E}[\mathcal{U}_n(d) - \mathcal{U}_n(d_j)]| = o\left(\left(\frac{\log n}{n\mathcal{G}_{\mathcal{T}}(m)}\right)^{1/2}\right).$$

Now consider the middle term,

$$\sup_{|d|\le A} |\mathcal{U}_n(d_j) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d_j) - \mathcal{D}_n]|.$$

Define

$$\mathcal{U}_n(d_j) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d_j) - \mathcal{D}_n] = \frac{1}{n\mathbb{E}[M_1]}\sum_{i=1}^{n}\Psi_i,$$

where

$$\Psi_i = \frac{1}{\mathbb{E}[M_1]}|S_i^{-1}|\left(\mathbb{1}_{S_i\le RER(\mathcal{T})} - \mathbb{1}_{S_i\le d_j+RER(\mathcal{T})}\right)M_i - \mathbb{E}\left[\left(\mathbb{1}_{S_i\le RER(\mathcal{T})} - \mathbb{1}_{S_i\le d_j+RER(\mathcal{T})}\right)M_i\right].$$

The variables Ψ_i satisfy

$$|\mathbb{E}[\Psi_i^p]| \le C\mathcal{G}_{\mathcal{T}}^{-(p+1)}(m).$$

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

$$\begin{aligned}
&\sum_n \mathbb{P}\left(\sup_{|d|\le A} |\mathcal{U}_n(d_{j(d)}) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d_{j(d)}) - \mathcal{D}_n]| \ge \eta\sqrt{\frac{\log n}{n\mathcal{G}_{\mathcal{T}}(m)}}\right)\\
&\le \sum_n d_n \max_j \mathbb{P}\left(|\mathcal{U}_n(d_j) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d_j) - \mathcal{D}_n]| \ge \eta\sqrt{\frac{\log n}{n\mathcal{G}_{\mathcal{T}}(m)}}\right) < \infty,
\end{aligned}$$

which completes the proof of (A6).

For the bias term (A7), we employ similar arguments. For all $d \in [-M, M]$,

$$\begin{aligned}
\mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n] &= -\frac{1}{\mathbb{E}[M_1]}\mathbb{E}\left[|S^{-1}|\left(\mathbb{1}_{[S_1\le d+RER(\mathcal{T})]} - \mathbb{1}_{[S_1\le RER(\mathcal{T})]}\right)M_1\right]\\
&= -\frac{1}{\mathbb{E}[M_1]}\mathbb{E}[(U_1(d+RER(\mathcal{T}),\mathcal{T}) - U_1(RER(\mathcal{T}),T|))M_1]\\
&+ \frac{1}{\mathbb{E}[M_1]}\mathbb{E}[(U_2(d+RER(\mathcal{T}),\mathcal{T}) - U_2(RER(\mathcal{T}),T|))M_1]\\
&= -\frac{1}{\mathbb{E}[M_1]}\mathbb{E}[(U_1(d+RER(\mathcal{T}),\mathcal{T}) - U_1(RER(\mathcal{T}),\mathcal{T}|))M_1]\\
&+ \frac{1}{\mathbb{E}[M_1]}\mathbb{E}[(U_2(d+RER(\mathcal{T}),\mathcal{T}) - U_2(RER(\mathcal{T}),\mathcal{T}|))M_1] + O(m^{b_1}) + O(m^{b_2})\\
&= -U((RER(\mathcal{T}),\mathcal{T})d + O(m^{b_1}) + O(m^{b_2}) + o(d).
\end{aligned}$$

Thus,

$$\mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n] = U((RER(\mathcal{T}),\mathcal{T})d + O(m^{b_1}) + O(m^{b_2}) + o(d).$$

 doi:10.20944/preprints202512.0947.v1

Therefore, the conditions of Lemma 1 are satisfied, yielding the Bahadur representation,

$$\widehat{RER}(\mathcal{T}) - RER(\mathcal{T}) = d_n = \frac{1}{U((RER(\mathcal{T}), \mathcal{T})} \mathcal{D}_n + O\left(\sup_{|d| \leq A} |\mathcal{U}_n(d) + \beta d - \mathcal{D}_n| \right).$$

□

Author Contributions: The authors contributed approximately equally to this work. Formal analysis, M.R.; Writing—review & editing, I.A. and A. L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deanship of Scientific Research and Graduate Studies at King Khalid University for funding this work through Research Groups under grant number RGP2/456/46.

Data Availability Statement: The data used in this study are available through the link http://www.models.life.ku.dk (accessed on 20 October 2025).

Acknowledgments: The authors thank and extend their appreciation to the funder of this work: Deanship of Scientific Research and Graduate Studies at King Khalid University for funding this work through Research Groups under grant number RGP2/456/46.

