Int. J. Adv. Eng. Pure Sci. 2025, 37(UYIK 2024 Special Issue): <1-6>
DOI: 10.7240/jeps.1496164

RESEARCH ARTICLE / ARAŞTIRMA MAKALESİ

Robust Function-on-Function Regression: A Penalized Tau-based Estimation Approach

Ufuk BEYAZTAŞ [1]

[1]*Marmara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 34722, İstanbul, Türkiye*

Abstract
This study introduces a novel penalized estimation method tailored for function-on-function regression models, combining the robustness of the Tau estimator with penalization techniques to enhance resistance to outliers. Function-on-function regression is essential for modeling intricate relationships between functional predictors and response variables across diverse fields. However, traditional methods often struggle with outliers, leading to biased estimates and diminished predictive performance. Our proposed approach addresses this challenge by integrating robust Tau estimation with penalization, promoting both robustness and parsimony in parameter estimation. Theoretical foundations of the penalized Tau estimator within function-on-function regression are discussed, along with empirical validations through simulation studies and an empirical data analysis. By incorporating penalization, our method not only ensures robust estimation of regression parameters but also promotes model simplicity, offering enhanced interpretability and generalization capabilities in functional data analysis.
Keywords: Functional data, Penalization, Regression, Tau estimator

Öz
Bu çalışma, fonksiyon-fonksiyon regresyon modellerine yönelik yeni bir cezalandırılmış tahmin yöntemini tanıtmaktadır ve Tau tahmin edicisinin sağlamlığını cezalandırma teknikleriyle birleştirerek aykırı değerlere karşı direnci artırmaktadır. Fonksiyon-fonksiyon regresyon, fonksiyonel bağımsız değişkenler ile yanıt değişkenleri arasındaki karmaşık ilişkileri modellemek için çeşitli alanlarda gereklidir. Ancak geleneksel yöntemler genellikle aykırı değerlerle başa çıkmakta zorlanır ve bu durum yanlı tahminlere ve zayıd tahmin performansına yol açar. Önerilen yaklaşımımız, sağlam Tau tahminini cezalandırma ile birleştirerek bu zorluğun üstesinden gelmekte ve parametre tahmininde hem sağlamlığı hem de tutarlılığı sağlamaktadır. Fonksiyon-fonksiyon regresyon içinde cezalandırılmış Tau tahmin edicisinin teorik temelleri tartışılmakta, simülasyon çalışmaları ve ampirik veri analizleri yoluyla ampirik doğrulamalar sunulmaktadır. Cezalandırmayı dahil ederek, yöntemimiz yalnızca regresyon parametrelerinin sağlam tahminini sağlamakla kalmaz, aynı zamanda modelin basitliğini teşvik ederek fonksiyonel veri analizinde daha iyi yorumlanabilirlik ve genelleme yetenekleri sunar.
Anahtar Kelimeler: Fonksiyonel veri, Cezalandırma, Regresyon, Tau tahmincisi

I. INTRODUCTION

Recent advances in data collection technology have markedly enhanced access to high-dimensional, intricately structured datasets, known as functional data. Consequently, there is a growing demand for analytical tools designed for functional data analysis. For a comprehensive review of the latest theoretical and practical advancements in this field, refer to [1], [2], and [3]. One notable method within this domain is function-on-function regression (FoFR), which has gained popularity for examining the relationships between a response and predictors, where both the response and predictors are represented as random curves.

Let (Y,X) denote a pair where Y is the response and X is the predictor. Here, Y and X are pressumed to be stochastic processes whose elements belong to the L_2 Hilbert space; specifically, Y∈L_2 (I) and X∈L_2 (S), where I and S are bounded and closed intervals on the real line. Consider {Y_i (t),X_i (s);i=1,…n} to be a random sample (Y,X), with t∈I and s∈S. The FoFR is then defined as follows:

$$Y_i(t) = \alpha(t) + \int_{-\infty}^{S} X_i(s)\beta(s,t)ds + \epsilon_i(t), \quad (1)$$

where α(t) is the constant function, β(s,t) is the slope function, and $\epsilon_i(t)$ is the functional noise. We pressume that this functional noise is independen of the predictor variable.

