

Article

Nonparametric Analysis of Functional Time Series Data Using Least Absolute Relative Error Regression

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

¹ Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

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

* Correspondence: faalmulhim@pnu.edu.sa

Abstract

In this paper, we introduce a novel kernel-based estimator for the regression operator of a scalar response variable $\mathcal{R}$ given a functional covariate $\mathcal{F}$ taking values in a semi-metric space. The estimator is constructed through the minimization of the least absolute relative error (LARE) criterion, which provides an invariant scale and more balanced measure of predictive performance than conventional squared error methods. By focusing on relative deviations, the LARE approach effectively reduces the influence of extreme response values and enhances robustness in the presence of heteroscedasticity. From a theoretical point of view, we investigate the asymptotic behavior of the proposed estimator under strong mixing conditions for functional time series data. We show that, despite the temporal dependence structure, the estimator remains consistent and achieves convergence rates comparable to those obtained under independence. In the computational part, we show that the proposed method is computationally efficient and straightforward to implement. Its empirical performance is evaluated through simulation studies conducted under different dependence scenarios. In addition, the applicability of the method is illustrated through the analysis of a real data set.

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

MSC: 62R20; 62G05; 62G08



Academic Editors: JiuJun Zhang and Xuejun Wang

Received: 14 April 2026

Revised: 19 May 2026

Accepted: 21 May 2026

Published: 25 May 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Investigating the relationship between a functional predictor and a scalar response variable is a common issue in functional data analysis. This relationship is most often studied through regression models, often estimated via the least squares criterion. However, this classical approach can be influenced by scale variability and atypical observations, which may affect its stability and accuracy. In this paper, we propose an alternative smoothing framework based on the least absolute relative error (LARE). By working with relative deviations rather than least square error, the LARE criterion yields a more robust estimation procedure and generally leads to improved predictive performance compared with standard regression methods.

The general setting of this paper concerns nonparametric prediction for functional data. It constitutes an active research area that has gained substantial progress in recent years. The pioneering contribution in this direction was made by [1], who established

the almost complete pointwise consistency of kernel-based regression estimators under the assumption of independent and identically distributed (i.i.d.) observations. Next, ref. [2] derived convergence results in the (L^p) norm, while [3] established the moment integrability of the same estimator in the same independence framework. The asymptotic normality of the kernel estimator under strong mixing dependence conditions was proven in [4]. Ref. [5] considers the k-nearest neighbor (kNN) estimators for a nonparametric functional regression model under negatively associated dependence. They proved the almost complete consistency. More recently, ref. [6] addresses a more general setting for functional covariates by assuming that the regressors take values in Riemannian manifolds. Under a strong mixing dependence framework, explicit expressions for the bias and variance of the proposed estimator are established. As an alternative to the classical kernel approaches, more recent studies have explored (M)-estimation techniques as well as local linear smoothing methods. A comprehensive overview of nonparametric functional prediction can be found in [7,8], which provide detailed surveys of these developments.

In contrast to these classical approaches, we focus on relative error (RE) regression, where the loss function is defined through relative deviations rather than conventional scoring functions. Such a formulation is particularly attractive in applications where scale invariance plays an important role. Despite its practical appeal, the development of nonparametric RE regression remains relatively limited, as most existing contributions have been devoted to parametric or semiparametric settings. One of the earliest contributions in this direction is given by [8], who investigated estimation procedures based on the relative squared error loss. From the applied studies, RE regression has been successfully employed in a variety of fields, including medical studies [9] and financial modeling [10]. Furthermore, ref. [11] examined RE-based estimation within multiplicative regression models, while ref. [12] extended the methodology to the nonparametric framework and established consistency and convergence results for local linear estimators under relative error loss.

In recent years, increasing attention has been given to RE regression in the presence of dependent data structures. For instance, ref. [13] studied RE-based estimation for spatial time series models. The first nonparametric functional RE regression framework was introduced by [14], who established strong consistency and derived the asymptotic distribution of the associated estimator. This result was extended in [15], where a kernel-based estimator was developed for truncated and incomplete functional data. Additional recent developments and methodological advances in this area can be found in [16,17].

While all the aforementioned studies estimate the regression operator using either the least squares (LS) or least squares relative error (LSRE) loss function, the present work adopts the absolute relative error loss within a kernel local weighting framework to construct a robust estimator for nonparametric functional regression. The resulting local absolute relative error (LARE) approach extends the classical LSRE methodology developed in [15,16] by providing enhanced robustness and improved estimation accuracy. These features are particularly important in functional time series settings, where the data are often contaminated by noise and characterized by high-frequency fluctuations. In such situations, extreme observations and local irregularities occur frequently and may substantially affect the quality of estimation. Consequently, the proposed estimator yields more stable and reliable inference for the underlying functional relationship. The robustness of the proposed LARE framework makes it especially suitable for high-frequency functional data and supports its applicability in a broad range of fields, including economics, finance, signal processing, and environmental sciences. The present contribution addresses both theoretical and practical aspects of the problem. From a theoretical perspective, we show that the nonparametric analysis of the LARE regression model is considerably more

challenging than that of the LSRE framework. In particular, unlike the LSRE estimator, which admits an explicit representation in terms of ratios involving the first inverse conditional moment functionals, the LARE estimator does not possess a closed-form expression. As a consequence, its asymptotic properties cannot be derived through direct analytical arguments and instead require the development of a refined Bahadur-type representation.

Despite these technical difficulties, we establish the almost complete convergence of the proposed estimator and derive its rate of convergence under mild and standard assumptions commonly used in nonparametric functional regression. In particular, the assumptions accommodate both the functional nature of the covariates and the strong mixing dependence structure of the underlying functional time series. Such strong mixing frameworks provide a realistic setting for many modern applications in which temporal dependence plays a fundamental role, thereby making robust predictive procedures essential for reliable inference and forecasting in the presence of dependence and data contamination. The obtained results extend recent findings of [18] from the independent framework to the dependent functional time series setting. Finally, the practical relevance of the proposed methodology is illustrated through extensive simulation studies and real-data applications, both of which demonstrate the robustness and strong empirical performance of the proposed estimator.

The remainder of this paper is organized as follows. Section 2 introduces the functional time series frame and the proposed estimation methodology. Section 3 is devoted to the underlying assumptions and the establishment of the main asymptotic results. In Section 4, we discuss the applicability of the estimator in the functional prediction. Section 5 presents a comprehensive empirical assessment of the proposed estimator through simulation experiments and real data applications. Finally, Section 7 concludes the paper with some final remarks and perspectives for future research. The technical proofs are collected in Appendix A.

2. Nonparametric Functional Data Analysis

2.1. Data

The principal contribution of this work is the development of a more general framework that relaxes the conventional independence assumption, which is rarely satisfied in real-world applications. Even when the underlying dependence is weak, neglecting such structures can result in biased estimates and potentially invalid statistical inference. In this context, modelling functional time series data provides a more realistic and theoretically sound alternative, as it explicitly take into account the temporal dependence. Consequently, the proposed approach offers a more meaningful and reliable framework than methods based on independent observations. It is worth noting that functional time series (FTS) have attracted considerable attention in the functional statistics literature. Comprehensive treatments of both theoretical developments and practical applications can be found in the monographs by [19–22].

To fix ideas, we begin by recalling the definition of the strong mixing property. For this purpose, we introduce the following notation. Let $\mathcal{A}_i^k$ denote the σ -algebra generated by the collection $\{\mathcal{X}_j : i \leq j \leq k\}$. The strong mixing coefficient, originally introduced by [23], is defined as follows; see also [24] for a comprehensive treatment.

Definition 1. Let $X = \{\mathcal{X}_i : i = 1, 2, \dots\}$ be a strictly stationary sequence of random variables. For any positive integer n , define

$$\alpha(n) = \sup_{k \in \mathbb{N}^*} \sup \left\{ |\mathbb{P}(\mathcal{A} \cap \mathcal{B}) - \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})| : \mathcal{A} \in \mathcal{A}_1^k, \mathcal{B} \in \mathcal{A}_{k+n}^\infty \right\}.$$

The sequence X is said to be α -mixing (or strongly mixing) if $\alpha(n) \rightarrow 0$ as $n \rightarrow \infty$.

In this study, our analysis is based on data drawn from a sequence $X = \{(\mathcal{F}_i, \mathcal{R}_i)\}_{i \geq 1}$ of strong functional time series observations, assumed to be identically distributed as the pair $(\mathcal{F}, \mathcal{R})$. Specifically, the random pair $(\mathcal{F}, \mathcal{R})$ is assumed to be valued in the product space $\mathfrak{S} \times \mathbb{R}$, where $\mathfrak{S}$ represents a functional space equipped with a suitable semi-metric D .

