VILNIUS
TECH
Vilnius Gediminas
Technical University

Function-on-Function Linear Quantile Regression

Ufuk Beyaztas[a] and Han Lin Shang[b]

[a] *Department of Statistics, Marmara University*
Goztepe Campus, 34722 Istanbul, Turkey
[b] *Department of Actuarial Studies and Business Analytics, Macquarie University*
Sydney, 2109 NSW, Australia

E-mail(*corresp.*): `ufuk.beyaztas@marmara.edu.tr`

E-mail: `hanlin.shang@mq.edu.au`

Abstract. In this study, we propose a function-on-function linear quantile regression model that allows for more than one functional predictor to establish a more flexible and robust approach. The proposed model is first transformed into a finite-dimensional space via the functional principal component analysis paradigm in the estimation phase. It is then approximated using the estimated functional principal component functions, and the estimated parameter of the quantile regression model is constructed based on the principal component scores. In addition, we propose a Bayesian information criterion to determine the optimum number of truncation constants used in the functional principal component decomposition. Moreover, a stepwise forward procedure and the Bayesian information criterion are used to determine the significant predictors for including in the model. We employ a nonparametric bootstrap procedure to construct prediction intervals for the response functions. The finite sample performance of the proposed method is evaluated via several Monte Carlo experiments and an empirical data example, and the results produced by the proposed method are compared with the ones from existing models.

Keywords: function-on-function regression, functional principal component analysis, median regression, quantile regression.

AMS Subject Classification: 62G08; 65D10.

1 Introduction

With advancements in technology and data storage, it is progressively common to obtain data whose sample elements are collected and saved over a continuum, such as space grids, time, depth, and wavelength. By interpolation and smoothing, these data are represented in curves, images, shapes, or more general objects called functional data. The availability of functional data has regularly increased in almost all scientific branches. The need for developing statistical approaches for this kind of data is also increasing. Consult [9,11,14,15,30,31,32] and [21] for the theoretical results and case studies of functional data analysis methods. Among many others, function-on-function regression (FFR) models have received considerable attention among researchers from the statistics community to investigate the relationship between a functional response and one or more functional explanatory variables. See, e.g., [6,8,13,17,26,35] and [2].

Let $\{\mathcal{Y}_i, \mathcal{X}_i : i = 1, \dots, n\}$ denote a random sample from a pair $(\mathcal{Y}, \mathcal{X})$, where both the response $\mathcal{Y}$ and predictor $\mathcal{X}$ variables are square-integrable random functions, $\mathcal{Y} = (\mathcal{Y}(t))_{t \in \mathcal{I}}$ and $\mathcal{X} = (\mathcal{X}(s))_{s \in S}$, on the bounded and closed intervals $\mathcal{I}, S \in \mathbb{R}$. Without loss of generality, we assume that both the response and predictor are mean-zero processes so that $\mathrm{E}[\mathcal{Y}(t)] = \mathrm{E}[\mathcal{X}(s)] = 0$. Then, the FFR model of $\mathcal{Y}$ on $\mathcal{X}$ is defined as follows:

$$\mathrm{E}[\mathcal{Y}(t)|\mathcal{X}(s)] = \int_{s \in S} \mathcal{X}(s)\beta(s,t)ds + \epsilon(t), \qquad (1.1)$$

where $\beta(s,t)$ is the smooth bivariate coefficient function and $\epsilon(t)$ is the error function with mean-zero. In the FFR model, the conditional mean of the functional response $\mathcal{Y}$ given a functional predictor $\mathcal{X}$, $\mathrm{E}[\mathcal{Y}|\mathcal{X}]$, is recovered via optimizing least squares (LS) loss, i.e., $\left[\mathcal{Y}(t) - \int_{s \in S} \mathcal{X}(s)\beta(s,t)ds\right]^2$. However, the LS loss is sensitive to outliers, which are observations far from the bulk of the data. Outliers are very common in empirical applications. In such cases, biased regression estimates are obtained under the LS loss, leading to unreliable inferences. Also, optimal regression estimates for the FFR model in (1.1) may not be obtained when the error term follows a non-normal heavy-tailed distribution. To remedy these problems, the quantile regression may be used as an alternative to model (1.1).

Quantile regression (QR), which was introduced by [20], has become a general framework to investigate the effects of predictors at different quantile levels of the response. QR renders it possible to characterize the entire conditional distribution of the response variable. Thus, it provides a more general picture of the response variable's conditional distribution function that might not be reflected by mean regression. There are several advantages of QR over mean regression: 1) QR is more robust to outliers than mean regression when outlying points are present in the response variable since it belongs to a robust model family [18]; 2) Compared with mean regression, QR produces more efficient results when the errors follow a non-normal heavy-tailed distribution; and 3) QR provides more efficient results than mean regression in the presence of conditional heteroscedasticity, the case when the variance depends on one or more predictors. Besides, QR allows obtaining prediction intervals based on

several quantiles of the response. Consult [18] for the theoretical materials and empirical applications of QR.

In FRMs, several QR models have been proposed when one of the observed response or predictor variables consists of random functions rather than scalar observations. See, e.g., [4, 5, 10, 25, 34, 36, 37] and [23]. The numerical results obtained from the studies mentioned above have shown that the QR produces better prediction and estimation accuracy than standard linear FRMs when the errors follow a non-normal heavy-tailed distribution and in the presence of outliers.

To the best of our knowledge, there has been no comprehensive work for QR in the case of function-valued response and predictors (function-on-function regression). In this paper, we propose a function-on-function linear quantile regression (FFLQR) model, which allows for more than one functional predictor by extending the traditional FFR model into the QR setting to establish a more flexible and robust approach. Our proposal also aims to present a more comprehensive description of the relationship between the functional response and predictors by focusing on conditional quantiles at different quantile levels. In traditional QR, directly modeling conditional quantiles and estimating the regression coefficients is performed by minimizing the check loss function. Although they are practically observed in discrete-time points in the functional data framework, the functional random variables intrinsically belong to an infinite-dimensional space. Thus, as in FRM, the direct estimation of the proposed method is an ill-posed problem. The common practical approach to remedy this problem is based on projecting infinite-dimensional functional objects onto a finite-dimensional space using dimension reduction methods. For this purpose, several general basis expansion methods, such as B-spline, Fourier, and wavelet basis have been proposed [17, 30, 32]. However, general basis expansion methods may require a large number of basis functions to project functional objects onto a finite-dimensional space, leading to poor estimation and prediction accuracy. On the other hand, the dimension reduction techniques including the functional principal component (FPC) analysis provide more informative approximation since they are data-driven and uses the information of functional predictors gathered from their covariance functions [12]. Therefore in this paper, the FPC analysis is first used to transform the functional random variables into a finite-dimensional space. This approximation allows converting infinite-dimensional FFLQR into a multivariate regression model of principal component scores. In what follows, we estimate the proposed FFLQR model using the FPC approximations of the functional response and predictor variables. The prediction performance of the proposed method is based on the truncation constants used in the decompositions of the functional objects. In this paper, the optimum values of the truncation constants are determined via a Bayesian information criterion (BIC). In practice, the exact form of the actual model is unspecified since the model's significant predictors are unknown. To this end, we employ a stepwise forward procedure along with the BIC to determine significant predictors. We apply a nonparametric bootstrap method to construct pointwise prediction intervals for the response function at different quantile levels.

The remainder of this paper is organized as follows. Section 2 presents the methodology of the proposed method. The results of Monte Carlo experiments and empirical data analysis are given in Section 3. Section 4 concludes the paper.

2 Methodology

We consider a random sample $\{\mathcal{Y}_i, \boldsymbol{\mathcal{X}}_i : i = 1, \ldots, n\}$ from a pair $(\mathcal{Y}, \boldsymbol{\mathcal{X}})$, where $\mathcal{Y} = (\mathcal{Y}(t))_{t \in \mathcal{I}}$ is a functional response and $\boldsymbol{\mathcal{X}} = [\mathcal{X}_{i1}, \ldots, \mathcal{X}_{iM}]^\top$ with $\mathcal{X}_m = (\mathcal{X}_m(s))_{s \in S}$ is an M dimensional vector of functional predictors, defined on the bounded and closed intervals $\mathcal{I}, S \in \mathbb{R}$. Without loss of generality, we assume that $\mathcal{I}, S = [0, 1]$ and both the response and predictors are mean-zero processes so that $\mathrm{E}[\mathcal{Y}(t)] = \mathrm{E}[\mathcal{X}_m(s)] = 0$ for $m = 1, \ldots, M$.