Corresponding Author: UFUK BEYAZTAŞ, **Tel:** 0212 777 33 24, **E-mail:** ufuk.beyaztas@marmara.edu.tr
Submitted: 05.06.2024, **Revised:** 15.07.2024, **Accepted:** 23.07.2024

The main objective in model (1) is to estimate the slope function β(s,t). Various methodologies have been developed for this purpose, including those by [4], [1], [5], [6], [7], and [8]. However, these methods typically rely on least squares estimation procedures, which are significantly affected by outliers—observations that deviate markedly from the bulk of the data. In the presence of outliers, least squares type estimators may cause biased estimates and unreliable inference.

To address this issue, several robust estimation methods for the slope function in model (1) have been proposed. In their work, [9] developed a regression model that maintains Fisher consistency and incorporates a decomposition method using functional principal components for the observed functions. In a different study, [10] introduced a method for robustly estimating model parameters using functional partial least squares, designed to handle outliers effectively. However, the efficacy of both approaches hinges on the selected basis dimension for the predictor. This choice dictates the degree of smoothness in the estimated functional parameter. As referenced in [7], this decision can lead to significant under-smoothing, particularly when the functional parameter inherently possesses a smoother nature in contrast to the higher-order components obtained from the partial least squares and principal component methods.

To achieve robust and smooth estimates for the regression coefficient function β(s,t), [11] recently proposed a robust penalized M-estimation strategy. Their numerical analyses demonstrated that the robust penalized M-estimator provides improved parameter estimates and model inferences in the presence of outliers compared to available methods. Conversely, the robust penalized M-estimator presented by [11] lacks the integration of both a high breakdown point and high efficiency. The breakdown point measures the estimator's resistance to the influence of outliers, whereas high efficiency refers to the estimator maintaining a variance comparable to that of the least squares estimator under normal distribution conditions, as elaborated by [12].

In this study, we propose a robust penalized tau estimator designed to yield a smooth and robust estimate for β(s,t), integrating both a high breakdown point and high-efficiency characteristics. The method represents the slope functions using a tensor product of B-spline expansion. In addition, the quadratic penalties are applied to the expansion coefficients to ensure smooth estimates. The regression functions are obtained using the tau estimator from [13], known for its high breakdown point and asymptotic efficiency under normal conditions. Our approach surpasses unpenalized estimators by ensuring a level of smoothness that mitigates overfitting. Unlike the M-estimator described by [11], our method produces estimates that maintain both high breakdown point and asymptotic efficiency under normality. The optimal smoothness degree is governed by the penalization term, and the optimum values of the smoothing parameters are determined through a grid-search approach with the Bayesian Information Criterion (BIC).

The rest of this paper is structured as follows. Section 2 introduces the proposed robust penalized tau estimator. In Section 3, the empirical performance of the proposed method is evaluated via Monte-Carlo experiments. Section 4 presents the results of empirical data analysis results. Finally, Section 5 concludes the paper.

II. METHODOLOGY

Let us consider the FoFR model in (1). To derive penalized estimates for the model's parameters, specifically $\alpha(t)$ and $\beta(s,t)$, we tackle the following minimization problem:

$$\underset{\beta_0,\beta}{\operatorname{argmin}} \sum_{i=1}^{n} \rho\left[Y_i(t) - \alpha(t) - \int_S X_i(s)\beta(s,t)ds\right] + \frac{\lambda_1}{2}J_1(\alpha) + \frac{\lambda_2}{2}J_2(\beta). \quad \textbf{(2)}$$

Here, ρ stands for a loss function, J_1 and J_2 represent penalty functions applied to α and β, respectively, and λ_1 and λ_2 serve as smoothing parameters regulating the degree of shrinkage in α and β, respectively.

