Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

Type of the Paper (Review)

A Literature Review of Semi-Functional Partial Linear Regression models

Mohammad Fayaz [1,*]

[1] PhD in Biostatistics, Shahid Beheshti University of Medical Sciences, Tehran, Iran; E-mail: Mohammad.Fayaz.89@gmail.com
* Mohammad.Fayaz.89@gmail.com; ORCID: 0000-0002-5643-9763.

Background: In the functional data analysis (FDA), the hybrid or mixed data are scalar and functional datasets. The semi-functional partial linear regression model (SFPLR) is one of the first semiparametric models for the scalar response with hybrid covariates. Various extensions of this model are explored and summarized. **Methods**: Two first research articles, including "semi-functional partial linear regression model", and "Partial functional linear regression" have more than 300 citations in the Google Scholar. Finally, only 106 articles remained according to the inclusion and exclusion criteria such as 1) including the published articles in the ISI journals and excluding 2) non-English and 3) preprints, slides and conference papers. We use PRISMA standard for systematic review. **Results**: The articles are categorized into the following main topics: estimation procedures, confidence regions, time series, and panel data, Bayesian, spatial, robust, testing, quantile regression, varying Coefficient Models, Variable Selection, Single-index model, Measurement error, Multiple Functions, Missing values, Rank Method and Others. There are different applications and datasets such as Tecator dataset, air quality, electricity consumption, and Neuroimaging ,among others. **Conclusions**: SFPLR is one of the most famous regression modeling methods for hybrid data that has a lot of extensions among other models.

Keywords: Functional Data Analysis (FDA); Hybrid Data; Semi-Functional Partial Linear Regression Model (SFPLR); Partial Functional Linear Regression ;Literature Review;

1. Introduction

The regression models in statistics estimate the relations between dependent variables (continuous, categorical, counts, or time-to-event) and independent variables. They are linear regression models, generalized linear models (GLM), generalized additive models (GAM), etc. for modeling different relations [1].

Recent advances in theoretical and application yield the new term "functional data analysis (FDA)" that it provides methodologies to summarize and model the high-dimension variables with underlying functional structure. In many cases when both functional and non-functional (such as scalar) are in interest, the new term is mixed ,or hybrid data [2, 3].

One of the first regression models that formulate the relationship between scalar and hybrid data is semi-functional partial linear regression [4]. Since then, there have been many extensions and applications of this model that is published, and this systematic review collects and summaries the most important of them. It first introduces the model in the mathematical formula, in the next part, it categorizes the models and applications. Finally, it has some conclusions about computing and usefulness of the model.

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

2. Systematic review

The research articles titled "Semi-Functional Partial Linear Regression" by German Aneiros-Perez and Philippe Vieu [4] and "Partial functional linear regression" by Hyejin Shin [5] published in 2006 and 2009 in the Statistics and Probability Letters and Journal of Statistical Planning and Inference, respectively.

The main idea of these models is to consider both the functional and non-functional, or mixed and hybrid, covariates to predict the real-valued scalar response. It used the nonparametric method for functional covariate with the weights of the functional version of the Nadaraya–Watson and the parametric method for the non-functional covariate with the linear relation. In the next few years, the authors and some researchers extended this model ,and recently ,many other extensions published.

In this study, we search these extensions with Google Scholar from 2006 to 2021 ,and there are at least 300 results. In the next step, we select among them based on the following inclusion criteria: published in the ISI-indexed journals and the exclusion criteria are 1) written in a non-English language, 2) published as preprints, thesis, conference proceedings, and slides. Finally, 106 research articles remained. (Figure 1) [6]

3. Models and their extensions

3.1. Semi-Functional Partial Linear Regression Model

Suppose that $\{(Y_i, X_{i1}, \dots, X_{ip}, T_i)\}_{i=1}^{n}$ is a sample of n independent and identically distributed vectors valued in $\mathbb{R}^{p+1} \times \ell \; (\ell \subset \mathcal{H})$. The ℓ is some given compact subset of $\mathcal{H}$ such that $\ell \subset \bigcup_{k=1}^{\tau_n} \mathcal{B}(z_k, l_n)$ that $\tau_n l_n^{\gamma} = C \; (\gamma \; and \; C \; denote \; real \; positive \; constants)$, $\tau_n \to \infty \; and \; l_n \to 0 \; as \; n \to \infty$ and $\mathcal{B}(t,h) = \{t' \in \mathcal{H}; d(t',t) < h\}$. The semi-functional partial linear regression model (SFPLR model) is [4] as

$$Y = r\left(X_1, \dots, X_p, T\right) + \varepsilon = \sum_{j=1}^{p} X_j \beta_j + m(T) + \varepsilon \qquad (1)$$

where

$$\begin{cases} Y & real \; valued \; response \; variable \\ X_j (j = 1, \dots, p) & real \; explanatory \; variable \\ T & functional \; explanatory \\ \varepsilon & random \; error, E\left(\varepsilon \middle| X_1, \dots, X_p, T\right) = 0 \\ \boldsymbol{\beta} = \left(\beta_1, \dots, \beta_p\right)^T & unknown \; real \; parameters \\ m & unkonwn \; smooth \; real \; function \end{cases}$$