Given any $\tau \in (0, 1)$, let $Q_\tau[\mathcal{Y}(t)|\boldsymbol{\mathcal{X}}]$ denote the conditional quantile of the response function $\mathcal{Y}$ given the vector of functional predictors $\boldsymbol{\mathcal{X}}$. Let us assume that $Q_\tau[\mathcal{Y}(t)|\boldsymbol{\mathcal{X}}]$ can be written as a linear functional of $\boldsymbol{\mathcal{X}}$ along with smooth bivariate functions $[\beta_{1\tau}(s, t), \ldots, \beta_{M\tau}(s, t)]$. Then, we define the proposed FFLQR model as follows:

$$Q_\tau[\mathcal{Y}_i(t)|\boldsymbol{\mathcal{X}}_i] = \sum_{m=1}^{M} \int_0^1 \mathcal{X}_{im}(s)\beta_{m\tau}(s, t)ds. \tag{2.1}$$

We consider minimizing the check loss function, $\rho_\tau(u) = u\{\tau - \mathbb{K}(u < 0)\}$ where $\mathbb{K}\{\cdot\}$ denotes an indicator function [18, 20], to solve the τth conditional quantile $Q_\tau[\mathcal{Y}_i(t)|\boldsymbol{\mathcal{X}}_i]$ as follows:

$$\underset{\beta_{1\tau}(s,t),\ldots,\beta_{M\tau}(s,t)}{\arg\min} \sum_{i=1}^{N} \rho_\tau \left[\mathcal{Y}_i(t) - \sum_{m=1}^{M} \int_0^1 \mathcal{X}_{im}(s)\beta_{m\tau}(s, t)ds \right]. \tag{2.2}$$

As stated in Section 1, the functional objects intrinsically belong to an infinite-dimensional space. Consequently, the minimizing problem stated in (2.2) is an ill-posed problem due to the infinite-dimensional nature of $\mathcal{Y}$ and $\boldsymbol{\mathcal{X}}$. To overcome this problem, we consider FPC decomposition for each functional object in (2.1). Let $\mathcal{C}_\mathcal{Y}(t_1, t_2) = \mathrm{Cov}[\mathcal{Y}(t_1), \mathcal{Y}(t_2)]$, $\mathcal{C}_{\mathcal{X}_m}(s_1, s_2) = \mathrm{Cov}[\mathcal{X}(s_1), \mathcal{X}(s_2)]$, respectively, denote the covariance functions of $\mathcal{Y}(t)$ and $\mathcal{X}_m(s)$ satisfying $\int_0^1 \int_0^1 \mathcal{C}_\mathcal{Y}^2(t_1, t_2)dt_1dt_2 < \infty$ and $\int_0^1 \int_0^1 \mathcal{C}_{\mathcal{X}_m}^2(s_1, s_2)ds_1ds_2 < \infty$. Then, by Mercer's theorem, the covariance functions can be represented as follows:

$$\mathcal{C}_\mathcal{Y} = \sum_{k \geq 1} w_k \phi_k(t_1)\phi_k(t_2), \qquad \forall t_1, t_2 \in [0, 1],$$

$$\mathcal{C}_{\mathcal{X}_m} = \sum_{l \geq 1} \kappa_{ml} \psi_{ml}(s_1)\psi_{ml}(s_2), \qquad \forall s_1, s_2 \in [0, 1],$$

where $\{\phi_k(t) : k = 1, 2, \ldots\}$ and $\{\psi_{ml}(s) : l = 1, 2, \ldots\}$ denote the orthonormal eigenfunctions corresponding to the non-negative eigenvalues $\{w_k : k = 1, 2, \ldots\}$ and $\{\kappa_{ml} : l = 1, 2, \ldots\}$ with $w_k \geq w_{k+1}$ and $\kappa_{m1} \geq \kappa_{ml+1}$, respectively. To

obtain FPC decompositions of functional variables, we use predetermined truncation constants $K_{\mathcal{Y}}$ (for $\mathcal{Y}(t)$) and $K_{\mathcal{X}_m}$ (for $\mathcal{X}_m(s)$). By Karhunen-Loève expansion, the functional response and functional predictors can be represented as follows:

$$\mathcal{Y}_i(t) \approx \sum_{k=1}^{K_{\mathcal{Y}}} \xi_{ik}\phi_k(t) = \boldsymbol{\xi}_i^\top \boldsymbol{\phi}(t), \qquad \forall t \in [0,1],$$

$$\mathcal{X}_{im}(s) \approx \sum_{l=1}^{K_{\mathcal{X}_m}} \zeta_{iml}\psi_{ml}(s) = \boldsymbol{\zeta}_{im}^\top \boldsymbol{\psi}_m(s), \qquad \forall s \in [0,1],$$

where the random variables $\xi_{ik} = \int_0^1 \mathcal{Y}_i(t)\phi_k(t)dt$, $\zeta_{iml} = \int_0^1 \mathcal{X}_{im}(s)\psi_{ml}(s)ds$ are the projections of $\mathcal{Y}_i(t)$ and $\mathcal{X}_{im}(s)$ onto their corresponding orthonormal bases, respectively. Similarly, the smooth bivariate coefficient functions $\beta_{m\tau}(s,t)$ can be expressed in terms of eigenfunctions as follows:

$$\beta_{m\tau}(s,t) \approx \sum_{l=1}^{K_{\mathcal{X}_m}} \sum_{k=1}^{K_{\mathcal{Y}}} \beta_{mlk}^{(\tau)}\psi_{ml}(s)\phi_k(t) = \boldsymbol{\psi}_m^\top(s)\boldsymbol{\beta}_m^{(\tau)}\boldsymbol{\phi}(t), \qquad \forall t,s \in [0,1],$$

where $\beta_{mlk}^{(\tau)} = \int_0^1 \int_0^1 \beta_m(s,t)\psi_{ml}(s)\phi_k(t)dsdt$ links the smooth dependence of $\beta_{m\tau}(s,t)$ on the quantile τ.

Next, using a similar parameterization as in [6,16] and by orthonormalities of $\boldsymbol{\phi}(t)$ and $\boldsymbol{\psi}_m(s)$, the τth conditional quantile of the response function in (2.1) can be expressed as follows:

$$Q_\tau\left[\mathcal{Y}_i(t)|\boldsymbol{\mathcal{X}}_i\right] = \sum_{m=1}^{M} \int_0^1 \boldsymbol{\zeta}_{im}^\top \boldsymbol{\psi}_m(s)\boldsymbol{\psi}_m^\top(s)\boldsymbol{\beta}_m^{(\tau)}\boldsymbol{\phi}(t),$$

$$\boldsymbol{\xi}_i^\top \boldsymbol{\phi}(t) = \sum_{m=1}^{M} \boldsymbol{\zeta}_{im}^\top \boldsymbol{\beta}_m^{(\tau)}\boldsymbol{\phi}(t),$$

$$\boldsymbol{\xi}_i^\top \int_0^1 \boldsymbol{\phi}(t)\boldsymbol{\phi}^\top(t)dt = \sum_{m=1}^{M} \boldsymbol{\zeta}_{im}^\top \boldsymbol{\beta}_m^{(\tau)} \int_0^1 \boldsymbol{\phi}(t)\boldsymbol{\phi}^\top(t)dt,$$

$$\boldsymbol{\xi}_i^\top = \sum_{m=1}^{M} \boldsymbol{\zeta}_{im}^\top \boldsymbol{\beta}_m^{(\tau)}.$$