2.2. Nonparametric Model

As stated in the Introduction, the main aim of this work is to analyze the relationship between a functional explanatory variable $\mathcal{F}$ and a scalar response variable $\mathcal{R}$. For this aim, we fix a point $F \in \mathbb{F}$ and consider a neighborhood $\mathcal{V}_F$ of F in the functional space $\mathbb{F}$. Often, the dependence between $\mathcal{F}$ and $\mathcal{R}$ was described using a least squares regression framework, defined by

$$NR(F) = \arg \min_r \mathbb{E} \left[(\mathcal{R} - r)^2 \middle| \mathcal{F} = F \right].$$

It is clear that the main advantage of this loss function is its analytical simplicity, as it allows us to derive an explicit expression of the predictor. Indeed, by taking the derivative of the criterion with respect to s , we obtain

$$NR(F) = \mathbb{E}[\mathcal{R} \mid \mathcal{F} = F]. \quad (1)$$

Despite these advantages, standard regression methods have certain limitations. In particular, they are often sensitive to outliers, may perform poorly under model misspecification, and can be unstable in the presence of multicollinearity or high-dimensional data structures. To remedy this drawback, we consider the least absolute relative error (LARE) criterion, which yields a more robust and equilibrating treatment allowing us to reduce the influence of extreme observations. Specifically, the LARE regression is defined as

$$RE(F) = \arg \min_r \mathbb{E} \left[\left| \frac{\mathcal{R} - r}{\mathcal{R}} \right| \middle| \mathcal{F} = F \right]. \quad (2)$$

In contrast to the quadratic loss in (1), this formulation is based on an absolute loss function, which improves robustness by penalizing deviations in a linear and more uniform manner. Consequently, the LARE criterion leads to a more stable and dependable evaluation of predictive performance, particularly when the data exhibit heavy tails or contain outlying observations.

For the estimation step, we use the strong functional time series observations $X = \{(\mathcal{F}_i, \mathcal{R}_i)\}_{i \geq 1}$ to construct a kernel-type estimator of $RE(F)$, defined as follows:

$$\widehat{RE}(F) = \arg \min_r \frac{\sum_{i=1}^n \mathbf{Q}(f^{-1}D(F, \mathcal{F}_i)) \left| \frac{\mathcal{R}_i - r}{\mathcal{R}_i} \right|}{\sum_{i=1}^n \mathbf{Q}(f^{-1}D(F, \mathcal{F}_i))}, \quad (3)$$

where $\mathbf{Q}$ is a kernel function and $f := f_n$ is a sequence of positive real numbers tending to zero as $n \rightarrow \infty$.

The algorithm for computing this estimator is described below.

1. **Input functional sample:** $X = \{(\mathcal{F}_i, \mathcal{R}_i)\}_{i=1}^n$, target curve F , kernel $\mathbf{Q}(\cdot)$, bandwidth $f_n > 0$.

2. **Compute functional distances:** For each observation $\mathcal{F}_i$, compute the semi-metric (or distance)

$$D(F, \mathcal{F}_i), \quad i = 1, \dots, n,$$

which measures the proximity between the target function F and each sample curve $\mathcal{F}_i$.

3. **Construct kernel weights:** Transform the distances into normalized kernel weights using the bandwidth f_n :

$$w_i(F) = \mathbf{Q}\left(\frac{D(F, \mathcal{F}_i)}{f_n}\right), \quad i = 1, \dots, n.$$

These weights assign larger importance to observations whose functional predictors are closer to F .

4. **Define and minimize the weighted LARE criterion:** For any candidate value r , compute the weighted loss function:

$$L_n(r; F) = \frac{\sum_{i=1}^n w_i(F) \left| \frac{\mathcal{R}_i - r}{\mathcal{R}_i} \right|}{\sum_{i=1}^n w_i(F)}.$$

Obtain the estimator by solving the one-dimensional optimization problem:

$$\widehat{RE}(F) = \arg \min_r L_n(r; F).$$

In practice, this step can be implemented using a grid search or a convex optimization routine.

5. **Output the estimator:** Return the minimizer $\widehat{RE}(F)$ as the estimated relative error regression function at F .

We point out that, in this contribution, we assume the uniqueness of the parameter $RF(F)$, whereas the uniqueness of the estimator is not necessarily required. Thus, $\widehat{RE}(F)$ denotes any solution of (3). Recall that the primary aim of the theoretical part of this paper is to investigate the asymptotic behaviour of the estimator $\widehat{RE}(F)$ of $RE(F)$, in the setting of functional time series case. To the best of our knowledge, this work provides the first construction and theoretical analysis of such an estimator within this general framework, including the multivariate case as a particular situation. Although the finite-dimensional setting can be viewed as a special case of the functional framework developed here, the infinite-dimensional context is especially relevant due to its increased mathematical complexity.

3. Main Asymptotic Result

Now, to state the almost complete convergence, we introduce generic strictly positive constants denoted by C or C' . We also define

$$\mathbf{Q}_i = \mathbf{Q}(f^{-1}D(F, \mathcal{F}_i)), \quad i = 1, \dots, n.$$

The closed ball of radius f centered at F is given by

$$B(F, f) = \{\mathcal{F} \in \mathbb{F} : D(\mathcal{F}, F) < f\},$$

and we define

$$U_1(r, F) = \mathbb{E}\left[|\mathcal{R}|^{-1} \mathbf{1}_{\{\mathcal{R} \leq r\}} \mid T = F\right], \quad U_2(r, F) = \mathbb{E}\left[|\mathcal{R}|^{-1} \mathbf{1}_{\{\mathcal{R} \geq r\}} \mid T = F\right].$$

We now state the following assumptions:

(RE1) The small-ball probability satisfies

$$\mathbb{P}(T \in B(F, f)) =: \mathcal{P}_F(f) > 0, \quad \forall f > 0,$$

such that

$$\lim_{f \rightarrow 0} \mathcal{Q}_f \mathcal{P}_F(f) = 0.$$

(RE2) For any $F_1, F_2 \in \mathcal{V}_F$, the functions U_γ are continuously differentiable with respect to r and satisfy

$$|U_\gamma(r, F_1) - U_\gamma(r, F_2)| \leq C d^{b_\gamma}(F_1, F_2), \quad b_\gamma > 0, \quad \gamma = 1, 2.$$

(RE3) The response variable admits bounded inverse moments: for all $p \geq 2$,

$$\mathbb{E}[|\mathcal{R}|^{-p} \mid T = F] \leq C(p, F) < \infty \quad \text{almost surely.}$$

(RE4) The sequence $(\mathcal{F}_i, \mathcal{R}_i)_{i \in \mathbb{N}}$ satisfies: $\exists a > 0, \exists c > 0 : \forall n \in \mathbb{N}, \alpha(n) \leq cn^{-a}$ and

$$\begin{cases} \forall i \neq j, & \mathbb{E}[\mathcal{R}_i \mathcal{R}_j \mid \mathcal{F}_i, \mathcal{F}_j] \leq C < \infty, \quad \text{a.s.} \\ \mathbb{P}((\mathcal{F}_i, \mathcal{F}_j) \in B(x, f) \times B(x, f)) = \mathcal{Q}_F(f) > 0; \end{cases}$$

(RE5) The kernel function $\mathbf{Q}$ is measurable, supported on $(0, 1)$, and uniformly bounded, i.e.,

$$0 < C_2 \leq \mathbf{Q}(\cdot) \leq C_3 < \infty.$$

(RE6) There exists $\eta > 0$, such that

$$C n^{\frac{2(2-a)}{(a+1)} + \eta} \leq \Phi_F(f) \leq n^{\frac{2a}{(2-a)p}},$$

$$\text{where } \Phi_F(f) = \max(\mathcal{P}_F^2(f), \mathcal{Q}_F(f)) \text{ and } a > \max\left(2 + 1/2b_2, p/2b_2, \frac{4p+1+\sqrt{24p+1}}{p-1}\right).$$

Discussion of the Assumptions

In general, these assumptions are mild and standard in the literature on functional time series analysis. The theoretical framework for nonparametric functional data analysis is essentially built upon two main components. The first, **(RE1)**, concerns the small-ball probability function, which characterizes the concentration properties of the probability measure induced by the functional covariate. This assumption is fundamental, as it controls the stochastic fluctuations of the estimator and plays a pivotal role in the variance component of the convergence rate. The second component, **(RE2)**, specifies the regularity of the nonparametric regression function. It is a classical smoothness assumption in functional nonparametric statistics and has been widely adopted in related works. This condition is mainly used to control the bias term in the asymptotic expansion. Assumption **(RE3)** is a well-known moment condition in regression analysis and is commonly employed to establish almost sure (or almost complete) consistency; see, for example, [1]. It is worth noting that **(RE3)** replaces the finite-variance condition typically used to ensure convergence in probability. Since almost complete convergence is a stronger form of convergence, it requires more restrictive moment assumptions. This reflects the standard trade-off between generality and strength: sharper asymptotic results generally come at the cost of stronger conditions.

Overall, the set of assumptions in this work is designed to capture three structural aspects of the model. The functional nature of the data is addressed through the concentration condition, the nonparametric structure is characterized via a Hölder-type regularity assumption, and the dependence structure is controlled by (RE4), which accounts for both global dependence across observations and local dependence within the functional trajectories. The remaining assumptions, (RE5)–(RE6), are mainly technical in nature; they are introduced to facilitate the proofs and to derive the convergence rate of the proposed estimator.