To derive estimates for α and β, we initially adopt the basis representation approach for the functional random variables, akin to the methodologies outlined in [7], [11], and [14]. Initially, we express $Y_i(t) = Y_i(t_{ij})$ and $X_i(s) = X_i(s_{ir})$, where $j = 1, \ldots, M_i$ and $r = 1, \ldots, G_i$, representing the number of observations on the response and predictor, respectively. For the basis representation of the functional objects in (2), we presume that $\beta_0(t)$ follows a B-spline basis expansion with K_0 basis expansion functions; $\alpha(t) = \sum_{k=1}^{K_0} \alpha_k \phi_k(t)$, where $\phi_k(t)$ (for $k = 1, \ldots, K_0$) denotes the B-spline basis expansion function, and α_k signifies the expansion coefficient. Additionally, we posit that $\beta(s,t)$ adopts a basis expansion representation with the truncation constants K_y and K_x as follows:

$$\beta(s,t) = \sum_{l=1}^{K_y} \sum_{p=1}^{K_x} b_{lp} \psi_l(t) \vartheta_p(s), \quad \textbf{(3)}$$

where $\psi_l(t)$ and $\vartheta_p(s)$ denote the expansion functions and b_{lp} denotes the corresponding expansion coefficient. Let $\Delta_r = s_{r+1} - s_r$ represent the length of r-th interval in S. Subsequently, the integral component in the minimization problem (3) can be approximated using numerical integration, which can be expressed as follows:

$$\begin{aligned}\int_S X_i(s)\beta(t,s)ds &\approx \sum_{r=1}^{G-1}\Delta_r\beta(t,s_r)X_i(s_r)\\ &= \sum_{r=1}^{G-1}\Delta_r\sum_{l=1}^{K_y}\sum_{p=1}^{K_x} b_{lp}\psi_l(t)\vartheta_p(s_r)X_i(s_r),\\ &= \sum_{l=1}^{K_y}\sum_{p=1}^{K_x} b_{lp}\psi_l(t)\tilde{\vartheta}_{p,i}\end{aligned} \quad (4)$$

where $\tilde{\vartheta}_{p,i} = \sum_{r=1}^{G-1}\Delta_r\,\vartheta_p(s_r)X_i(s_r)$. Replacing the basis expansion approximation in (1) gives:

$$Y_i(t) = \sum_{k=1}^{K_0}\alpha_k\phi_k(t) + \sum_{l=1}^{K_y}\sum_{p=1}^{K_x} b_{lp}\psi_l(t)\tilde{\vartheta}_{p,i}. \quad (5)$$

For the penalty functionals , $J_1(\alpha)$ and $J_2(\beta)$, we employ quadratic penalties based on the second derivatives of the basis functions. Let $\boldsymbol{\alpha} = \left[\alpha_1,\dots,\alpha_{K_0}\right]^T$ represent the vector consisting of expansion coefficient for the constant function. Then, the penalty functional $J_1(\alpha)$ is approximated as $\tilde{J}_1(\alpha) = \int_I\left[\alpha^{(2)}(t)\right]^2 dt = \boldsymbol{\alpha^T P_\alpha \alpha}$, where $\alpha^{(2)}(t)$ denotes the second derivative of $\alpha(t)$, and $\boldsymbol{P_\alpha}$ denotes penalty matrice with dimension $K_0\times K_0$ with entries $\int_I \phi_k^{(2)}(t)\phi_l^{(2)}(t)dt$. Let $\boldsymbol{b} = \left(b_{lp}\right)_{lp}$ denote the $K_y\times K_x$ dimensional basis expansion coefficients. Then, the penalty functional $J_2(\beta)$ is approximated as follows:

$$\tilde{J}_2(\beta) = \int_I\int_S\left[\frac{\partial^2}{\partial t^2}\beta(t,s)\right]^2 ds\,dt + \int_I\int_S\left[\frac{\partial^2}{\partial s^2}\beta(t,s)\right]^2 ds\,dt = \boldsymbol{b}^T(\boldsymbol{\vartheta}\otimes\boldsymbol{P}_y + \boldsymbol{P}_x\otimes\boldsymbol{\psi})\,\boldsymbol{b}, \quad (6)$$