Let $\boldsymbol{\Xi} = \left[\boldsymbol{\xi}_1^\top, \ldots, \boldsymbol{\xi}_n^\top\right]^\top$, $\boldsymbol{\Pi} = [\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_n]^\top$, where $\boldsymbol{\zeta}_i = \left[\boldsymbol{\zeta}_{i1}^\top, \ldots, \boldsymbol{\zeta}_{iM}^\top\right]^\top$ and $\boldsymbol{\mathcal{B}} = \left[\boldsymbol{\beta}_1^{(\tau)}, \ldots, \boldsymbol{\beta}_M^{(\tau)}\right]^\top$ denote the matrices of the FPC scores and coefficient vectors. Then, we have the following multivariate model in matrix form:

$$\boldsymbol{\Xi} = \boldsymbol{\Pi}\boldsymbol{\mathcal{B}}.$$

In what follows, the parameter matrix $\boldsymbol{\mathcal{B}}$ can be estimated by minimizing the check loss function as follows:

$$\widehat{\boldsymbol{\mathcal{B}}} = \underset{\boldsymbol{\mathcal{B}}}{\arg\min}\left[\sum_{i=1}^{N} \rho_\tau\left(\boldsymbol{\Xi}_i - \boldsymbol{\Pi}_i\boldsymbol{\mathcal{B}}\right)\right],$$

where $\widehat{\boldsymbol{\mathcal{B}}} = \left[\widehat{\beta}_1^{(\tau)}, \ldots, \widehat{\beta}_M^{(\tau)} \right]^\top$. Note that the input of ρ_τ in the minimization problem given above is a vector, and thus, minimization is done according to each coordinate of the vector separately to obtain the estimates of each element of $\boldsymbol{\mathcal{B}}$. Accordingly, the mth estimated coefficient function can be reconstructed as

$$\widehat{\beta}_{m\tau}(s,t) = \sum_{l=1}^{K_{\mathcal{X}_m}} \sum_{k=1}^{K_{\mathcal{Y}}} \psi_{ml}(s) \widehat{\beta}_{mlk}^{(\tau)} \phi_k(t) = \boldsymbol{\psi}_m^\top(s) \widehat{\boldsymbol{\beta}}_m^{(\tau)} \boldsymbol{\phi}(t),$$

and for a given $\tau \in (0,1)$, the estimated conditional quantile of response in (2.1) is obtained

$$\widehat{Q}_\tau \left[\mathcal{Y}_i(t) | \boldsymbol{\mathcal{X}}_i \right] = \sum_{m=1}^M \int_0^1 \mathcal{X}_{im}(s) \widehat{\beta}_{m\tau}(s,t) ds. \tag{2.3}$$

Note that, if the functional response and/or one or more functional predictors are not mean-zero processes, i.e., $\mathrm{E}[\mathcal{Y}(t)] \neq 0$ and/or $\mathrm{E}[\mathcal{X}_m(s)] \neq 0$ ($m \in [1, \ldots, M]$), then the conditional quantile of $\mathcal{Y}(t)$ given $\boldsymbol{\mathcal{X}}$ includes an intercept function, say $\beta_{0\tau}(t)$, as follows:

$$Q_\tau \left[\mathcal{Y}_i(t) | \boldsymbol{\mathcal{X}}_i \right] = \beta_{0\tau}(t) + \sum_{m=1}^M \int_0^1 \mathcal{X}_{im}(s) \beta_{m\tau}(s,t) ds.$$

In such a case, the estimate of the intercept function can be obtained by including the FPC decomposition of $\beta_{0\tau}(t)$, i.e., $\beta_{0\tau}(t) = \sum_{k=1}^{K_{\mathcal{Y}}} g_k \phi_k(t)$, into the check loss function.

The results given above demonstrate that: 1) the infinite-dimensional minimization problem (2.2) can be reduced to a simple finite-dimensional QR problem using the metrics $\{\boldsymbol{\psi}_m(s) : m = 1, \ldots, M\}$ and $\boldsymbol{\phi}(t)$ in the spaces of FPC scores $\left\{ \boldsymbol{\zeta}_{im}^\top : m = 1, \ldots, M \right\}$ and $\boldsymbol{\xi}_i^\top$; and 2) the coefficient functions in the proposed FFLQR model can be approximated using the FPC basis functions and the estimated parameter of the QR model constructed based on the FPC scores. This finite-dimensional QR model can be easily estimated using the available package "quantreg" [19].

The performance of the proposed FFLQR depends on the truncation parameters $K_{\mathcal{Y}}$ and $K_{\mathcal{X}_m}$ (for $m = 1, \ldots, M$). In this paper, the optimum numbers of truncation constants are determined based on the BIC.

Let us denote all the possible models by

$$\mathcal{J} = \left\{ K_{\mathcal{Y}}, K_{\mathcal{X}_m} | K_{\mathcal{Y}} = 1, \ldots, K_{\mathcal{Y}_{\max}}, \; K_{\mathcal{X}_m} = 1, \ldots, K_{\mathcal{X}_{m,\max}} \right\}.$$

Then, we assume that there exists an optimal estimate for the FFLQR associated with $K_{\mathcal{Y}_0}, \; K_{\mathcal{X}_{m_0}} \in \mathcal{J}$ as follows:

$$\widehat{Q}_\tau \left[\mathcal{Y}_i(t) | \boldsymbol{\mathcal{X}}_i \right] = \sum_{m=1}^M \int_0^1 \mathcal{X}_{im}(s) \widehat{\beta}_{m\tau}^{(K_{\mathcal{Y}_0}, K_{\mathcal{X}_0})}(s,t) ds,$$

where $\widehat{\beta}_{m\tau}^{(K_{\mathcal{Y}_0}, K_{\mathcal{X}_0})}(s,t)$ is the estimated coefficient function with $K_{\mathcal{Y}_0}$ and $K_{\mathcal{X}_{m_0}}$ and $K_{\mathcal{X}_0} = \left\{ K_{\mathcal{X}_{1_0}}, \ldots, K_{\mathcal{X}_{M_0}} \right\}$. For computational simplicity, we apply the

same truncation constant $K_{\mathcal{X}}$ to all the predictors. For each combination of $K_{\mathcal{Y}}$ and $K_{\mathcal{X}}$, the estimates of coefficient functions are obtained by

$$\widehat{\boldsymbol{\beta}}_{\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t) =$$

$$\underset{\beta_{1\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t),\ldots,\beta_{M\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t)}{\arg\min} \sum_{i=1}^{N} \rho_{\tau} \left[\mathcal{Y}_{i}(t) - \sum_{m=1}^{M} \int_{0}^{1} \mathcal{X}_{im}(s)\beta_{m\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t)ds \right],$$

where $\widehat{\boldsymbol{\beta}}_{\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t) = \left[\widehat{\beta}_{1\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t),\ldots,\widehat{\beta}_{M\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t)\right]^{\top}$. Then, according to the definition of BIC given by [33], we obtain the following BIC for the FFLQR:

$$\text{BIC}(K_{\mathcal{Y}},K_{\mathcal{X}}) = \left\| \ln \left[\sum_{i=1}^{n} \rho_{\tau}\left(\mathcal{Y}_{i}(t) - \sum_{m=1}^{M} \int_{0}^{1} \mathcal{X}_{im}(s)\widehat{\beta}_{m\tau}^{(K_{\mathcal{Y}},K_{\mathcal{X}})}(s,t)ds\right)\right]\right\|_{\mathcal{L}_{2}}$$
$$+ \omega \ln(n),$$

where $\omega = K_{\mathcal{Y}} + K_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{L}_{2}}$ denotes the $\mathcal{L}_{2}$ norm. In short, $K_{\mathcal{Y}_{\max}} \times K_{\mathcal{X}_{\max}}$ different models (i.e., $K_{\mathcal{Y}} = 1,\ldots,K_{\mathcal{Y}_{\max}}$, $K_{\mathcal{X}} = 1,\ldots,K_{\mathcal{X}_{\max}}$) are built and the optimum $K_{\mathcal{Y}}$ and $K_{\mathcal{X}}$ ($\widetilde{K}_{\mathcal{Y}}$ and $\widetilde{K}_{\mathcal{X}}$, respectively) are determined by

$$\left(\widetilde{K}_{\mathcal{Y}}, \widetilde{K}_{\mathcal{X}}\right) = \underset{K_{\mathcal{Y}},K_{\mathcal{X}}}{\arg\min} \text{BIC}(K_{\mathcal{Y}},K_{\mathcal{X}}).$$