Theorem 1. *Given assumptions (RE1)–(RE6), we obtain*

$$|\widehat{RE}(F) - RE(F)| = O(f_n^{b_1}) + O(f_n^{b_2}) O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n \mathcal{P}_F^2(f)}}\right), \quad a.co.. \quad (4)$$

4. Application to Prediction in Continuous Time Process

Clearly, a particular important case of functional data is when observations are generated by a continuous-time stochastic process. In this setting, the proposed methodology can be effectively employed to predict a future real-valued feature of the process by incorporating information from its entire past trajectory in a continuous manner. Indeed, let P_t denote a continuous time process observed in the interval $(0, b)$, and our goal is to predict its unobserved value at time b using the LARE approach. To this purpose, we define N functional observations $(\mathcal{F}_i)_{i=1,\dots,N}$ by

$$\forall t \in [0, b), \quad \mathcal{F}_i(t) = P_{((i-1)b+t)/N}.$$

The associated scalar responses are defined as

$$\mathcal{R}_i = P_{(ib)/N}.$$

Consequently, the kernel-based estimator of the LARE regression $\widehat{RE}(\mathcal{F}_N)$ is constructed from the sample $(\mathcal{F}_i, \mathcal{R}_i)_{i=1,\dots,N-1}$. The asymptotic properties of this estimator are summarized in the following corollary.

Corollary 1. *Under the conditions of Theorem 1, and for sufficiently large N , we obtain*

$$\widehat{RE}(\mathcal{F}_N) - RE(\mathcal{F}_N) \rightarrow 0 \quad a.co..$$

5. Computational Aspects

A Simulated Data-Driven Model

The purpose of this section is to emphasize the importance of LARE regression as a robust predictive tool. Recall that LARE regression provides a flexible, nonparametric approach to prediction, allowing to identify the complex and potentially nonlinear relationships that may be overlooked by standard regression models. Despite these advantages, the LARE model remains relatively not fully explored compared with conventional regression techniques. In many applied fields, researchers and practitioners consider the standard regression as a primary analytical tool for modeling and prediction. While widely used, standard regression often depends on restrictive requirements, namely, sensitivity to outliers or heteroscedasticity, which can limit its effectiveness in complex settings. In contrast, LARE regression offers a more robust and adaptable framework, particularly in situations characterized by asymmetry, heterogeneity, or heavy-tailed behavior. Therefore, LARE regression constitutes a powerful alternative to standard regression, providing enhanced

predictive performance and greater flexibility. This illustration is designed to address three primary objectives, which are the simplicity of implementation, the impact of correlation on predictive performance and the identification of the main characteristics of the models. For these aims, we apply the algorithm of Section 4 on the dynamic of a real time series, which is generated by GARCH model, GARCH(1,1). It is well known that this class of processes satisfies the strong mixing assumption (see [25]), which provides a suitable theoretical foundation for the developments presented in this contribution. Formally, in this illustration, we generate a GARCH process P_t as follows:

$$\begin{cases} P_t = \mu + \epsilon_t, & \text{where } \epsilon_t = \sigma_t Z_t, \\ \sigma_t^2 = w + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \end{cases}$$

where the parameters ω , α , and β satisfy the conditions $\omega > 0$ and $\alpha + \beta < 1$, ensuring the stationarity of the process. In order to covers different cases we simulate three distributions of the innovations Z_t that are Gaussian, Student-t distribution and Skewed Student-t distribution. Such cases explore characteristics sch the heavy-tailed distribution as well as the symmetry or asymmetry properties. The different generated GARCH process using the routine function *ugarchpath* of the R-package *rugarch* for R version 4.4.3. The trajectories of the simulated process are displayed in Figure 1.

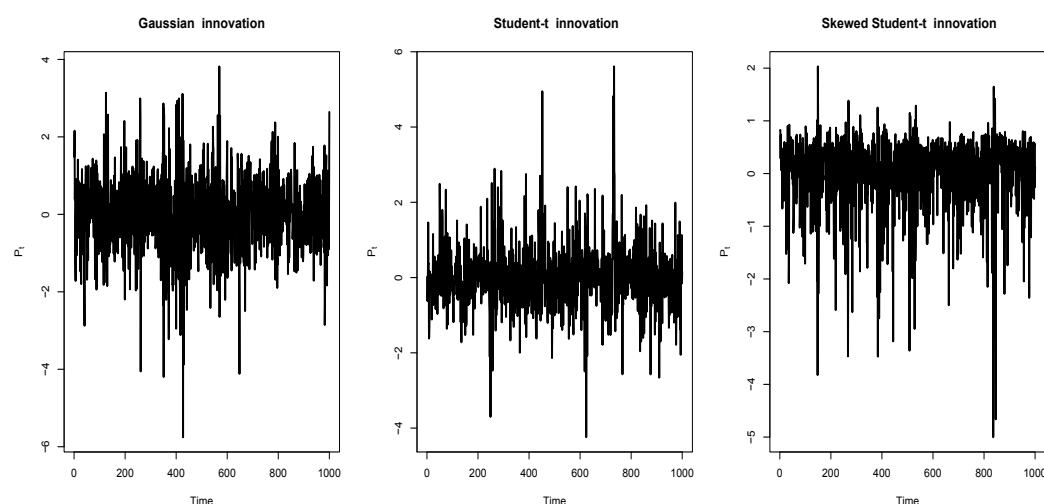


Figure 1. The GARCH process, $\mu = 0, w = 0.1, \alpha = 0.2, \beta = 0.7$.

The conditional volatility for the three cases are plotted in Figure 2.

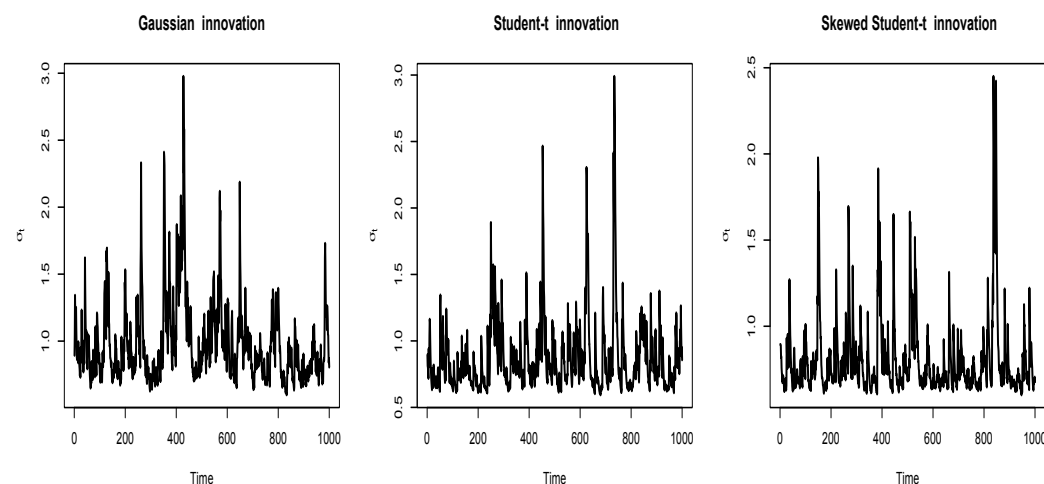


Figure 2. The σ_t process, $\mu = 0, w = 0.1, \alpha = 0.2, \beta = 0.7$.

Now, in order to highlight the relevance of the nonparametric functional approach, we compare the estimator $\widehat{RE}$ with the standard regression defined by

$$\widetilde{RE} = \frac{\sum_{i=1}^n \mathbf{Q}(f^{-1}D(F, \mathcal{F}_i)) \mathcal{R}_i}{\sum_{i=1}^n \mathbf{Q}(f^{-1}D(F, \mathcal{F}_i))}.$$

We point out that, this simulation study is conducted using a sample of 1000 observations of the pair $(\mathcal{R}_i = P_{(ib)/N}, \mathcal{F}_i = P_{s \in ((ib)/N, ((i-1)b)/N}))$, which is divided into two subsets: the first 800 observations are used for in-sample analysis, while the remaining 200 observations are reserved for out-of-sample evaluation. Regarding the choice of the parameters involved in estimator $\widehat{RE}$ and $\widetilde{RE}$, we employ a quadratic kernel defined on $(0, 1)$, along with a PCA-based metric, which is particularly suitable for handling discontinuous functional data. The optimal bandwidth is selected locally via a cross-validation procedure as follows

$$f_{\text{opt}}(F) = \arg \min_{f \in H_n(F)} \sum_{i=1}^n |\mathcal{R}_i - \widehat{RE}^{-l}(\mathcal{F}_i)|,$$

where $H_n(F)$ denotes the set of positive real values f_F such that the ball centered at F with radius f_F contains exactly $l \in \{5, 10, 15, \dots, n/2\}$ nearest neighbors of F . This bandwidth selection strategy is standard in nonparametric functional statistics; see [1] for a comprehensive discussion. Furthermore, to evaluate the feasibility of the LARE regression, as well as the effect of correlation, we compare its performance with that of standard nonparametric $\widetilde{RE}$ approaches using different values w, α, b, β . Varying these parameters allows us to quantify the effect of correlation on the performance of the two estimators $\widehat{RE}$ and $\widetilde{RE}$. The mean absolute error (MAE) is used to assess the performance of both estimators.

$$MAE = \frac{1}{200} \sum_{j \text{ in the testing sample}} |\mathcal{R}_j - \text{Pred}(\mathcal{F}_i)|$$

where Pred means either $\widehat{RE}$ or $\widetilde{RE}$. The MAEs are reported in Table 1.

Table 1. MAE results under different scenarios.