where $\boldsymbol{\psi} = \int_{\boldsymbol{I}}\boldsymbol{\psi}(\boldsymbol{t})\boldsymbol{\psi}(\boldsymbol{t})^T dt$ with $\boldsymbol{\psi}(\boldsymbol{t}) = \left[\boldsymbol{\psi_1}(\boldsymbol{t}),\dots,\boldsymbol{\psi}_{K_y}\right]^T$, $\boldsymbol{\vartheta} = \int_{\boldsymbol{S}}\boldsymbol{\vartheta}(\boldsymbol{s})\boldsymbol{\vartheta}(\boldsymbol{s})^T ds$ with $\boldsymbol{\vartheta}(\boldsymbol{s}) = \left[\boldsymbol{\vartheta_1}(\boldsymbol{s}),\dots,\boldsymbol{\vartheta}_{K_x}(\boldsymbol{s})\right]^T$. Here, $\boldsymbol{P}_y$ and $\boldsymbol{P}_x$ are the matrices of penalty terms whose elements are computed from the derivatives of the expansion coefficients.

Subsequently, leveraging the approximate penalty functionals and the expansion representation of the functionals, we can rewrite the minimization problem in (2) as follows:

$$\underset{\boldsymbol{\alpha},\boldsymbol{b}}{\operatorname{argmin}}\sum_{i=1}^n\sum_{j=1}^M\rho\left[Y_i(t_j) - \phi^T(t_j)\boldsymbol{\alpha} - \left(\tilde{\boldsymbol{\vartheta}}_i^T\otimes\boldsymbol{\psi}^T(t_j)\right)\boldsymbol{b}\right] + \frac{\lambda_1}{2}\tilde{J}_1(\alpha) + \frac{\lambda_2}{2}\tilde{J}_2(\beta). \quad (7)$$

To obtain robust estimates for $\boldsymbol{\alpha}$ and $\boldsymbol{b}$, we consider the τ-estimator proposed by [13]. Let $\boldsymbol{\theta} = [\boldsymbol{\alpha}^T,\boldsymbol{b}^T]^T$, $\boldsymbol{\Pi} = [\boldsymbol{\Pi_1},\dots,\boldsymbol{\Pi_n}]^T$ with $\boldsymbol{\Pi_i} = \left[\boldsymbol{\phi^T}(t)\tilde{\boldsymbol{\vartheta}}_i^T\otimes\boldsymbol{\psi^T}(t)\right]^T$, and $\boldsymbol{P}(\lambda_1,\lambda_2)$ is a block diagonal matrix whose elements are $\lambda_1\boldsymbol{P}_\alpha$ and $\lambda_2(\boldsymbol{\vartheta}\otimes\boldsymbol{P}_y + \boldsymbol{P}_x\boldsymbol{\psi})$. Subsequently, we consider the following optimization problem in a matrix form:

$$arg\min_{\boldsymbol{\theta}}\sum_{i=1}^n\rho[Y_i(t) - \boldsymbol{\Pi}_i\boldsymbol{\theta}] + \boldsymbol{P}(\lambda_1,\lambda_2)\boldsymbol{\theta} \quad (8)$$

The τ-estimator for $\boldsymbol{\theta}$ is defined as follows:

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\,\tau(\boldsymbol{\theta}), \quad (9)$$

where the τ-scale estimator $\tau(\boldsymbol{\theta})$ is given by

$$\tau^2(\boldsymbol{\theta}) = s^2(\boldsymbol{\theta})\frac{1}{n\,u_2}\sum_{i=1}^n\rho_2\left[\frac{Y_i(t)-\boldsymbol{\theta}^T\boldsymbol{\Pi}_i}{s(\boldsymbol{\theta})}\right], \quad (10)$$

with $s(\boldsymbol{\theta})$ is an M-estimator that solves $\frac{1}{n}\sum_{i=1}^n\rho_1\left[\frac{Y_i(t)-\boldsymbol{\theta}^T\boldsymbol{\Pi}_i}{s(\boldsymbol{\theta})}\right] = u_1$. Here, the loss functions ρ_1 and ρ_2 are symmetric, continuously differentiable, and bounded functions. The parameters u_1 and u_2, conversely, act as tuning parameters utilized to achieve consistency under normally distributed error terms. The selection of loss functions ρ_1 and ρ_2 holds significant practical and theoretical relevance. In this investigation, we adopt the optimal loss function proposed by [15]:

$$\rho(v) = \begin{cases} 1.38\left(\frac{v}{c}\right)^2, & \left|\frac{v}{c}\right| \le \frac{2}{3}\\ 0.55 - 2.69\left(\frac{v}{c}\right)^2 + 10.76\left(\frac{v}{c}\right)^4 - 11.66\left(\frac{v}{c}\right)^6 + 4.04\left(\frac{v}{c}\right)^8, & \frac{2}{3} < \left|\frac{v}{c}\right| \le 1\\ 1, & \left|\frac{v}{c}\right| > 1\end{cases} \quad (11)$$

Following the recommendation of [15], we opt for $c_1 = 1.214$ and $b_1 = 0.5$ for ρ_1 and $c_2 = 3.270$ and $b_2 = 0.128$ for ρ_2. With these parameter selections, the τ-estimator achieves a 50% breakdown and 95% efficiency under normally distributed error terms, as demonstrated by [15].

The τ-estimator $\widehat{\boldsymbol{\theta}}$ is computed via an iterative algorithm. We employ a random resampling-based fast estimation algorithm for this purpose. Initially, random subsamples are drawn from the entire dataset. For each subsample, the iteratively reweighted least squares algorithm is iterated multiple times to obtain potential estimates. This process continues until convergence, and the final estimator is selected from the potential estimates, providing the minimum scale estimate. Let $\widehat{\boldsymbol{\theta}} = \left[\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{b}}\right]^T$ represent the τ-estimate of $\boldsymbol{\theta}$. Subsequently, the robust estimates of the intercept function and bivariate coefficient function are obtained as follows:

$$\hat{\alpha}(t) = \widehat{\boldsymbol{\phi}}^T\widehat{(t)}\,\widehat{\boldsymbol{\alpha}}, \qquad \hat{\beta}(t,s) = [\boldsymbol{\vartheta}^T(s)\otimes\boldsymbol{\psi}^T(t)]\,\widehat{\boldsymbol{b}}. \quad (12)$$

III. MONTE CARLO SIMULATIONS

We implement a series of simulations to assess the estimation and predictive performance of the proposed method, referred to as "tau." This method's empirical performance is benchmarked against functional principal component regression (fpcr), functional partial least squares regression (fpls), and the penalized function-on-function linear regression model introduced by [7] (pffr). For the simulations, we adopt the data generation process outlined in [11].

In the data generation process, the predictor is created at fifty equally-spaced grid points within the unit interval. Similarly, the elements of response variable are simulated at sixty equally distance points within the same interval. The generation of the functional predictor follows:

$$X_i(s) = \sum_{r=1}^{50} \frac{1}{r^2} \{\zeta_{i1,r} \sqrt{2} \sin(r\,\pi s) \zeta_{i2,r} \sqrt{2} \cos(r\,\pi s)\},$$

where $\zeta_{i1,r}$ and $\zeta_{i2,r}$ i.i.d. random variables from the normal distribution with zero mean and unit variance. Following this, the functional response is generated using the specified methodology:

$$Y_i(t) = \alpha(t) + \int_0^1 X_i(s)\beta(s,t)ds + \epsilon_i(t)\,,$$

where $\alpha(t) = 2\exp(-(t-1)^2)$, $\beta(s,t) = 4\cos(2\,\pi t)\sin(\pi s)$, and ϵ_i is the random noise where each $\epsilon_i(t_j) \sim N(0,0.01)^2$. Replacing 5% and 10% of the data points with outliers, we utilize $\alpha^*(t) = 4\exp(-(t)^2)$, $\beta^*(s,t) = 6\sin(4\,\pi t)\sin(2\pi s)$ to generate atypical observations. Figure 1 presents a visual depiction of the generated data alongside the parameter functions utilized in the process.