2.1 Variable selection procedure

When considering the proposed FFLQR model (2.1), the vector of functional predictors $\boldsymbol{\mathcal{X}}$ may include too many variables, and not all of them may have a significant effect on the model. Consequently, there may need a variable selection procedure to determine only the significant functional predictors. To this end, we consider a forward variable selection procedure together with the extension of BIC introduced by [22] to the proposed method. Denote by $\mathcal{D} = \{d_{1},\ldots,d_{a}\} \subset \{1,\ldots,M\}$ the candidate model having the functional predictors $\{\mathcal{X}_{d_{1}},\ldots,\mathcal{X}_{d_{a}}\}$ in the model. Using the check loss function, the estimates of coefficient function for this model are given by

$$\widehat{\boldsymbol{\beta}}_{\tau}^{\mathcal{D}}(s,t) = \underset{\beta_{d_{1}\tau}^{\mathcal{D}}(s,t),\ldots,\beta_{d_{a}\tau}^{\mathcal{D}}(s,t)}{\arg\min} \sum_{i=1}^{N} \rho_{\tau}\left[\mathcal{Y}_{i}(t) - \sum_{m\in\mathcal{D}} \int_{0}^{1} \mathcal{X}_{im}(s)\beta_{m\tau}^{\mathcal{D}}(s,t)ds\right],$$

where $\widehat{\boldsymbol{\beta}}_{\tau}^{\mathcal{D}}(s,t) = \left[\widehat{\beta}_{d_{1}\tau}^{\mathcal{D}}(s,t),\ldots,\widehat{\beta}_{d_{a}\tau}^{\mathcal{D}}(s,t)\right]^{\top}$. Let $|\mathcal{D}|$ denote the cardinality a of $\mathcal{D}$. Then, following the definition of BIC presented by [22], the BIC for the candidate model is defined as follows:

$$\text{BIC}(\mathcal{D}) = \left\| \ln \left[\sum_{i=1}^{n} \rho_{\tau}\left(\mathcal{Y}_{i}(t) - \sum_{m\in\mathcal{D}} \int_{0}^{1} \mathcal{X}_{im}(s)\widehat{\beta}_{m\tau}^{\mathcal{D}}(s,t)ds\right)\right]\right\|_{\mathcal{L}_{2}} + |\mathcal{D}|\frac{\ln(n)}{2n}.$$

Accordingly, we present the considered variable selection procedure as follows.

Step 1. Construct M - FFLQR models based on the typical functional response and a functional predictor variable:

$$Q_\tau \left[\mathcal{Y}_i(t) | \mathcal{X}_{im} \right] = \int_0^1 \mathcal{X}_{im}(s) \beta_{m\tau}(s,t) ds, \quad m = 1, \ldots, M$$

and calculate the $\mathrm{BIC}\,(\mathcal{D})$ for each of these models. Then, determine the initial model to having smallest $\mathrm{BIC}\,(\mathcal{D})$. Let $\mathcal{X}_i^{(1)}(s)$ and $\mathrm{BIC}^{(1)}\,(\mathcal{D})$ denote the functional predictor in the initial model and its $\mathrm{BIC}\,(\mathcal{D})$ value, respectively.

Step 2. Construct $(M-1)$ - FFLQR model based on the common functional response and a vector of functional predictors:

$$Q_\tau \left[\mathcal{Y}_i(t) | \boldsymbol{\mathcal{X}}_{im} \right] = \sum_{m=1}^2 \int_0^1 \mathcal{X}_{im}(s) \beta_{m\tau}(s,t) ds,$$

where $\boldsymbol{\mathcal{X}}_{im} = \left[\mathcal{X}_i^{(1)}(s), \mathcal{X}_{im}(s) \right]$ and $\mathcal{X}_{im}(s) \neq \mathcal{X}_i^{(1)}(s)$, and calculate the $\mathrm{BIC}\,(\mathcal{D})$ for each of these models. The vector of functional predictors having the smallest $\mathrm{BIC}\,(\mathcal{D})$, say $\boldsymbol{\mathcal{X}}^{(2)}$, is chosen as the predictor vector for the current model if $\mathrm{BIC}^{(2)}\,(\mathcal{D}) / \mathrm{BIC}^{(1)}\,(\mathcal{D}) < 0.95$, where $\mathrm{BIC}^{(2)}\,(\mathcal{D})$ is the computed $\mathrm{BIC}\,(\mathcal{D})$ when $\boldsymbol{\mathcal{X}}^{(2)}$ is used in the current model. In other words, the second predictor enters the model if it contributes at least 5% to the model (the threshold value 5% is determined via an extensive Monte Carlo experiment). Repeat this procedure until all the significant functional predictors are included in the model.

Note that applying the above variable selection procedure, together with the procedure of determination of optimum truncation constants discussed in Section 2, may not be computationally efficient when there are a large number of predictors to be considered. In our numerical analyses, we first determine the significant variables using fixed truncation constants, say $K_{\mathcal{Y}} = K_{\mathcal{X}} = 2$. Then, the optimum values of $K_{\mathcal{Y}}$ and $K_{\mathcal{X}}$ are determined based on the significant functional predictors. In this context, we perform several simulations (but not reported in the paper for the sake of space), and the results have shown that the choices of $K_{\mathcal{Y}}$ and $K_{\mathcal{X}_m}$ do not have a significant effect on the determination of significant predictors. In high-dimensional settings, a parallel implementation of the determination of optimum truncation constants and significant variables (such as using "doParallel" [7] and "foreach" [27] packages) can be used to reduce computational costs.

2.2 Computational framework

In this section, we summarize the computational framework of the proposed method. We note that all the numerical calculations for the proposed method are performed using version 4.1.1 on an Intel Core i7 6700HQ 2.6 GHz PC. In the proposed method, first, the FPC decomposition of the functional variables are obtained using some available functions such as `create.bspline.basis`

and `pca.fd` in the package "fda" [29]. Then, the `rqs.fit` function in the package "quantreg" [19] along with the computed principal component scores is used to approximate the coefficient functions in the finite-dimensional space. Finally, the functional forms of the estimated coefficient functions are obtained using the estimated FPC basis functions. We build some functions for other calculations such as BIC and variable selection procedure. An example code (with definitions) for the proposed method can be found at `https://github.com/UfukBeyaztas/FFLQR`.

3   Numerical results

In this section, several Monte Carlo experiments under different scenarios and empirical data analysis are performed to investigate the finite-sample predictive performance of our proposed FFLQR model. We compare our results with those calculated via the LS [32], functional partial least squares (FPLS) [3], and classical FPC based FFR models. The LS method uses a general basis function expansion method (B-spline basis function is used in our calculations) to project the functional variables into the finite-dimensional space. Then, the LS loss function is used to estimate the model parameters. In FPLS, an iterative procedure similar to FPC is used to obtain the FPLS basis functions and corresponding coefficients. Then, the LS estimator is used to approximate the functional coefficients.

3.1   Monte Carlo experiments

In the Monte Carlo experiments, we perform MC = 200 Monte Carlo simulation runs, each of which consists of $M = 5$ functional predictors $\boldsymbol{\mathcal{X}}_i = \{\mathcal{X}_{i1}(s), \ldots, \mathcal{X}_{i5}(s)\}$ with $i = 1, 2, \ldots, 500$ observed at 100 equally spaced points in the unit interval $s \in [0, 1]$. We consider the following process, which is a modified version of those considered by [24], to generate the functional predictors:

$$\mathcal{X}_{im}(s) = 10 + \sum_{j=0}^{\ell} \frac{V_{i,m+j}(s)}{\sqrt{\ell + 1}}, \tag{3.1}$$

where $[V_{i,1}(s), V_{i,2}(s), \ldots, V_{i,9}(s)]$ are generated from a Gaussian process with mean-zero and a positive definite covariance function $\boldsymbol{\Sigma}_V(s, s') = e^{-100(s-s')^2}$. In (3.1), the parameter ℓ ($\ell > 0$) denotes the lag parameter that controls the correlation level between predictor functions. Herein, the large ℓ corresponds to high correlations between the predictors. In our analyses, we consider $\ell = 4$ to generate functional predictors. In addition, we consider the following processes to generate the smooth bivariate regression coefficients:

$$\beta_1(s,t) = (1-s)^2(t-0.5)^2, \quad \beta_2(s,t) = e^{-3(s-1)^2 - 5(t-0.5)^2},$$
$$\beta_3(s,t) = e^{-5(s-0.5)^2 - 5(t-0.5)^2} + 8e^{-5(s-1.5)^2 - 5(t-0.5)^2},$$
$$\beta_4(s,t) = \sin(1.5\pi s)\sin(\pi t), \quad \beta_5(s,t) = \sqrt{st}.$$