The Z's Distribution	The Lag b	The Parameters (w, α, β)	$\widehat{RE}$	$\widetilde{RE}$
N (0,1)	5	(0.1, 0.2, 0.7)	0.024	0.032
		(0.1, 0.05, 0.2)	0.014	0.028
		(0.2, 0.05, 0.2)	0.018	0.025
		(0.1, 0.05, 0.8)	0.027	0.034
		(0.1, 0.5, 0.2)	0.025	0.033
		(0.2, 0.5, 0.8)	0.042	0.038
	15	(0.1, 0.2, 0.7)	0.028	0.035
		(0.1, 0.05, 0.2)	0.019	0.032
		(0.2, 0.05, 0.2)	0.023	0.025
		(0.1, 0.05, 0.8)	0.027	0.038
		(0.1, 0.5, 0.2)	0.029	0.038
		(0.2, 0.5, 0.8)	0.052	0.044
	30	(0.1, 0.2, 0.7)	0.033	0.038
		(0.1, 0.05, 0.2)	0.024	0.036
		(0.2, 0.05, 0.2)	0.028	0.029
		(0.1, 0.05, 0.8)	0.031	0.040
		(0.1, 0.5, 0.2)	0.037	0.046
		(0.2, 0.5, 0.8)	0.058	0.053

Table 1. Cont.

The Z's Distribution	The Lag b	The Parameters (w, α, β)	$\widehat{RE}$	$\widehat{RE}$
STD (shape = 0.4)	5	(0.1, 0.2, 0.7)	0.047	0.036
		(0.1, 0.05, 0.2)	0.041	0.023
		(0.2, 0.05, 0.2)	0.049	0.028
		(0.1, 0.05, 0.8)	0.052	0.037
		(0.1, 0.5, 0.2)	0.068	0.041
		(0.2, 0.5, 0.8)	0.086	0.030
	15	(0.1, 0.2, 0.7)	0.095	0.039
		(0.1, 0.05, 0.2)	0.061	0.037
		(0.2, 0.05, 0.2)	0.075	0.041
		(0.1, 0.05, 0.8)	0.122	0.042
		(0.1, 0.5, 0.2)	0.127	0.041
		(0.2, 0.5, 0.8)	0.182	0.052
	30	(0.1, 0.2, 0.7)	0.121	0.041
		(0.1, 0.05, 0.2)	0.123	0.045
		(0.2, 0.05, 0.2)	0.157	0.052
		(0.1, 0.05, 0.8)	0.194	0.054
		(0.1, 0.5, 0.2)	0.207	0.061
		(0.2, 0.5, 0.8)	0.256	0.072
SSTD (shape = 0.4, skew = 0.2)	5	(0.1, 0.2, 0.7)	0.089	0.042
		(0.1, 0.05, 0.2)	0.062	0.031
		(0.2, 0.05, 0.2)	0.155	0.039
		(0.1, 0.05, 0.8)	0.179	0.042
		(0.1, 0.5, 0.2)	0.207	0.046
		(0.2, 0.5, 0.8)	0.286	0.051
	15	(0.1, 0.2, 0.7)	0.398	0.054
		(0.1, 0.05, 0.2)	0.471	0.061
		(0.2, 0.05, 0.2)	0.494	0.065
		(0.1, 0.05, 0.8)	0.538	0.071
		(0.1, 0.5, 0.2)	0.587	0.079
		(0.2, 0.5, 0.8)	0.622	0.052
	30	(0.1, 0.2, 0.7)	0.657	0.062
		(0.1, 0.05, 0.2)	0.695	0.071
		(0.2, 0.05, 0.2)	0.729	0.078
		(0.1, 0.05, 0.8)	0.784	0.074
		(0.1, 0.5, 0.2)	0.827	0.084
		(0.2, 0.5, 0.8)	0.956	0.087

The results of this table indicate that the nonparametric functional LARE regression estimator achieves higher accuracy compared with the standard nonparametric approach. The effect of correlation on the estimation performance is important, in the sense that the MAE values increase under highly volatile settings, particularly when $\alpha + \beta$ is close to 1. However, it is clear that the LARE improves performance under heavy-tailed or skewed contamination, but it may lose efficiency under Gaussian/light-tailed settings. Overall, the main strength of the nonparametric functional LARE estimator is reflected in its flexibility, ease of implementation, and reduced sensitivity to small deviations induced by the dependence structure.

In the second part of this illustration, we investigate the robustness of the proposed estimator $\widehat{RE}$. Of course, robustness may be evaluated through several well-established criteria, including the influence function, gross-error sensitivity, breakdown point, local shift sensitivity, and bias–variance behavior under contamination, among others (see [26])

for a comprehensive review). Although some of these concepts are not straightforward to generalize to functional settings, they all share the same goal of measuring the stability of an estimator and its sensitivity to different deviation in the appropriate situation. For this study, we consider the standard data (corresponds to an innovation distributed by the Standard Gaussian distribution, the lag = 30 and ($w = 0.1, \alpha = 0.2, \beta = 0.7$)). Now, we assess the robustness of the proposed method through a contamination stress test. This approach is designed to examine how the estimator behaves when the observed data are perturbed by introducing controlled levels of contamination. In particular, the stress test provides a practical framework for evaluating the sensitivity and stability of the estimator under progressively increasing contamination proportions. Specifically, we investigate the estimator's robustness by varying the contamination level and measuring its performance under each scenario. The following algorithm describes the detailed procedure used in this part.

Step 1 Define large scale C

Step 2 For given percentage $m\%$ we create a contaminated sample $\mathcal{X}_C$ by multiplying $m\%$ response values $\mathcal{R}_i$ by C

Step 3 Split the sample $\mathcal{X}_C$ into part 75% of the learning sample and 25% for the testing sample

Step 4 Compute $\hat{\theta}_{m,C} = \text{either } \widehat{RE} \text{ or } \widetilde{RE}$ for the testing sample using the learning sample of $\mathcal{X}_C$

Step 5 Calculate MAE for both estimators using the testing sample

Table 2 reports the MAE for various values of C , m and n .

Table 2. MAE values under different contamination levels.

n	m%	C	$\widetilde{RE}$	$\widehat{RE}$
200	10%	5	0.032	0.095
200	15%	5	0.043	0.758
200	20%	5	0.048	0.945
200	10%	10	0.063	1.789
200	15%	10	0.072	2.172
200	20%	10	0.081	2.089
300	10%	5	0.029	0.089
300	15%	5	0.038	0.142
300	20%	5	0.042	0.437
300	10%	10	0.059	0.782
300	15%	10	0.069	1.863
300	20%	10	0.072	1.967
500	10%	5	0.026	0.086
500	15%	5	0.026	0.119
500	20%	5	0.039	0.231
500	10%	10	0.054	0.564
500	15%	10	0.067	1.557
500	20%	10	0.070	1.742

As expected, the results confirm the superior robustness of $\widehat{RE}$ compared with $\widetilde{RE}$. In particular, the MAE associated with $\widehat{RE}$ exhibits low variability across different contamination levels, indicating that the LARE regression estimator is both stable and highly robust. Specifically, its MAE ranges from 0.026 to 0.081, reflecting only mild sensitivity to increasing contamination. By contrast, the estimator $\widetilde{RE}$ shows substantially higher variability, with MAE values ranging from 0.086 to 2.089, which indicates a considerable degradation in

performance under contamination. These results clearly highlight the robustness advantage of the proposed LARE approach through its consistently low error variability.

6. Real Data Application

This section introduces a new predictive approach for financial time series using the studied LARE regression, with particular attention to its feasibility in volatility forecasting and option pricing. We perform a comparative analysis of several regression-based predictors designed to enhance the empirical prediction of asset returns, which play a central role in derivative valuation. This framework offers considerable flexibility for exploring the complex and the irregular dynamics that characterize financial markets. Therefore, the robustness of the studied model is particularly important in option pricing, where an accurate representation of tail behavior and volatility dynamics is crucial for reliable valuation.

For the empirical analysis, we consider closed-prices (of frame of 5 min) of Bitcoin Cryptocurrency. The data was retrieved from the Stooq database (available in <https://stooq.com/db/h/>, accessed on 18 March 2026) and it covers the period from 31 October 2025 to 15 March 2026. The resulting time series displays clear evidence of volatility clustering, rapid fluctuations, and strong temporal dependence features that present substantial challenges for financial models. These properties highlight the practical relevance of robust and nonparametric approaches for both volatility forecasting and derivative pricing. In practice, financial analysis is generally performed using log-returns rather than raw prices. Accordingly, we define the log-return process as

$$P(t) := 100 \left(\log(Z(t)) - \log(Z(t-1)) \right),$$

where $Z(t)$ represents the asset price at time t . The resulting series $(Z(t), P(t))$ are illustrated in Figures 3 and 4.