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

References

1 Ferraty, F. and Vieu, P., 2002. The functional nonparametric model and application to spectrometric data. Computational Statistics, 17(4), pp.545-564.
2 Dabo-Niang, S. and Rhomari, N., 2003. Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. Comptes rendus. Mathématique, 336(1), pp.75-80.
3 Delsol, L., 2007. Régression non-paramétrique fonctionnelle: Expressions asymptotiques des moments. In Annales de l'ISUP (Vol. 51, No. 3, pp. 43-67).
4 Masry, E., 2005. Nonparametric regression estimation for dependent functional data: asymptotic normality. Stochastic Processes and their Applications, 115(1), pp.155-177.
5 Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P., 2010. Rate of uniform consistency for nonparametric estimates with functional variables. Journal of Statistical planning and inference, 140(2), pp.335-352.
6 Attouch, M.K., Laksaci, A. and Saïd, E.O., 2013. Robust regression for functional time series data. Journal of the Japan Statistical Society, 42(2), pp.125-143.
7 Barrientos-Marin, J., Ferraty, F. and Vieu, P., 2010. Locally modelled regression and functional data. Journal of Nonparametric Statistics, 22(5), pp.617-632.
8 Azzi, A., Belguerna, A., Laksaci, A. and Rachdi, M., 2024. The scalar-on-function modal regression for functional time series data. Journal of Nonparametric Statistics, 36(2), pp.503-526.
9 Narula, S.C. and Wellington, J.F., 1977. Prediction, linear regression and the minimum sum of relative errors. Technometrics, 19(2), pp.185-190.
10 Chatfield, C., 2007. The joys of consulting. Significance, 4(1), pp.33-36.
11 Chen, K., Guo, S., Lin, Y. and Ying, Z., 2010. Least absolute relative error estimation. Journal of the American Statistical Association, 105(491), pp.1104-1112.
12 Yang, Y. and Ye, F., 2013. General relative error criterion and M-estimation. Frontiers of Mathematics in China, 8(3), pp.695-715.
13 Jones, M.C., Park, H., Shin, K.I., Vines, S.K. and Jeong, S.O., 2008. Relative error prediction via kernel regression smoothers. Journal of Statistical Planning and Inference, 138(10), pp.2887-2898.
14 Mechab, W. and Laksaci, A., 2016. Nonparametric relative regression for associated random variables. Metron, 74(1), pp.75-97.
15 Attouch, M., Laksaci, A. and Messabihi, N., 2017. Nonparametric relative error regression for spatial random variables. Statistical papers, 58(4), pp.987-1008.
16 Demongeot, J., Hamie, A., Laksaci, A. and Rachdi, M., 2016. Relative-error prediction in nonparametric functional statistics: Theory and practice. Journal of Multivariate Analysis, 146, pp.261-268.

 doi:10.20944/preprints202512.0947.v1

17 Altendji, B., Demongeot, J., Laksaci, A. and Rachdi, M., 2018. Functional data analysis: estimation of the relative error in functional regression under random left-truncation model. Journal of Nonparametric Statistics, 30(2), pp.472-490.

18 Chikr Elmezouar, Z., Alshahrani, F., Almanjahie, I.M., Kaid, Z., Laksaci, A. and Rachdi, M., 2023. Scalar-on-Function Relative Error Regression for Weak Dependent Case. Axioms, 12(7), p.613.

19 Benzamouche, S., Ould Saïd, E. and Sadki, O., 2025. Nonparametric estimation of the relative error regression for twice censored and dependent data. Communications in Statistics-Theory and Methods, 54(5), pp.1492-1525.

20 Xia, X., Ming, H. and Li, J., 2026. Relative error model average for multiplicative models. Statistics and Computing, 36(1), p.18.

21 Wang, J.L., Chiou, J.M. and Müller, H.G., 2016. Functional data analysis. Annual Review of Statistics and its application, 3(1), pp.257-295.

22 Aneiros, G., Bongiorno, E.G., Goia, A. and Hušková, M. eds., 2025. New Trends in Functional Statistics and Related Fields. Springer Nature..

23 Ndiaye, M., Dabo-Niang, S., Ngom, P., Thiam, N., Brehmer, P. and El Vally, Y., 2024, May. Nonparametric prediction and supervised classification for spatial dependent functional data under fixed sampling design. In Nonlinear Analysis, Geometry and Applications: Proceedings of the Third NLAGA-BIRS Symposium, AIMS-Mbour, Senegal, August 21–27, 2023 (pp. 69-100). Cham: Springer Nature Switzerland.

24 Ferraty, F. and Vieu, P., 2006. Nonparametric functional data analysis: theory and practice. New York, NY: Springer New York.

25 Rachdi, M. and Vieu, P., 2007. Nonparametric regression for functional data: automatic smoothing parameter selection. Journal of statistical planning and inference, 137(9), pp.2784-2801.

26 Benaglia, T., Chauveau, D., Hunter, D.R. and Young, D.S., 2010. mixtools: an R package for analyzing mixture models. Journal of statistical software, 32, pp.1-29.

27 Huber, P. J., and Ronchetti, E. M. (2009). Robust Statistics (2nd ed.). John Wiley & Sons.

28 Clark, C. J., et al. 2003. Dry matter and soluble solids in onion (Allium cepa L.): use of near-infrared spectroscopy as a screening tool. Journal of the Science of Food and Agriculture, 83(5), 371-379.

29 Koenker, R. and Zhao, Q., 1996. Conditional quantile estimation and inference for ARCH models. Econometric theory, 12(5), pp.793-813.