Then, the functions of the response variable are generated as follows:

$$\mathcal{Y}_i(t) = \sum_{m \in \mathcal{D}} \int_0^1 \mathcal{X}_{im}(s)\beta_m(s,t)ds + \epsilon_i(t),$$

where $\mathcal{D} = \{2, 4, 5\}$ denotes the index set of significant functional predictors. The error functions $\epsilon_i(t)$ are generated from the Ornstein-Uhlenbeck process:

$$\epsilon_i(t) = \gamma + [\epsilon_0(t) - \gamma]\, e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-u)}dW_u,$$

where γ, θ and $\sigma > 0$ are real constants, and W_u denotes the Wiener process. Given $\epsilon_0(t)$, which is the initial value of $\epsilon_i(t)$ and independently taken from W_u, the functions $\epsilon_i(t)$ are generated from the joint distribution of $\{\epsilon_i(t)\}_{i=1}^n$. To investigate the robustness of the QR over the mean regression, we consider two distributions for the error terms: (i) $N(0,1)$ (symmetric) and (iii) $\chi^2_{(1)}$ (skewed). In addition, we consider two signal-to-noise ratio levels: $\sigma = 0.1$ and $\sigma = 1$. Examples of the generated functions for the response and predictor variables are presented in Figure 1.
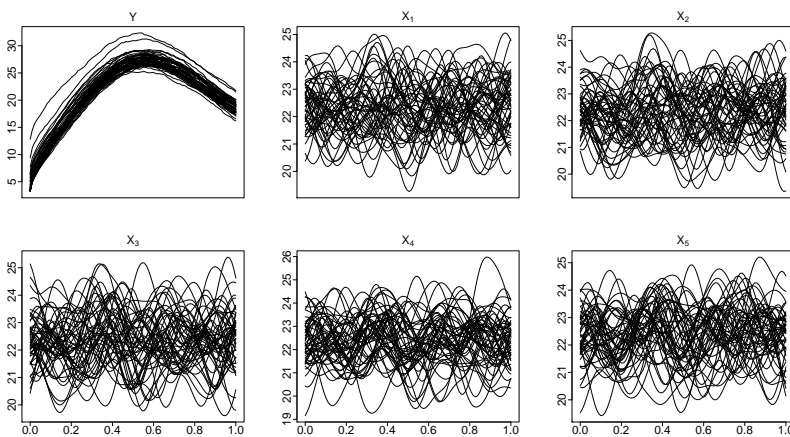

Figure 1. Plots of the generated 50 functions of the functional variables when $\sigma = 1$ and error terms follow $\chi^2_{(1)}$ distribution. The functions of the response variable are generated as $\mathcal{Y}_i(t) = \sum_{m \in \mathcal{D}} \int_0^1 \mathcal{X}_{im}(s)\beta_m(s,t)ds + \epsilon_i(t)$, where $\mathcal{D} = \{2, 4, 5\}$.

To further investigate the robust nature of the proposed method, we consider a second scenario for the response variable, where the observations associated with the response variable are contaminated by outliers at $[5\%, 10\%, 20\%]$ contamination levels. Since the QR characterizes the conditional distribution of the response variable, only the generated functional response is contaminated by the outliers. The outlier-contaminated response functions are generated by contaminating $n \times [5\%, 10\%, 20\%]$ randomly selected functions by a random function; $\widetilde{\mathcal{Y}}_i = \mathcal{Y}_i + |N(10, 0.04)|$.

The mean squared prediction error (MSPE) metric is considered to compare the methods' predictive performance. In doing so, the generated data are first

divided into training and test samples with sizes $n_{\text{train}} = 200$ and $n_{\text{test}} = 300$. Then, the training samples are used to construct models, and the functions in the test sample are used for validation. The MSPE is computed as follows:

$$\text{MSPE} = \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \left\| \mathcal{Y}_i(t) - \widehat{\mathcal{Y}}_i(t) \right\|_{\mathcal{L}_2}^2 ,$$

where $\widehat{\mathcal{Y}}_i(t)$ is the prediction of the ith observation in the test sample and is calculated using (2.3). To construct pointwise prediction intervals for the response function in the test sample, we employ a case-sampling-based bootstrap. In doing so, $R = 100$ bootstrap pseudo-samples, $(\mathcal{Y}^*, \boldsymbol{\mathcal{X}}^*)$, with sizes n_{train} are drawn with replacement from $(\mathcal{Y}, \boldsymbol{\mathcal{X}})$. Then, based on the bootstrap pseudo-samples, the smooth bivariate functional coefficients are estimated to obtain $R = 100$ sets of bootstrap replicates of the predicted response functions $\left\{ \widehat{\mathcal{Y}}_i^{*,1}(t), \ldots, \widehat{\mathcal{Y}}_i^{*,R}(t) \right\}$. Finally, the $100(1 - \alpha)\%$ bootstrap prediction interval for $\mathcal{Y}_i(t)$ is obtained taking the $\alpha/2^{\text{th}}$ and $1 - \alpha/2^{\text{th}}$ quantiles of the generated R sets of bootstrap replicates of the ith predicted response:

$$\left[Q_{\alpha/2}^i(t), Q_{1-\alpha/2}^i(t) \right],$$

where $Q_\alpha^i(t)$ denote the αth quantile of the bootstrap replicates. We consider the coverage probability deviance (CPD) and interval score (score) metrics to evaluate the performance of the methods:

$$\text{CPD} = (1 - \alpha) - \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \Vdash \left\{ Q_{\alpha/2}^i(t) \leq \mathcal{Y}_i(t) \leq Q_{1-\alpha/2}^i(t) \right\},$$

$$\text{score} = \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \left\| \left[\left\{ Q_{1-\alpha/2}^i(t) - Q_{\alpha/2}^i(t) \right\} + \frac{2}{\alpha} \left(Q_{\alpha/2}^i(t) - \mathcal{Y}_i(t) \right) \right.\right.$$

$$\left.\left. \Vdash \left\{ \mathcal{Y}_i(t) < Q_{\alpha/2}^i(t) \right\} + \frac{2}{\alpha} \left(\mathcal{Y}_i(t) - Q_{1-\alpha/2}^i(t) \right) \Vdash \left\{ \mathcal{Y}_i(t) > Q_{1-\alpha/2}^i(t) \right\} \right] \right\|_{\mathcal{L}_2}.$$

The CPD is the absolute difference between the nominal and empirical coverage probabilities. The small CPD value indicates that the prediction interval constructed by the method covers most of the observations in the test sample. In contrast, its large values indicate the prediction interval fails to cover observations. The interval score simultaneously evaluates the coverage probability and width of the constructed prediction interval, and a smaller interval score corresponds to a sharp prediction interval (more accurate and narrower). Note that we also calculate pointwise prediction intervals for the response variable using the proposed FFLQR by fitting the same model on the generated data for two quantile levels $\tau_1 = \alpha/2$ and $\tau_2 = (1 - \alpha/2)$, and compare its performance with the metrics obtained from the bootstrap method.

Before presenting results, we note that we compare our proposed FFLQR method with the LS, FPLS, FPC under three models: 1) the *full model*, where all five predictors are included in the model; 2) the *true model*, where only the significant predictors in the index set $\mathcal{D} = \{2, 4, 5\}$ are included to the

model; and 3) the *selected model*, where the variables used in the model are determined using the variable selection procedure discussed in Section 2.1. To make a logical comparison, throughout the simulations, we set $\tau = 0.5$ under our FFLQR model, and we focus on evaluating function-on-function median regression versus mean regression.
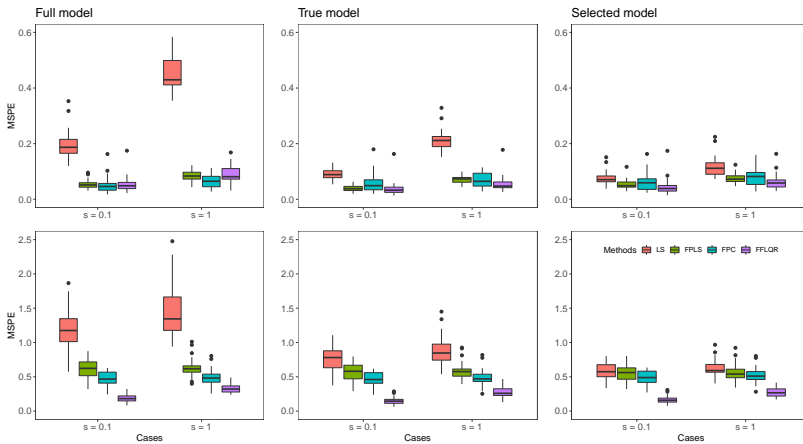

Figure 2. Predictive model performances: Calculated MSPE values of the LS, FPLS, FPC, and FFLQR methods when no outliers are present in the data; full model (first column), true model (second column), and selected model (third column). Data are generated based on two cases where the signal-to-noise ratio levels are $s = 0.1$ and $s = 1$ and two error distributions; $N(0,1)$ (first row) and $\chi^2_{(1)}$ (second row).