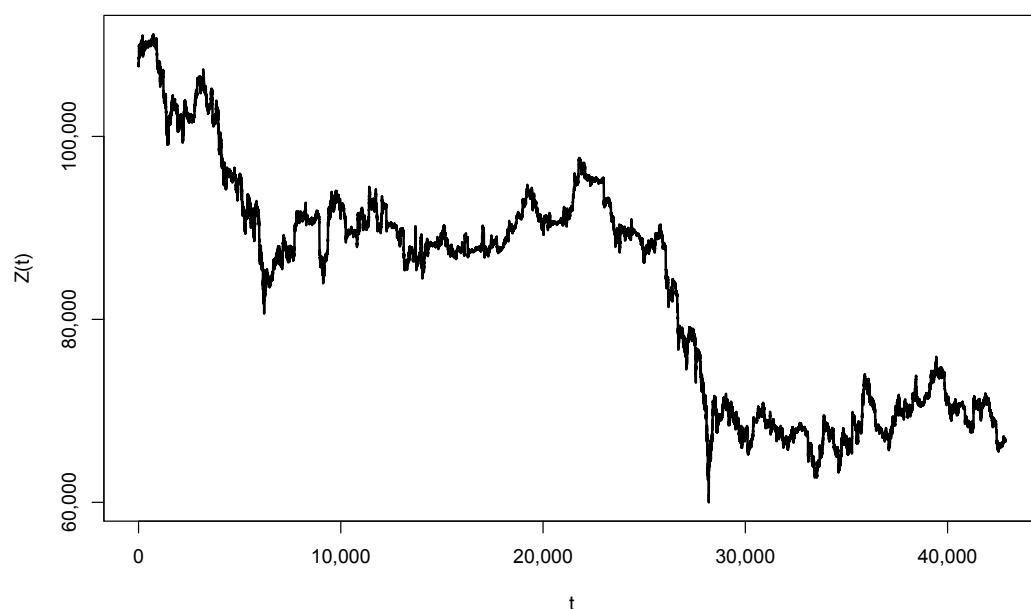


Figure 3. The $Z(t)$ -process.

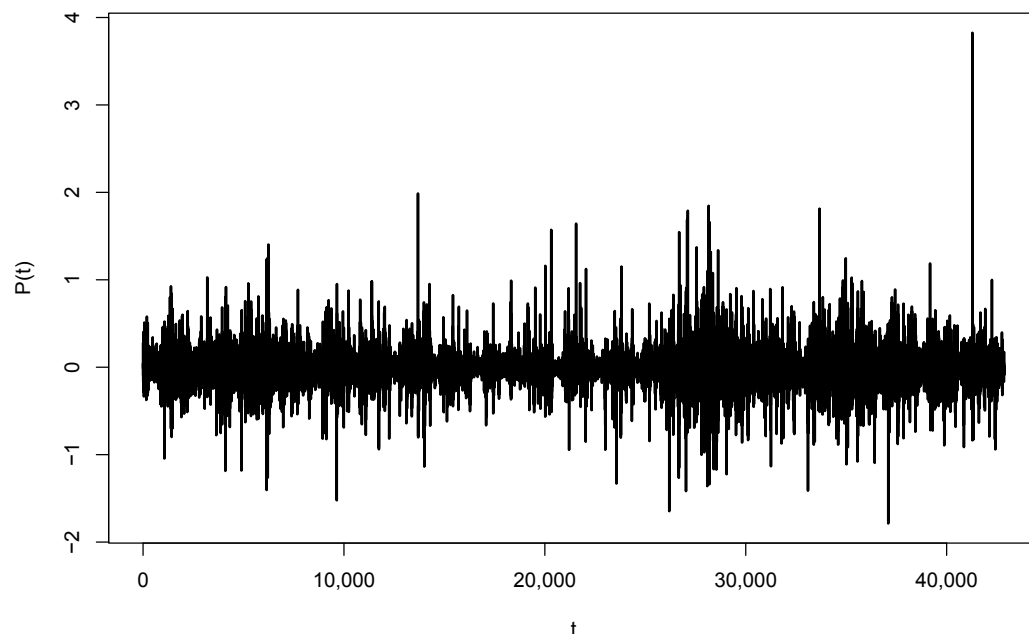


Figure 4. The $P(t)$ -process.

It is evident that the series $P(t)$ is affected by certain outliers. To quantify these observations, we employ the median absolute deviation (MAD) function. Using this procedure, a total of 11.71% observations are considered as outliers in the dataset. Our main objective is to forecast one-day ahead, which motivates the examination of the stochastic process $\{P(s)\}_{s \in [0, T]}$. It is well established that financial return series typically display zero mean with substantial volatility. These two features highlight the nature dynamic of the asset returns. According to the algorithm of the functional time series data developed in Section 4, we construct a functional sample by cutting the series $\{P(s)\}$ into $N = 135$ curves, each representing one day of observations.

Although our primary goal is to predict the total day return, we further assess the performance of the proposed method by comparing it with a classical regression estimator based on the conditional expectation, denoted by $\widetilde{RE}$. To illustrate the added value of the LARE-regression approach, the prediction analysis is carried out under two scenarios: in the first, the analysis is performed on the original data $P(T)$ containing outliers, while in the second, outliers are removed. Keeping the same strategy of the previous section to select the principal parameters including the smoothing parameter, the metric D and the same kernel $\mathbf{Q}$ we obtain the following results (see Figures 5 and 6). In these figures we plot predicted values against the true values in both case (initial data and clean data).

The empirical results demonstrate that the LARE-regression estimator, $\widehat{RE}$, delivers more accurate predictions than the conventional least-squares estimator, $\overline{RE}$. This superiority is confirmed through the comparison of the mean absolute error (MAE) obtained under the two estimation procedures across both data configurations. For the original dataset, where outliers are retained, the MAE associated with the LARE-regression estimator is only 0.082, whereas the classical least-squares approach yields a substantially larger error of 0.37. A similar statement is observed after the removal of outliers: the MAE under the LARE framework remains low at 0.09, while the least-squares estimator still exhibits a markedly higher error of 1.03.

This conclusion clearly underline the robustness of the LARE-regression approach. In particular, the estimator maintains stable and reliable predictive performance even in the presence of extreme observations and irregular fluctuations, conditions that commonly

arise in financial time series. This robustness makes the proposed functional methodology especially well suited for modeling and forecasting data characterized by high volatility and heavy-tailed behavior.

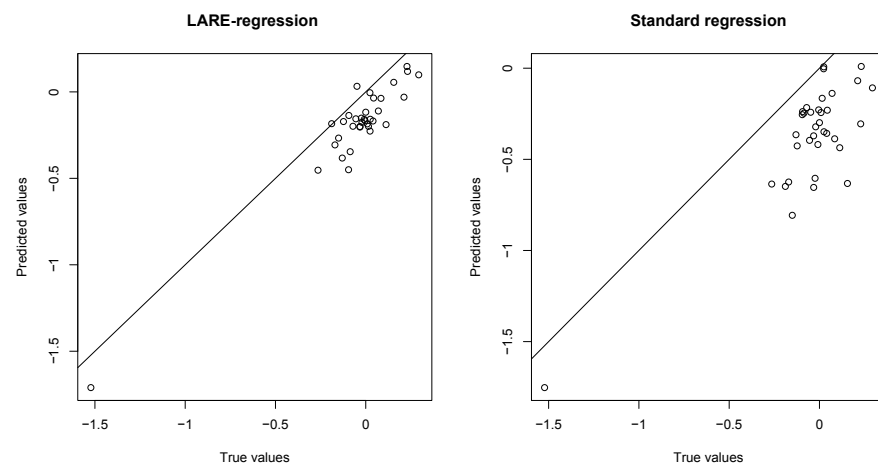


Figure 5. The cleaned data (data without outliers).

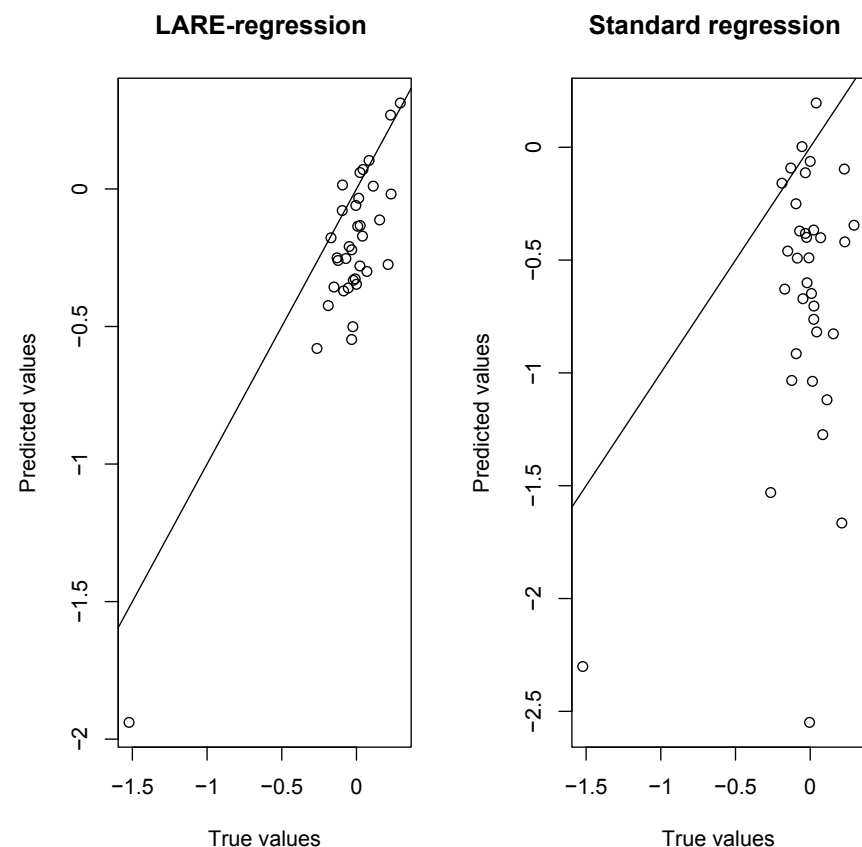


Figure 6. The initial data (data with outliers).