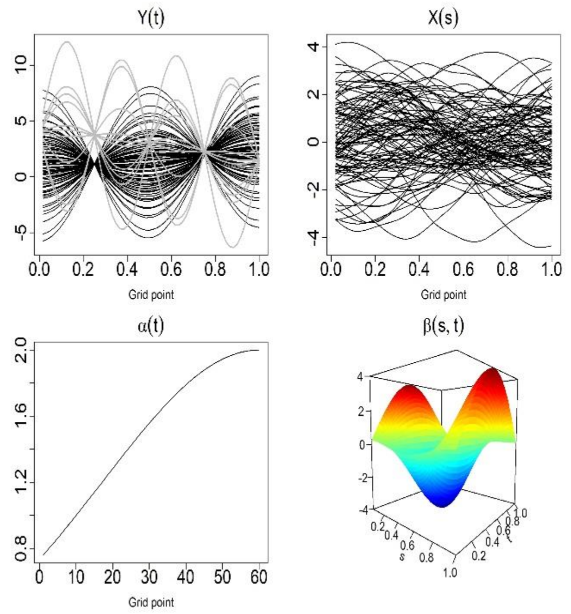

Figure 1. The graphical display showcases the response (left-top panel), predictor (right-top panel), constant function (left-bottom panel), and slope function (right-bottom panel).

In the experiments, a training sample of fixed size $n_{train} = 250$ is generated. Based on the training sample, we build the models and compute the root relative integrated squared percentage estimation errors (RISPEE) for the constant and slope functions for assessing the estimation performance of the methods:

$$RRISPEE(\hat{\alpha}) = 100 \times \sqrt{\frac{||\hat{\alpha}(t) - \alpha(t)||_2^2}{||\alpha(t)||_2^2}},$$

$$RRISPEE(\hat{\beta}) = 100 \times \sqrt{\frac{||\hat{\beta}(s,t) - \beta(s,t)||_2^2}{||\beta(s,t)||_2^2}},$$

where $||\cdot||_2$ denotes the L_2 norm. It's essential to mention that the methods fpcr and fpls presume that the response and predictor variables are centered, so that they have mean-zero. That is, $RRISPEE(\hat{\alpha})$ is computed only for pffr and tau methods. To assess the predictive performance of the methods, a test sample of size $n_{test} = 100$ is generated. The models fitted using the training sets are then applied to the test samples, and we compute the root mean squared percentage error (RMSPE) as follows:

$$RMSPE = 100 \times \sqrt{\frac{||\hat{Y}(t) - Y(t)||_2^2}{||Y(t)||_2^2}}$$

In the simulations, 100 Monte Carlo replications are conducted. For constructing the models using pffr, fpcr, fpls, and the proposed tau method, a fixed 15 basis expansion functions are employed.

The computed mean $RRISPEE(\hat{\alpha})$, $RRISPEE(\hat{\beta})$, and $RMSPE$, along with their standard errors, are presented in Table 1. When no outliers are present in the data, the proposed tau estimator demonstrates superior parameter estimation for the intercept function and achieves better prediction accuracy, as indicated by lower RMSPE values, compared to all other methods. This improvement may be due to the random data generation process, where a 0% contamination level might still produce small-magnitude outliers. The proposed method effectively mitigates the impact of these outliers, yielding enhanced results, whereas other non-robust methods are influenced by these outliers, resulting in biased outcomes. However, in this scenario, the tau method performs worst for $RRISPEE(\hat{\beta})$, with fpcr and fpls showing the best results. When outliers are introduced into the data, regardless of the contamination level, the proposed tau method consistently outperforms all competitors across all performance metrics. The non-robust methods exhibit significantly poorer estimation and predictive performance in the presence of outliers compared to their results with 0% contamination. In contrast, the proposed method effectively down-weights the influence of outliers, maintaining performance comparable to that achieved with no contamination.

Table 1. The mean $RRISPEE(\hat{\alpha})$, $RRISPEE(\hat{\beta})$, and $RMSPE$, along with their standard errors (given in brackets), are computed over 100 Monte Carlo replications.