The calculated MSPE values when no outliers are present in the data are given in Figure 2. The results demonstrate that: 1) Both true and selected models produce significantly smaller MSPE values than those of full models for all methods. An interesting but not too surprising result produced by our simulations is that, compared with true models, selected models tend to produce smaller MSPEs. This result could be because some true coefficient functions have a smaller effect on the response function than others. In such a case, to estimate the true model, the functional predictors corresponding to these coefficient functions have to be estimated, which increases the number of coefficients to be estimated and reduces the prediction accuracy. See, for example [24] for similar results. 2) For all settings, all the methods produce larger MSPE values with increasing signal-to-noise ratio levels. 3) When the error terms follow symmetric distribution (i.e., $N(0,1)$), the proposed FFLQR model produces competitive performance compared with the existing methods. However, when the errors follow the non-normal $\chi^2_{(1)}$ distribution, the proposed method significantly outperforms the LS, FPLS, and FPC methods.

When no outliers are present in the data, the calculated CPD and interval score values are shown in Figures 3 and 4, respectively.

The proposed method produces improved CPD and score values, among others, under the symmetric errors from these figures. This result indicates that the proposed method produces more accurate bootstrap prediction inter-
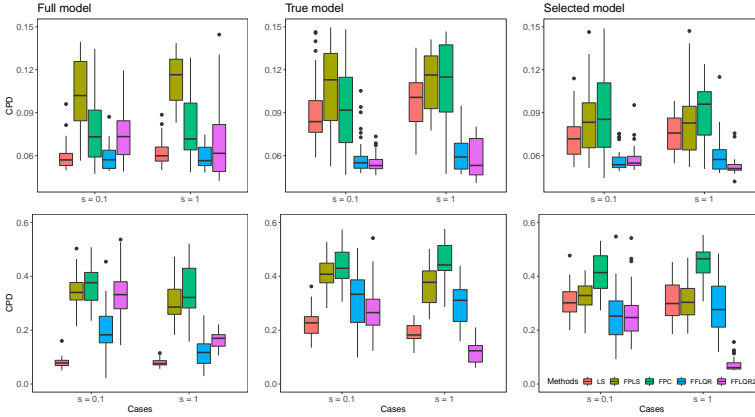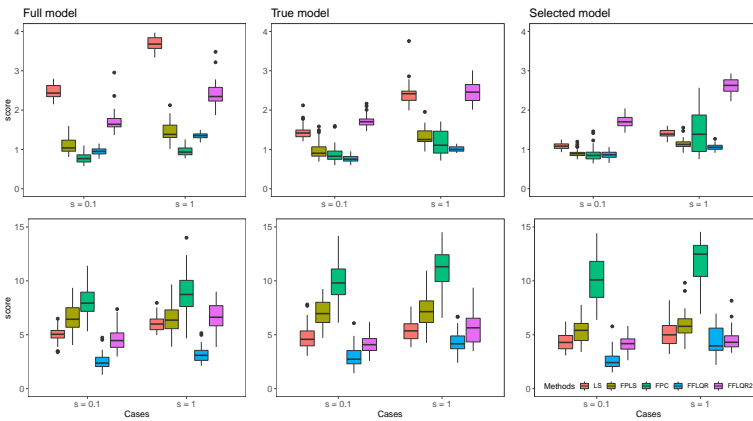
Figure 3. Predictive model performances: Calculated CPD values of the LS, FPLS, FPC, and FFLQR methods when no outliers are present in the data; full model (first column), true model (second column), and selected model (third column). Data are generated based on two cases where the signal-to-noise ratio levels are $s = 0.1$ and $s = 1$ and two error distributions; $N(0, 1)$ (first row) and $\chi^2_{(1)}$ (second row). "FFLQR2" denotes the performance metrics obtained directly from the proposed FFLQR.

Figure 4. Predictive model performances: Calculated score values of the LS, FPLS, FPC, and FFLQR methods when no outliers are present in the data; full model (first column), true model (second column), and selected model (third column). Data are generated based on two cases where the signal-to-noise ratio levels are $s = 0.1$ and $s = 1$ and two error distributions; $N(0, 1)$ (first row) and $\chi^2_{(1)}$ (second row). "FFLQR2" denotes the performance metrics obtained directly from the proposed FFLQR.

vals with narrower prediction interval lengths compared with other methods. The prediction intervals obtained directly from the proposed FFLQR model (not bootstrap-based) produce competitive or even better CPD values than the bootstrap-based prediction intervals constructed by the proposed method. However, they have generally larger score values than the bootstrap prediction intervals. In other words, the non-bootstrap-based prediction intervals are

wider than those of bootstrap-based prediction intervals.

The calculated MSPE, CPD, and interval score values, when the data are contaminated by outliers, are presented in Figure 5.
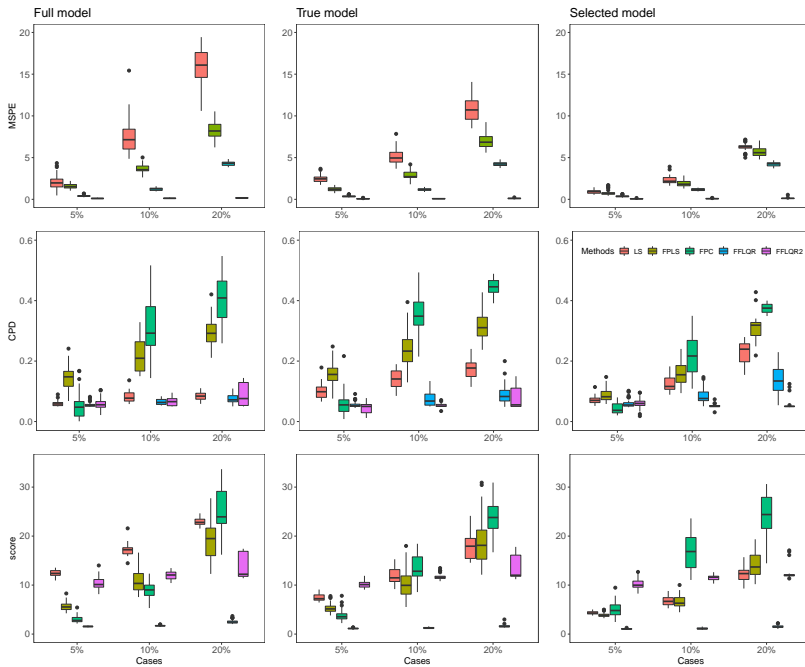

Figure 5. Predictive model performances: Calculated MSPE (first row), CPD (second row), and score (third row) values of the LS, FPLS, FPC, and FFLQR methods when [5%, 10%, 20%] of the data are contaminated by magnitude outliers; full model (first column), true model (second column), and selected model (third column). Data are generated when the signal-to-noise ratio level is [$s = 1$ and the error terms follow $N(0, 1)$ distribution. "FFLQR2" denotes the interval scores obtained directly from the proposed FFLQR.

The results indicate that the proposed method produces smaller prediction errors than the existing methods for all models and contamination levels. Also, the proposed method produces significantly smaller bootstrap-based interval score values than those obtained by LS, FPLS, and FPC methods. The non-bootstrap-based prediction intervals produce competitive CPD values with larger score values (i.e., wider prediction interval lengths) to the bootstrap-based prediction intervals. Therefore, the bootstrap procedure seems a better approach to construct prediction intervals.

We also compare the proposed method with LS, FPLS, and FPC methods regarding their computing times. In doing so, the data are generated under $N(0, 1)$ errors with signal-to-ratio level $\sigma = 1$. Then, all the methods under the selected model are performed, and the computing times of the methods are recorded. Our results indicate that the proposed method requires more computing time compared with the other three methods. For one Monte Carlo experiment, our proposed method requires approximately nine seconds, while other methods require approximately four seconds.

3.2 Data analysis: Drosophila life cycle gene expression time-series data

We consider the microarray data obtained by a cDNA microarray experiment, reported by [1]. The original microarray data consist of gene expression levels for 4028 genes involved in Drosophila melanogaster's life cycle during four stages of morphogenesis: embryo, larva, pupa, and adult. [1] reported gene expression patterns for nearly one-third of all Drosophila genes and identified the groups of co-expressed genes. In this study, we consider three specific groups of genes involved during the first three stages in the life cycle of Drosophila, e.g., embryo, larva, and pupa: 1) "transient early zygotic" consisting of 21 genes, 2) "muscle-specific" consisting of 23 genes, and 3) "eye-specific" consisting of 33 genes. Transient early zygotic genes, which are the blastoderm-specific genes, are expressed at high levels only during the critical period of development when cellularization of the syncytial blastoderm embryo occurs [1]. These genes have a single-peak expression pattern in the embryo stage. Muscle-specific genes, which are tissue-specific genes, present a two-peak expression pattern that coincides with larval and pupa muscle development. Eye-specific genes, on the other hand, have a single-peak expression pattern in the pupa stage. The graphical displays of the gene expression profiles for all three groups of genes are presented in Figure 6.
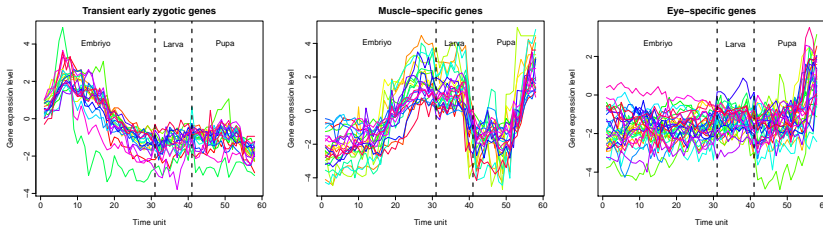

Figure 6. Graphical displays of the gene expression profiles; transient early zygotic genes (left panel), muscle-specific genes (middle panel), and eye-specific genes (right panel).