7. Conclusions and Prospects

The present work introduces a robust framework for functional regression estimation based on the least absolute relative error (LARE) loss function. We develop a novel estimator and establish its asymptotic properties, demonstrating its strong consistency. The functional structure is characterized using the small-ball probability function, which explores the topological properties of the probability measure associated with the functional explanatory

variable. Moreover, the use of the LARE criterion substantially enhances the robustness of the regression model. The asymptotic theory is developed under general assumptions, explicitly accommodating strong mixing functional time series. This guarantees that the results hold for a wide range of dependent functional data and continuous-time processes.

Empirical analyses illustrate the ease of implementing the proposed estimator. We demonstrate that it can be applied straightforwardly, with several established techniques available for selecting the smoothing parameter. Computational results further highlight the superiority of the LARE-based functional regression in terms of both robustness and predictive accuracy relative to alternative approaches.

Despite these advantages, some limitations of the proposed nonparametric LARE regression approach should be noted. In particular, the estimator may exhibit sensitivity when the response variable takes values close to zero. In such situations, the relative error term can become excessively large or unstable, which may affect the estimation procedure and generate numerical instability. Therefore, although the relative error framework provides improved robustness in many practical settings, special attention is required when analyzing data that contain small response values.

This contribution also opens several avenues for future research. A natural extension would be to develop a local linear variant of this robust modal regression estimator, as local linear modeling is known to improve convergence rates and reduce boundary bias. Establishing the asymptotic normality of the proposed estimator is another important open problem, essential for constructing confidence intervals and conducting hypothesis tests. Additional directions include exploring linear or partially linear formulations of the regression model. A particularly challenging but impactful extension would be adapting these methods to more complex data structures, such as censored data or ergodic functional time series. Of course addressing these challenges require novel mathematical tools and theoretical developments beyond the scope of the present study.

Author Contributions: The authors contributed approximately equally to this work. Formal analysis, F.A.A.; validation, M.B.A. writing—review and editing, A.L. 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 (PNURSP2026R515), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, and the Deanship of Scientific Research and Graduate Studies at King Khalid University through the Research Groups Program under grant number R.G.P. 2/17/47.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available through the link <https://stooq.com/db/h/>, accessed on 18 March 2026).

Acknowledgments: 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 (PNURSP2026R515), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, and the Deanship of Scientific Research and Graduate Studies at King Khalid University through the Research Groups Program under grant number R.G.P. 2/17/47.

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

Appendix A

This section is devoted to the proofs of the main results. For the reader's convenience, we first summarize the principal notation in Table A1 and recall the Fuk–Nagaev inequality in Lemma A1.

Table A1. Summary of the main notation used throughout the paper.

Notation	Description
$(\mathfrak{S}, D)$	semi-metric space
$\mathcal{F}_i$	The i th functional covariate.
$\mathcal{R}_i$	The i th scalar response variable.
$RE(F)$	The LARE regression evaluated at F .
$\widehat{RE}(F)$	The kernel estimator of the LARE regression .
$\mathbf{Q}_i = \mathbf{Q}(f^{-1}D(F, \mathcal{F}_i))$	The kernel weight assigned to the observation $\mathcal{F}_i$.
$\mathcal{P}_F(f)$	The small-ball probability around F with radius f .
$\alpha(n)$	The strong mixing coefficient r .

Lemma A1. Let $(Z_i)_{i \in \mathbb{N}}$ be an identically distributed algebraic α -mixing process. If $|Z_1| \leq M$ almost surely, then for all $r \geq 1$,

$$P\left(\left|\sum_{i=1}^n Z_i\right| > \varepsilon\right) \leq C \left[\left(1 + \frac{\varepsilon^2}{rs_n^2}\right)^{-r/2} + nr^{-1} \left(\frac{r}{\varepsilon}\right)^{a+1} \right].$$

Here

$$s_n^2 = \sum_{i=1}^n \sum_{j=1}^n |\text{Cov}(Z_i, Z_j)|.$$

Appendix B. Proof of Theorem 1

The proof is based on the applying Lemma 1 in [18] on

$$\mathcal{U}_n(d) = \frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} (0.5 - \mathbf{1}_{\mathcal{R}_i \leq (d + RE(F))}) \mathbf{Q}_i,$$

$$d_n = \widehat{RE}(F) - RE(F), \quad \mathcal{D}_n = \mathcal{U}_n(0).$$

Therefore, it suffices to verify the conditions of the cited lemma in order to obtain the desired Bahadur representation of $\widehat{RE}(F) - RE(F)$. Specifically, we need to establish that

$$\mathcal{D}_n = o_{a.co.}(1), \quad (\text{A1})$$

$$\sup_{|d| \leq A} |\mathcal{U}_n(d) + \beta d - \mathcal{D}_n| = o_{a.co.}(1), \quad \text{for some } A, \beta > 0. \quad (\text{A2})$$

and

$$\frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} (0.5 - \mathbf{1}_{\mathcal{R}_i \leq \widehat{RE}}) \mathbf{Q}_i = O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right). \quad (\text{A3})$$

Finally, the proof of Theorem 1 follows as a direct consequence of the lemmas stated below.

Lemma A2. Given assumptions of Theorem 1, we obtain

$$|\mathcal{D}_n| = O(f^{\min(b_1, b_2)}) + O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right) \quad (\text{A4})$$

Proof. To do that we employ a truncation argument, since $|\mathcal{R}_i|^{-1}$ is not necessarily bounded. More precisely, we introduce the following truncated quantities:

$$\mathcal{V}^* = \frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n \mathbf{Q}(f^{-1}D(F, \mathcal{F}_i)) \mathcal{R}_i^*,$$

where

$$\mathcal{R}_i^* = \left(0.5 - \mathbf{1}_{\{\mathcal{R}_i \leq RE(F)\}}\right) |\mathcal{R}_i|^{-1} \mathbf{1}_{\{|\mathcal{R}_i|^{-1} < \gamma_n\}}, \quad \text{with} \quad \gamma_n = n^{a/p}.$$

The desired result follows from the following three intermediate bounds:

$$|\mathbb{E}[\mathcal{V}^*] - \mathbb{E}[\mathcal{V}]| = O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right), \quad (\text{A5})$$

$$|\mathcal{V}^* - \mathcal{V}| = O_{a.co.}\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right), \quad (\text{A6})$$

and

$$|\mathcal{V}^* - \mathbb{E}[\mathcal{V}^*]| = O_{a.co.}\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right). \quad (\text{A7})$$

We begin by proving (A5). Observe that

$$|\mathbb{E}[\mathcal{V}^*] - \mathbb{E}[\mathcal{V}]| \leq C \frac{1}{\mathcal{P}_F(f)} \mathbb{E}\left[|\mathcal{R}|^{-1} \mathbf{1}_{\{|\mathcal{R}|^{-1} \geq \gamma_n\}} \mathbf{Q}(f^{-1}D(F, \mathcal{F}))\right].$$

Applying Hölder's inequality with $\alpha_0 = \frac{p}{2}$ and β_0 such that $\frac{1}{\alpha_0} + \frac{1}{\beta_0} = 1$, we obtain

$$\begin{aligned} \mathbb{E}\left[|\mathcal{R}|^{-1} \mathbf{1}_{\{|\mathcal{R}|^{-1} \geq \gamma_n\}} \mathbf{Q}(f^{-1}D(F, \mathcal{F}_1))\right] &\leq \mathbb{E}^{1/\alpha_0}\left[|\mathcal{R}|^{-\alpha_0} \mathbf{1}_{\{|\mathcal{R}|^{-1} \geq \gamma_n\}}\right] \\ &\quad \times \mathbb{E}^{1/\beta_0}\left[\mathbf{Q}^{\beta_0}(f^{-1}D(F, \mathcal{F}_1))\right] \\ &\leq \gamma_n^{-1} \mathbb{E}^{1/\alpha_0}\left[|\mathcal{R}|^{-2\alpha_0}\right] \\ &\quad \times \mathbb{E}^{1/\beta_0}\left[\mathbf{Q}^{\beta_0}(f^{-1}D(F, \mathcal{F}_1))\right] \\ &\leq \gamma_n^{-1} \mathbb{E}^{1/\alpha_0}\left[|\mathcal{R}|^{-p}\right] \\ &\quad \times \mathbb{E}^{1/\beta_0}\left[\mathbf{Q}^{\beta_0}(f^{-1}D(F, \mathcal{F}_1))\right] \\ &\leq C \gamma_n^{-1} \mathcal{P}_F^{1/\beta_0}(f). \end{aligned}$$

This establishes the required bound and completes the proof of (A5). Consequently,

$$|\mathcal{V}^* - \mathbb{E}[\mathcal{V}^*]| \leq n^{1/2-a/p} \mathcal{P}_F^{(1-\beta)/\beta}(f).$$

Since $a > p$, it follows that the right-hand side converges to zero, which yields (A5). We now turn to the proof of (A6). To this end, we apply Markov's inequality. For any $\epsilon > 0$, we have

$$\begin{aligned} \mathbb{P}(|\mathcal{V}^* - \mathcal{V}| > \epsilon) &\leq \sum_{i=1}^n \mathbb{P}\left(|\mathcal{R}_i|^{-1} > n^{a/p}\right) \\ &\leq n \mathbb{P}\left(|\mathcal{R}|^{-1} > n^{a/p}\right) \\ &\leq n^{1-a} \mathbb{E}\left[|\mathcal{R}|^{-p}\right]. \end{aligned}$$