%	Metric	pffr	fpcr	fpls	tau
0	$RRISPEE(\hat{\alpha})$	0.0958 (0.0940)	-	-	0.0253 (0.0052)
	$RRISPEE(\hat{\beta})$	2.3722 (0.0073)	1.7457 (0.6129)	0.8209 (0.1517)	3.5597 (1.0817)
	$RMSPE$	0.7463 (0.0279)	3.9089 (2.5550)	3.9166 (2.5646)	0.0729 (0.0035)
5	$RRISPEE(\hat{\alpha})$	6.8535 (1.1207)	-	-	0.0347 (0.0053)
	$RRISPEE(\hat{\beta})$	32.5246 (13.7395)	8.6258 (6.4341)	9.6036 (7.334)	4.1152 (1.4877)
	$RMSPE$	7.4112 (1.0931)	7.9222 (1.5537)	8.1117 (1.6796)	0.0850 (0.0059)
10	$RRISPEE(\hat{\alpha})$	12.3480 (0.6299)	-	-	0.0403 (0.0089)
	$RRISPEE(\hat{\beta})$	32.6744 (11.6223)	15.9920 (6.7921)	12.3848 (2.1659)	6.0510 (2.7341)
	$RMSPE$	12.2392 (0.8134)	12.4304 (1.0484)	12.2301 (1.0691)	0.0991 (0.0058)

Moreover, we compare the performance of the proposed method with existing non-robust methods, namely, pffr, fpcr, and fpls, in terms of their computing times. A single Monte Carlo simulation is performed with a sample size of 250, and the elapsed computing time for all the methods is recorded. The computations are executed on a desktop PC with an Intel® Core™ i5-9500 CPU at 3.00 GHz and 8 GB RAM. The computing times (in seconds) are recorded as 6.83, 0.25, 0.95, and 72.63 for pffr, fpcr, fpls, and the proposed tau method, respectively. From the results, it is evident that the classical non-robust methods require considerably less computing time than the proposed method. This result is due to the proposed method utilizing an iterative approach for estimating both model parameters and smoothing parameters (grid-search algorithm).

IV. EMPIRICAL DATA ANALYSIS

We employ the Oman weather dataset from the National Center for Statistics & Information (https://data.gov.om). This dataset comprises monthly maximum humidity (in percentage) and evaporation (in millimeters) measurements from 49 weather stations across Oman, spanning from January 2018 to December 2018. Each observation is treated as a function of the months, resulting in a total of n = 49 functional observations $\{Y(t), X(s): 1 \leq t, s \leq 24\}$. Figure 2 presents the graphical display of the functional observations.

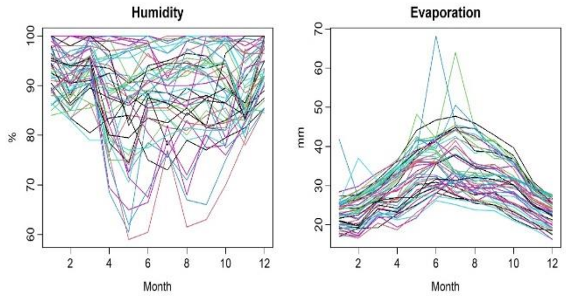

Figure 2. Graphical display of the maximum humidity (left panel) and evaporation (right panel) variables fort he Oman weather data.

To assess the predictive performance of the methods, we repeat the following procedure 100 times: 1) Randomly split the dataset into training and test samples, with sizes 33 and 16, respectively. 2) Construct a model using the training sample curves, employing 8 basis functions determined by the generalized cross-validation procedure. 3) Use this model to predict 13 curves in the test sample. 4) Calculate the RMSPE for each replication to compare the predictive accuracy of the methods.

Figure 2 reveals that the Oman weather data contains clear atypical observations in the response (humidity). Hence, it's expected that the proposed robust method would deliver superior prediction results compared to its non-robust counterparts, namely pffr, fpcr, and fpls. The mean RMSPE values computed from the methods and their standard errors given inbrackets are as follows: 5.9523 (0.9518) for pffr, 5.8333 (0.9255) for fpcr, 5.2234 (0.8309) for fpls, and 3.1239 (0.4478) for the proposed tau method. These findings suggest that our method achieves improved predictive performance, as indicated by lower RMSPE values compared with its competitors.