Our aim with the gene expression levels is to explore the relationship between temporal patterns of gene expression levels measured at different development stages of Drosophila melanogaster. More precisely, our goal is to predict the later life gene expression profiles (pupa stage) of the three groups above of genes based on the gene expression patterns measured at embryo and larval stages. Using a mean regression-based FFR model, [28] studied the 23 muscle-specific gene expression levels to explore the dependence of adult gene expression patterns on larval patterns. In addition, they studied 27 strictly maternal gene expression levels and assessed gene expression pattern dependencies between the early embryo and the female pattern adult germline. The results produced by the analyses of [28] demonstrated that the gene expression profiles for pupa and adult phases are strongly related to the profiles of the same genes obtained during the embryo phase. Furthermore, their results showed a positive relationship in expression for muscle development-related genes and a negative relationship for Drosophila's strictly maternal genes.

In this study, we consider the following FFR model to explore the dependence of pupa gene expression patterns on embryo and larval patterns:

$$\mathcal{Y}(t) = \beta_0(t) + \int_{s_1=1}^{31} \mathcal{X}_1(s_1)\beta_1(s_1,t)ds_1 + \int_{s_2=1}^{10} \mathcal{X}_2(s_2)\beta_2(s_2,t)ds_2, \;\; t \in [1,17],$$

where $\mathcal{Y}(t)$, $\mathcal{X}_1(s_1)$, and $\mathcal{X}_2(s_2)$ denote the gene expression levels measured at pupa, embryo, and larval stages, respectively. From Figure 6, the muscle-specific genes are generally expressed at similar levels. On the other hand, some of the transient early zygotic and eye-specific genes are observed at higher or lower levels (potential outliers). Therefore, our proposed FFLQR model may robustly predict these groups' later life gene expression profiles compared with the existing models. To compare the prediction performance of our proposed method with the LS, FPLS, and FPC-based FFR models, we repeat the following procedure 100 times. For each group of genes, the entire datasets are randomly divided into two parts: 1) roughly, half of the datasets are used to construct models and 2) the remaining observations are used for validation.

For each replication, the MSPE is computed for each method. In addition, We apply the case-sampling-based bootstrap to construct pointwise prediction intervals for the later life gene expression profiles. We calculate performance metrics based on five FPCs and B-spline basis expansion functions. Note that, for the gene expression levels, we calculate the performance metrics under two models: the full model, which includes both $\mathcal{X}_1$ and $\mathcal{X}_2$ in the model, and the selected model, which includes only the significant predictor variable(s) determined by the proposed variable selection procedure.

The results produced by our analyses are presented in Figure 7. This figure shows that the selected model produces smaller prediction performance metrics than the full model for all the methods. When considering the performance of selected models, compared with the LS FPLS, and FPC methods, FFLQR produces smaller MSPE values for transient early zygotic and eye-specific genes. In contrast, it produces slightly larger error values than the FPLS and FPC for muscle-specific genes. This finding is because some of the transient early zygotic and eye-specific genes are observed at higher or lower levels (potential outliers). At the same time, all the muscle-specific genes are generally expressed at similar levels. Therefore, our proposed method produces robust results for the transient early zygotic and eye-specific genes. In addition, the proposed method produces generally smaller CPD and interval score values than the LS, FPLS, and FPC methods.

We construct a model for all genes using the observed gene expression profiles to determine the significant variables. For transient early zygotic genes, the only larval stage ($\mathcal{X}_2$) is selected into the final model by the LS, FPLS, and FPC methods. In contrast, only the proposed method selected the embryo ($\mathcal{X}_1$) as significant. For muscle-specific and eye-specific genes, on the other hand, the only larval stage ($\mathcal{X}_2$) is selected as significant by the LS, FPC, and FFLQR methods, while both embryo ($\mathcal{X}_1$) and larval ($\mathcal{X}_2$) stages are selected as significant by the FPLS method. The estimated regression coefficient functions by the proposed method for all genes are presented in Figure 8.
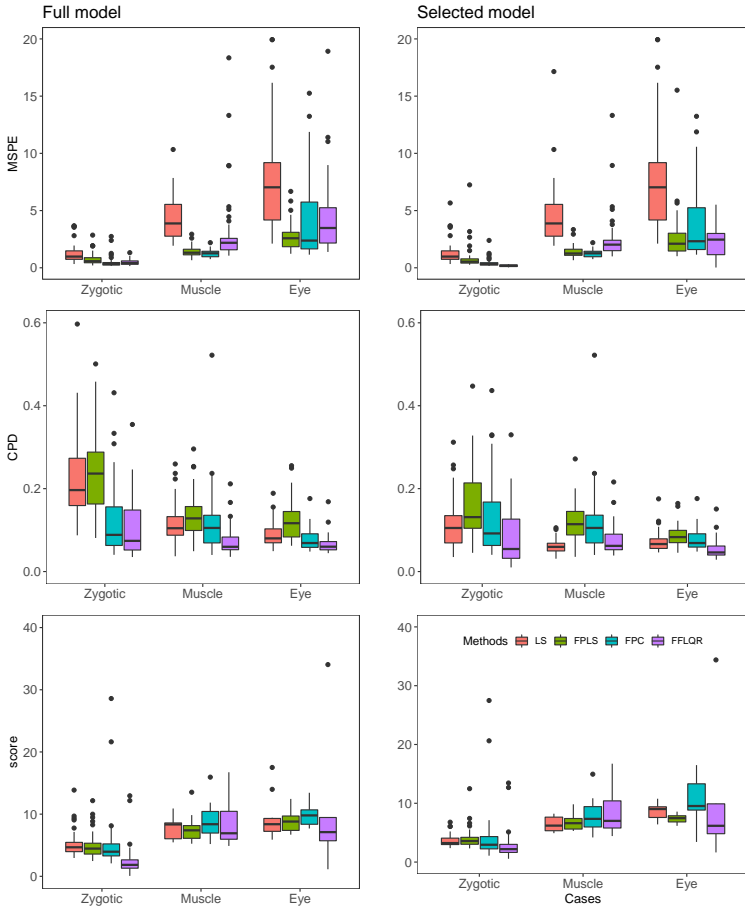
Figure 7. Predictive model performances: Calculated MSPE (first row), CPD (second row), and interval score (third row) values of the LS, FPLS, FPC, and FFLQR methods for the Drosophila life cycle gene expression time-series data; full model (first column) and selected model (second row).
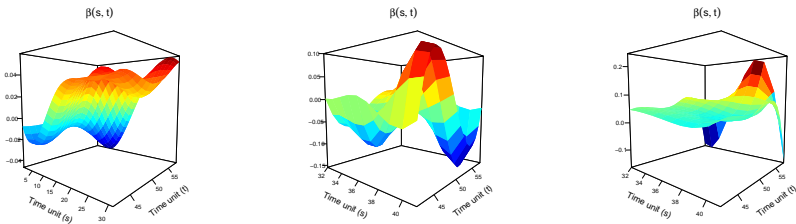

Figure 8. Estimated regression coefficient functions $\widehat{\beta}_\tau(s, t)$ when $\tau = 0.5$ for the Drosophila life cycle gene expression time-series data; transient early zygotic genes (left panel), muscle-specific genes (middle panel), and eye-specific genes (right panel).