In particular, choosing

$$\epsilon = \epsilon_0 \left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right),$$

and using the condition $a > \frac{5}{2}$, we obtain

$$\mathbb{P}\left(|\mathcal{V} - \mathcal{V}^*| > \epsilon_0 \left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n \mathcal{P}_F^2(f)}} \right)\right) \leq n^{3/2-a} < Cn^{-1-\nu},$$

for some constants $C > 0$ and $\nu > 0$. We now proceed to prove (A7). Define

$$\Lambda_i = \mathbf{Q}_i \mathcal{R}_i^* - \mathbb{E}[\mathbf{Q}_1 \mathcal{R}_1^*].$$

Then, for any $\varepsilon > 0$, we have

$$\begin{aligned} \mathbb{P}\{|\mathcal{V}^* - \mathbb{E}[\mathcal{V}^*]| > \varepsilon\} &= \mathbb{P}\left\{\left|\frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n \Lambda_i\right| > \varepsilon\right\} \\ &\leq \mathbb{P}\left\{\left|\sum_{i=1}^n \Lambda_i\right| > \varepsilon n \mathbb{E}[\mathbf{Q}_1]\right\}. \end{aligned}$$

To bound this probability, we apply the Fuk–Nagaev inequality (see Lemma A1), which requires controlling the asymptotic behavior of

$$S_n'^2 = \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\Lambda_i, \Lambda_j) = \sum_{i=1}^n \sum_{i \neq j} \text{Cov}(\Lambda_i, \Lambda_j) + n \text{Var}(\Lambda_1).$$

To this end, we decompose the double sum into two parts. Let

$$S_1' = \{(i, j) : 1 \leq |i - j| \leq u_n\}, \quad S_2' = \{(i, j) : u_n + 1 \leq |i - j| \leq n - 1\},$$

and denote by $J_{1,n}'$ and $J_{2,n}'$ the corresponding sums of covariances over S_1' and S_2' , respectively.

On S_1' , we have

$$\begin{aligned} J_{1,n}' &= \sum_{S_1'} |\text{Cov}(\Lambda_i, \Lambda_j)| \\ &\leq C \sum_{S_1'} (|\mathbb{E}[\mathbf{Q}_i \mathbf{Q}_j]| + |\mathbb{E}[\mathbf{Q}_i] \mathbb{E}[\mathbf{Q}_j]|), \end{aligned}$$

which yields

$$J_{1,n}' \leq C n u_n \Phi_F(f).$$

On S_2' , we apply Davydov–Rio’s inequality (in the L^∞ case) to obtain

$$|\text{Cov}(\Lambda_i, \Lambda_j)| \leq C \gamma_n^2 \alpha(|i - j|).$$

Consequently,

$$J_{2,n}' = \sum_{S_2'} |\text{Cov}(\Lambda_i, \Lambda_j)| \leq \frac{n \gamma_n^2 u_n^{-a+1}}{a-1}.$$

Choosing

$$u_n = \left(\frac{\gamma_n^2}{\Phi_F(f)} \right)^{1/a},$$

we deduce that

$$\sum_{i=1}^n \sum_{i \neq j} \text{Cov}(\Lambda_i, \Lambda_j) = O\left(n \gamma_n^{2/a} \Phi_F(f)^{(a-1)/a}\right).$$

On the other hand, since

$$\text{Var}(\Lambda_1) \leq \mathbb{E}[(\mathbf{Q}_1 \mathcal{R}_1^*)^2] \leq C \mathbb{E}[\mathbf{Q}_1^2] = O(\mathcal{P}_F(f)),$$

it follows that

$$S_n'^2 = O(n \Phi_F^{1/2}(f)). \quad (\text{A8})$$

We now apply the Fuk–Nagaev inequality to the sequence $\{\Lambda_i\}$. For any $\ell > 0$ and $\varepsilon > 0$, we have

$$\begin{aligned} \mathbb{P}\left\{\left|\mathbb{E}[\hat{\Psi}_N^*] - \hat{\Psi}_N^*\right| > \varepsilon\right\} &\leq \mathbb{P}\left\{\left|\sum_{i=1}^n \Lambda_i\right| > \varepsilon n \mathbb{E}[\mathbf{Q}_1]\right\} \\ &\leq C(A_1' + A_2'), \end{aligned}$$

where

$$A_1' = \left(1 + \frac{\varepsilon^2 n^2 (\mathbb{E}[\mathbf{Q}_1])^2}{S_n'^2 \ell}\right)^{-\ell/2}, \quad A_2' = n \ell^{-1} \left(\frac{\ell}{\varepsilon n \mathbb{E}[\mathbf{Q}_1]}\right)^{a+1}.$$

Choosing

$$\varepsilon = \lambda' \frac{\sqrt{n \log n \Phi_F^{1/2}(f)}}{n \mathbb{E}[\mathbf{Q}_1]} \quad \text{and} \quad \ell = C(\log n)^2,$$

we obtain, under assumption (RE6),

$$A_2' \leq C n^{1-(a+1)/2} \Phi_F(f)^{-(a+1)/4} (\log n)^{(3a-1)/2} \leq C n^{-1-\nu_1'}, \quad (\text{A9})$$

for some $\nu_1' > 0$.

Similarly, using $\ell = O((\log n)^2)$, we derive

$$A_1' \leq C \left(1 + \frac{\lambda'^2 \log n}{\ell}\right)^{-\ell/2} \leq C n^{-1-\nu_2'}, \quad (\text{A10})$$

for some $\nu_2' > 0$.

Combining (A9) and (A10), we obtain the desired bound, which establishes (A7).

Finally, we control $\mathcal{D}_n$. We have

$$\begin{aligned} \mathbb{E}[|\mathcal{D}_n|] &= \frac{1}{\mathbb{E}[\mathbf{Q}_1]} \mathbb{E}\left[|\mathcal{R}_1|^{-1} \left(0.5 - \mathbf{1}_{[\mathcal{R}_1 \leq RE(F)]}\right) \mathbf{Q}_1\right] \\ &\leq \frac{1}{\mathbb{E}[\mathbf{Q}_1]} \mathbb{E}[\mathbf{Q}_1 | U_1(RE(F), F) - U_1(RE(F), T_1)|] \\ &\quad + \frac{1}{\mathbb{E}[\mathbf{Q}_1]} \mathbb{E}[\mathbf{Q}_1 | U_2(RE(F), F) - U_2(RE(F), T_1)|] \\ &= O(f^{b_1}) + O(f^{b_2}). \end{aligned}$$

It follows that

$$|\mathcal{D}_n| = O(f^{\min(b_1, b_2)}) + O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n \mathcal{P}_F^2(f)}}\right).$$

□

Lemma A3. Given assumptions of Theorem 1, we obtain

$$\sup_{|d| \leq A} |\mathcal{U}_n(d) + \beta d - \mathcal{D}_n| = O(f^{\min(b_1, b_2)}) + O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n \mathcal{P}_F^2(f)}}\right).$$

for $\beta = U(s, F) = \frac{\partial(U_2(s, F) - U_1(s, F))}{\partial s}$ and

$$\frac{1}{n \mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} (0.5 - \mathbf{1}_{\mathcal{R}_i \leq \widehat{R\hat{E}}}) \mathbf{Q}_i = O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n \mathcal{P}_F^2(f)}}\right). \quad (\text{A11})$$

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

$$\sup_{|d| \leq A} |\mathcal{U}_n(d) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n]| = O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n \mathcal{P}_F^2(f)}}\right), \quad (\text{A12})$$

and the bias term

$$\sup_{|d| \leq A} |\mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n] + U(RE(F), F)d| = O(f^{b_1}) + O(f^{b_2}), \quad (\text{A13})$$

where

$$U(s, F) = \frac{\partial}{\partial s} (U_2(s, F) - U_1(s, F)).$$

To establish (A12), we exploit the compactness of the interval $[-A, A]$ and construct a finite covering

$$[-A, A] \subset \bigcup_{j=1}^{d_n} [d_j - l_n, d_j + l_n], \quad d_j \in [-A, A],$$

with $l_n = d_n^{-1} = n^{-1/2}$. For any $d \in [-A, A]$, let

$$j(d) = \arg \min_j |d - d_j|.$$

We then decompose the supremum as

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

Using the inequality

$$|\mathbf{1}_{[S < a]} - \mathbf{1}_{[S < b]}| \leq \mathbf{1}_{[|S - b| \leq |a - b|]},$$

the first term can be bounded as

$$\sup_{|d| \leq A} |\mathcal{U}_n(d) - \mathcal{U}_n(d_j)| \leq \frac{1}{n \mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n Z_i,$$

where

$$Z_i = \sup_{|d| \leq A} |\mathcal{R}_i^{-1}| \mathbf{1}_{[|\mathcal{R}_i - d_{j(d)} - RE(F)| \leq Cl_n]} \mathbf{Q}_i.$$

Using arguments similar to those employed in the first lemma, we obtain

$$\sup_{|d| \leq A} |\mathcal{U}_n(d) - \mathcal{U}_n(d_j)| = O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right).$$

For the third term,

$$\sup_{|d| \leq A} |\mathbb{E}[\mathcal{U}_n(d) - \mathcal{U}_n(d_j)]| \leq \frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \mathbb{E}[Z_1] \leq Cl_n.$$

Since

$$l_n = o\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right), \quad (\text{A14})$$

we deduce that

$$\sup_{|d| \leq A} |\mathbb{E}[\mathcal{U}_n(d) - \mathcal{U}_n(d_j)]| = O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right).$$

We now consider the middle term

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

Define

$$\mathcal{U}_n(d_j) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d_j) - \mathcal{D}_n] = \frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n \Psi_i,$$

where

$$\Psi_i = |\mathcal{R}_i^{-1}| \left(\mathbf{1}_{\{\mathcal{R}_i \leq RE(F)\}} - \mathbf{1}_{\{\mathcal{R}_i \leq d_j + RE(F)\}} \right) \mathbf{Q}_i - \mathbb{E}[\cdot].$$

Applying once again the truncation argument together with the Fuk–Nagaev inequality, as in the previous lemma, yields

$$\sup_{|d| \leq A} |\mathcal{U}_n(d_j) - \mathcal{D}_n - \mathbb{E}[\mathcal{U}_n(d_j) - \mathcal{D}_n]| = O\left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}}\right).$$

Combining the previous bounds establishes (A12).