V. CONCLUSION

The FoFR model has emerged as a pivotal tool for investigating the functional relationship between a functional response and a set of functional predictor variables. Numerous methods have been put forth to accurately estimate the parameters of this model. However, many of these methods suffer from a lack of robustness and can be substantially influenced by the existence of outliers. Consequently, traditional approaches may produce biased estimates for the regression parameters, leading to subpar predictive performance.

This study introduces a novel penalized robust estimation method, named "tau," tailored to acquire outlier-resistant estimates for the slope function of the FoFR model. The proposed method employs B-spline expansion to project functional object into a finite-dimensional space and the penalty functionals obtained from their second derivatives are applied to the expansion coefficients to control the smoothness of the estimates. To assess the estimation and predictive performance, a series of Monte Carlo experiments and empirical data analyses are conducted, comparing the results favorably with existing methods. The findings indicate that the proposed method yields comparable estimation and predictive performance to existing non-robust methods in outlier-free data scenarios. However, notably, our method demonstrates improved estimation and predictive accuracy when data are contaminated by outliers, surpassing the performance of existing methods in such scenarios.

The estimation approach proposed in this study can be extended in several research directions. For example, the current functional regression model includes only one functional predictor. The proposed method can be easily extended to models that include multiple functional predictors. Additionally, the considered model includes only a functional predictor. However, fields such as health and medicine often require both functional and scalar predictors. Therefore, the proposed method can be extended to robustly estimate model parameters when both functional and scalar predictors are included. Moreover, our Monte Carlo experiments indicate that the proposed method requires significantly more computing time than existing methods. To address this, several algorithms, such as parallel computing, can be applied to reduce the computational burden of the proposed method.

REFERENCES

[1] Ramsay, J.O., & Silverman, B.W. (2006). *Functional data analysis*, 2nd edn. Springer, New York.

[2] Horvath, L., & Kokoszka, P. (2012). *Inference for functional data with applications*, Springer, New York.

[3] Kokoszka, P., & Reimherr, M. (2015). *Introduction to functional data analysis*, CRC Press, London.

[4] Yamanishi, Y., & Tanaka, Y. (2003). Geographically weighted functional multiple regression analysis: A numerical investigation, *Journal of the Japanese Society of Computational Statistics*, 15, 307–317.

[5] Yao, F., Müller, H.-G., & Wang, J.-L. (2005). Functional linear regression analysis for longitudinal data, *The Annals of Statistics*, 33, 2873–2903.

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

[7] Ivanescu, A. E., Staicu, A.-M., Scheipl, F., & Greven, S. (2015). Penalized function-on-function regression, *Computational Statistics*, 30, 539–568.

[8] Beyaztas, U., & Shang, H.L. (2020). On function-on-function regression: Partial least squares approach, *Environmental and Ecological Statistics*, 27, 95–114.

[9] Hullait, H., Leslie, D.S., Pavlidis, N. G., & King, S. (2021). Robust function-on-function regression, *Technometrics*, 63, 396–409.

[10] Beyaztas, U., & Shang, H.L. (2022). A robust partial least squares approach for function-on-function regression, *Brazilian Journal of Probability and Statistics*, 36, 199-219.

[11] Cai, X., Xue, L., & Cao, J. (2021). Robust penalized M-estimation for function-on-function linear regression, *Stat* 10, e390.

[12] Beyaztas, U., Shang, H.L., & Mandal, A. (2023). Robust function-on-function interaction regression, *Statistical Modelling*, in-press.

[13] Yohai, V.J., & Zamar, R.H. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale, *Journal of the American Statistical Association: Theory and Methods*, 83, 406–413.

[14] Beyaztas, U., Shang, H.L., & Saricam, S. (2024). Penalized function-on-function linear quantile regression, *Computational Statistics*, in-press.

[15] Salibian-Barrera, M., Willems, G., & Zamar, R. (2008). The fast-τ estimator for regression, *Journal of Computational and Graphical Statistics*, 17, 659-682.