4   Conclusions

We propose a QR approach for the FFR models, an extension of the traditional FFR model to the QR framework. In our proposed method, all the functional objects are first transformed into a finite-dimensional space using the FPC method to overcome the ill-posed problem caused by the infinite-dimensional nature of the functional variables. It approximates the regression coefficient functions using the FPC decomposition of the functional response and predictor variables. A forward variable selection procedure and the proposed FFLQR method are introduced to determine the significant functional predictors for the model. Moreover, we employ a nonparametric bootstrap approach to investigate further the uncertainty of predictions produced by the proposed FFLQR method. Our method's finite-sample performance is investigated via several Monte Carlo experiments and empirical data analysis, and they are compared favorably with existing methods.

All the numerical analyses performed in our study have demonstrated that the proposed FFLQR method produces superior performance over the existing function-on-function mean regression models when the error terms follow a non-normal distribution or in the presence of outliers.

Acknowledgements

References

[1] M.N. Arbeitman, E.E.M. Furlong, F. Imam, E. Johnson, B.H. Null, B.S. Baker, M.A. Krasnow, M.P. Scott, R.W. Davis and K.P. White. Gene expression during the life cycle of Drosophila melanogaster. *Science*, **297**(5590):2270–2275, 2002. https://doi.org/0.1126/science.1072152.

[2] U. Beyaztas and H.L. Shang. A comparison of parameter estimation in function-on-function regression. *Communications in Statistics - Simulation and Computation*, **0**(0):1–31, 2020. https://doi.org/10.1080/03610918.2020.1746340.

[3] U. Beyaztas and H.L. Shang. On function-on-function regression: Partial least squares approach. *Environmental and Ecological Statistics*, **27**(1):95–114, 2020. https://doi.org/10.1007/s10651-019-00436-1.

[4] H. Cardot, C. Crambes and P. Sarda. Quantile regression when the covariates are functions. *Journal of Nonparametric Statistics*, **17**(7):841–856, 2005. https://doi.org/10.1080/10485250500303015.

[5] K. Chen and H-G. Müller. Conditional quantile analysis when covariates are functions, with application to growth data. *Journal of the Royal Statistical Society: Series B*, **74**(1):67–89, 2012. https://doi.org/10.1111/j.1467-9868.2011.01008.x.

[6] J.-M. Chiou, Y.-F. Yang and Y.-T. Chen. Multivariate functional linear regression and prediction. *Journal of Multivariate Analysis*, **146**:301–312, 2016. https://doi.org/10.1016/j.jmva.2015.10.003.

[7] Microsoft Corporation and S. Weston. *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*. R package version 1.0.16, 2020. Available from Internet: `https://CRAN.R-project.org/package=doParallel`.

[8] D. Şentürk and H.-G. Müller. Generalized varying coefficient models for longitudinal data. *Biometrika*, **95**(3):653–666, 2008. https://doi.org/10.1093/biomet/asn006.

[9] A. Cuevas. A partial overview of the theory of statistics with functional data. *Journal of Statistical Planning and Inference*, **147**:1–23, 2014. https://doi.org/10.1016/j.jspi.2013.04.002.

[10] F. Ferraty, A. Rabhi and P. Vieu. Conditional quantiles for dependent functional data with application to the climatic El Nino phenomenon. *Sankhyā: The Indian Journal of Statistics*, **67**(2):378–398, 2005.

[11] F. Ferraty and P. Vieu. *Nonparametric Functional Data Analysis*. Springer, New York, 2006. https://doi.org/10.1007/0-387-36620-2.

[12] P. Hall and J.L. Horowitz. Methodology and convergence rates for functional linear regression. *The Annals of Statistics*, **35**(1):70–91, 2007. https://doi.org/10.1214/009053606000000957.

[13] J. Harezlak, B.A. Coull, N.M. Laird, S.R. Magari and D.C. Christiani. Penalized solutions to functional regression problems. *Computational Statistics and Data Analysiss*, **51**(10):4911–4925, 2007. https://doi.org/10.1016/j.csda.2006.09.034.

[14] L. Horváth and P. Kokoszka. *Inference for Functional Data with Applications*. Springer, New York, 2012. https://doi.org/10.1007/978-1-4614-3655-3.

[15] T. Hsing and R. Eubank. *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. John Wiley & Sons, Chennai, India, 2015. https://doi.org/10.1002/9781118762547.

[16] H. Hullait, D.S. Leslie, N.G. Pavlidis and S. King. Robust function-on-function regression. *Technometrics*, **63**(3):396–409, 2021. https://doi.org/10.1080/00401706.2020.1802350.

[17] A.E. Ivanescu, A.M. Staicu, F. Scheipl and S. Greven. Penalized function-on-function regression. *Computational Statistics*, **30**(2):539–568, 2015. https://doi.org/10.1007/s00180-014-0548-4.

[18] R. Koenker. *Quantile Regression*. Cambridge University Press, Cambridge, 2005. https://doi.org/10.1017/CBO9780511754098.

[19] R. Koenker. *quantreg: Quantile Regression*. R package version 5.85, 2021. Available from Internet: `https://CRAN.R-project.org/package=quantreg`.

[20] R. Koenker and Jr. G. Bassett. Regression quantiles. *Econometrica*, **46**(1):33–50, 1978. https://doi.org/10.2307/1913643.

[21] P. Kokoszka and M. Reimherr. *Introduction to Functional Data Analysis*. CRC Press, Boca Raton, 2017. https://doi.org/10.1201/9781315117416.

[22] E.R. Lee, H. Noh and B.U. Park. Model selection via Bayesian information criterion for quantile regression models. *Journal of the American Statistical Association: Theory and Methods*, **109**(505):216–229, 2014. https://doi.org/10.1080/01621459.2013.836975.

[23] Y. Liu, M. Li and J.S. Morris. Function-on-scalar quantile regression with application to mass spectrometry proteomics data. *The Annals of Applied Statistics*, **114**(2):521–541, 2020. https://doi.org/10.1214/19-AOAS1319.

[24] R. Luo and X. Qi. Interaction model and model selection for function-on-function regression. *Journal of Computational and Graphical Statistics*, **28**(2):309–322, 2019. https://doi.org/10.1080/10618600.2018.1514310.

[25] H. Ma, T. Li, H. Zhu and Z. Zhu. Quantile regression for functional partially linear model in ultra-high dimensions. *Computational Statistics and Data Analysis*, **129**:135–147, 2019. https://doi.org/10.1016/j.csda.2018.06.005.

[26] H. Matsui, S. Kawano and S. Konishi. Regularized functional regression modeling for functional response and predictors. *Journal of Math-for-Industry*, **1**(A3):17–25, 2009.

[27] Microsoft and S. Weston. *foreach: Provides Foreach Looping Construct*. R package version 1.5.1, 2020. Available from Internet: `https://CRAN.R-project.org/package=foreach`.

[28] H.G. Müller, J.M. Chiou and X. Leng. Inferring gene expression dynamics via functional regression analysis. *BMC Bioinformatics*, **9**(60):1–20, 2008. https://doi.org/10.1186/1471-2105-9-60.

[29] J. O. Ramsay, S. Graves and G. Hooker. *fda: Functional Data Analysis*. R package version 5.1.9, 2020. Available from Internet: `https://CRAN.R-project.org/package=fda`.

[30] J.O. Ramsay and C.J. Dalzell. Some tools for functional data analysis. *Journal of the Royal Statistical Society, Series B*, **53**(3):539–572, 1991. https://doi.org/10.1111/j.2517-6161.1991.tb01844.x.

[31] J.O. Ramsay and B.W. Silverman. *Applied Functional Data Analysis 2nd edition*. Springer, New York, 2002. https://doi.org/10.1007/b98886.

[32] J.O. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer, New York, 3 edition, 2006. https://doi.org/10.1007/b98888.

[33] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, **6**(2):461–464, 1987. https://doi.org/10.1214/aos/1176344136.

[34] Q. Tang and L. Cheng. Partial functional linear quantile regression. *Science China Mathematics*, **57**(12):2589–2608, 2014. https://doi.org/10.1007/s11425-014-4819-x.

[35] F. Yao, H.-G. Müller and J.-L. Wang. Functional linear regression analysis for longitudinal data. *The Annals of Statistics*, **33**(6):2873–2903, 2005. https://doi.org/10.1214/009053605000000660.

[36] F. Yao, S. Sue-Chee and F. Wang. Regularized partially functional quantile regression. *Journal of Multivariate Analysis*, **156**:39–56, 2017. https://doi.org/10.1016/j.jmva.2017.02.001.

[37] D. Yu, L. Kong and I. Mizera. Partial functional linear quantile regression for neuroimaging data analysis. *Neurocomputing*, **195**:74–87, 2016. https://doi.org/10.1016/j.neucom.2015.08.116.