We now turn to the bias term (A13). For any $d \in [-A, A]$,

$$\begin{aligned} \mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n] &= -\frac{1}{\mathbb{E}[\mathbf{Q}_1]} \mathbb{E}\left[|\mathcal{R}^{-1}| (\mathbf{1}_{[\mathcal{R}_1 \leq d + RE(F)]} - \mathbf{1}_{[\mathcal{R}_1 \leq RE(F)]}) \mathbf{Q}_1\right] \\ &= -\frac{1}{\mathbb{E}[\mathbf{Q}_1]} \mathbb{E}\left[(U_1(d + RE(F), F) - U_1(RE(F), F)) \mathbf{Q}_1\right] \\ &\quad + \frac{1}{\mathbb{E}[\mathbf{Q}_1]} \mathbb{E}\left[(U_2(d + RE(F), F) - U_2(RE(F), F)) \mathbf{Q}_1\right] \\ &\quad + O(f^{b_1}) + O(f^{b_2}). \end{aligned}$$

Using a first-order expansion around $RE(F)$ gives

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

Hence

$$\sup_{|d| \leq A} |\mathbb{E}[\mathcal{U}_n(d) - \mathcal{D}_n] + U(RE(F), F)d| = O(f^{b_1}) + O(f^{b_2}),$$

which establishes (A13).

Now, we return to the proof of (A11). Similarly to [27,28] we employ the subgradients and weighted median arguments to prove that

$$\frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} (0.5 - \mathbf{1}_{\mathcal{R}_i \leq \widehat{RE}}) \mathbf{Q}_i \leq \frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} \mathbf{1}_{\widehat{RE} - l_n \leq \mathcal{R}_i \leq \widehat{RE} + l_n} \mathbf{Q}_i$$

By a simple analytical arguments we prove that the bias part of

$$\mathbb{E} \left[\frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} \mathbf{1}_{\widehat{RE} - l_n \leq \mathcal{R}_i \leq \widehat{RE} + l_n} \mathbf{Q}_i \right] = O(l_n) = o \left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}} \right)$$

Further, the dispersion term is evaluated by using the Fuk–Nagaev inequality

$$\begin{aligned} \frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} \mathbf{1}_{\widehat{RE} - l_n \leq \mathcal{R}_i \leq \widehat{RE} + l_n} \mathbf{Q}_i - \mathbb{E} \left[\frac{1}{n\mathbb{E}[\mathbf{Q}_1]} \sum_{i=1}^n |\mathcal{R}_i|^{-1} \mathbf{1}_{\widehat{RE} - l_n \leq \mathcal{R}_i \leq \widehat{RE} + l_n} \mathbf{Q}_i \right] \\ = O_{a.co.} \left(\sqrt{\frac{\Phi_F^{1/2}(f) \log n}{n\mathcal{P}_F^2(f)}} \right) \end{aligned}$$

which complete the proof of the lemma. $\square$

Therefore, the conditions of Lemma 1 in [18] are satisfied, yielding the Bahadur representation

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

Finally, the convergence rates established in the two lemmas imply the result stated in Theorem 1.

References

1. Ferraty, F.; Vieu, P. The functional nonparametric model and application to spectrometric data. *Comput. Stat.* **2002**, *17*, 545–564. [CrossRef]
2. Dabo-Niang, S.; Rhomari, N. Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *Comptes Rendus Math.* **2003**, *336*, 75–80. [CrossRef]
3. Delsol, L. Régression non-paramétrique fonctionnelle: Expressions asymptotiques des moments. In *Annales de l'ISUP; LSTA*: Paris, France, 2007; Volume 51, pp. 43–67.
4. Masry, E. Nonparametric regression estimation for dependent functional data: Asymptotic normality. *Stoch. Process. Their Appl.* **2005**, *115*, 155–177. [CrossRef]
5. Hu, X.; Wang, J.; Wang, L.; Yu, K. K-nearest neighbor estimation of functional nonparametric regression model under NA samples. *Axioms* **2022**, *11*, 102. [CrossRef]
6. Wiem, N.; Salah, K.; Françoise, Y.A. Nonparametric regression on Riemannian manifolds under an α -mixing process. *Commun. Stat.—Theory Methods* **2026**, *1*–16. [CrossRef]
7. Barrientos-Marin, J.; Ferraty, F.; Vieu, P. Locally modelled regression and functional data. *J. Nonparametr. Stat.* **2010**, *22*, 617–632. [CrossRef]

8. Narula, S.C.; Wellington, J.F. Prediction, linear regression and the minimum sum of relative errors. *Technometrics* **1977**, *19*, 185–190. [[CrossRef](#)]
9. Chatfield, C. The joys of consulting. *Significance* **2007**, *4*, 33–36. [[CrossRef](#)]
10. Chen, K.; Guo, S.; Lin, Y.; Ying, Z. Least absolute relative error estimation. *J. Am. Stat. Assoc.* **2010**, *105*, 1104–1112. [[CrossRef](#)] [[PubMed](#)]
11. Yang, Y.; Ye, F. General relative error criterion and M-estimation. *Front. Math. China* **2013**, *8*, 695–715. [[CrossRef](#)]
12. Jones, M.C.; Park, H.; Shin, K.I.; Vines, S.K.; Jeong, S.O. Relative error prediction via kernel regression smoothers. *J. Stat. Plan. Inference* **2008**, *138*, 2887–2898. [[CrossRef](#)]
13. Attouch, M.; Laksaci, A.; Messabihi, N. Nonparametric relative error regression for spatial random variables. *Stat. Pap.* **2017**, *58*, 987–1008. [[CrossRef](#)]
14. Demongeot, J.; Hamie, A.; Laksaci, A.; Rachdi, M. Relative-error prediction in nonparametric functional statistics: Theory and practice. *J. Multivar. Anal.* **2016**, *146*, 261–268. [[CrossRef](#)]
15. Altendji, B.; Demongeot, J.; Laksaci, A.; Rachdi, M. Functional data analysis: Estimation of the relative error in functional regression under random left-truncation model. *J. Nonparametr. Stat.* **2018**, *30*, 472–490. [[CrossRef](#)]
16. Benzamouche, S.; Ould Saïd, E.; Sadki, O. Nonparametric estimation of the relative error regression for twice censored and dependent data. *Commun. Stat.—Theory Methods* **2025**, *54*, 1492–1525. [[CrossRef](#)]
17. Xia, X.; Ming, H.; Li, J. Relative error model average for multiplicative models. *Stat. Comput.* **2026**, *36*, 18. [[CrossRef](#)]
18. Laksaci, A.; Almanjahie, I.M.; Rachdi, M. Nonparametric Functional Least Absolute Relative Error Regression: Application to Econophysics. *Mathematics* **2026**, *14*, 268. . [[CrossRef](#)]
19. Wang, J.L.; Chiou, J.M.; Müller, H.G. Functional data analysis. *Annu. Rev. Stat. Its Appl.* **2016**, *3*, 257–295. [[CrossRef](#)]
20. Bosq, D. *Linear Processes in Function Spaces: Theory and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2000; Volume 149. [[CrossRef](#)]
21. Horváth, L.; Kokoszka, P.; Reeder, R. Estimation of the mean of functional time series and a two-sample problem. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2013**, *75*, 103–122. [[CrossRef](#)]
22. Amato, U.; Antoniadis, A.; De Feis, I.; Gijbels, I. Functional time series forecasting: A systematic review. *Stat. Pap.* **2025**, *66*, 21. [[CrossRef](#)]
23. Rosenblatt, M. A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **1956**, *42*, 43–47. . [[CrossRef](#)] [[PubMed](#)]
24. Bradley, R.C.; Jurek, Z.J. The strong mixing and the selfdecomposability properties. *Stat. Probab. Lett.* **2014**, *84*, 67–71. [[CrossRef](#)]
25. Francq, C.; Zakoian, J.M. Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econom. Theory* **2006**, *22*, 815–834. [[CrossRef](#)]
26. Maronna, R.A.; Martin, R.D.; Yohai, V.J.; Salibián-Barrera, M. *Robust Statistics: Theory and Methods (with R)*; John Wiley & Sons: Hoboken, NJ, USA, 2019. . [[CrossRef](#)]
27. Hallin, M.; Lu, Z.; Yu, K. Local linear spatial quantile regression. *Bernoulli* **2009**, *15*, 659–686. [[CrossRef](#)]
28. Ruppert, D.; Carroll, R.J. Trimmed least square estimation in the linear model. *J. Am. Stat. Assoc.* **1980**, *75*, 828–838. [[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.