The functional variables are modeled in the semi-metric space. The model's parameters are estimated with the following formulas with h as a smoothing parameter and kernel function K is a function from $[0,\infty)$ into $[0,\infty)$ has support $[0,1]$, is Lipschitz continuous on $[0,\infty]$ and $\exists \theta$ such that $\forall u \in [0,1], -K'(u) > \theta > 0$:

$$\hat{\beta}_h = \left(\tilde{X}_h^T \tilde{X}_h\right)^{-1} \tilde{X}_h^T \tilde{X}_h \qquad (2)$$

$$\widehat{m}_h(t) = \sum_{i=1}^{n} w_{n,h}(t, T_i)(Y_i - X_i^T \hat{\beta}_h) \qquad (3)$$

$$w_{n,h}(t, T_i) = \frac{K(d(t, T_i)/h)}{\sum_{j=1}^{n} K(d(t, T_j)/h)} \qquad (4)$$

(functional version of the Nadaraya-Watson-type weights)

3.2. Partial functional linear regression

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

In a more general model, "suppose Y , the response is a real-valued r.v. on a probability space $(\Omega, \mathcal{B}, P)$. And $\boldsymbol{z}$,non-functional covariates, is a p-dimensional vector of r.v. with zero means and finite second moments. And $\{X(t): t \in \mathrm{T}\}$ be a zero mean, the second-order stochastic process defined on $(\Omega, \mathcal{B}, P)$ with sample paths in $L^2(T)$ the set of all square-integrable functions on T " [5] :

$$Y = \boldsymbol{\beta}^T \boldsymbol{z} + \int_0^1 \gamma(t) X(t) dt + \varepsilon, \qquad (5)$$

$$\begin{cases} \gamma \ \ is\ a\ square\ integrable\ function\ on\ [0,1] \\ \varepsilon \ \ is\ a\ random\ error\ with\ zero\ mean\ and\ variance\ \sigma^2\ and\ independent\ of\ z\ and\ X \end{cases}$$

It is a general model, such that:

$$\begin{cases} if\ \gamma = 0 \rightarrow classical\ linear\ regression\ model \\ if\ \boldsymbol{\beta} = \boldsymbol{0} \rightarrow functioanl\ linear\ regression\ model \\ if\ m(X) = \int_0^1 \gamma(t) X(t) dt \rightarrow semi-functioanl\ partial\ linear\ regresion\ model \\ (Y = \boldsymbol{\beta^T z} + m(X) + \varepsilon) \end{cases}$$

3.3. Other extensions

There are many other extensions for models (1) and (5) ([4] and [5]), we categorize them:

- **Estimation procedures:** there are different methods and studies for estimation procedures of SFPLR model such as: fully automatic estimation procedure with the data-driven method and cross-validation for bandwidth selection of the smoothing parameter of nonparametric component [7], the asymptotic normality of linear part is studied [8], in a situations when the observation number of each subject is completely flexible and studying the convergence rate of the nonparametric part [9], the spline estimator of nonparametric part with studying their convergence rate [10], the two estimation procedure: 1-functional principal components regression (FPCR) and 2- functional ridge regression (FRR) based on the Tikhonov regularization [11-13], the new estimators for the parametric component called semiparametric least squares estimator (SLSE) [14], the nonparametric component approximated by a B-spline function [15], polynomial spline [16] and the slope function is estimated with the functional principal component basis [15, 16], the *k*-nearest-neighbours (*k*NN) estimates with the local adaptive property that is better in the practice than kernel methods and some computations properties of this estimator [17-19], the Functional Semiparametric Additive Model via COmponent Selection and Smoothing Operator (FSAM-COSSO) in sparse setting [20], the sufficient dimension reduction methods such as sliced inverse regression (SIR) and sliced average variance estimation (SAVE) [21], the estimations are from the reproducing kernel Hilbert spaces (RKHS) [22], the frequentist and optimal model averaging [23], the latent group structure with K-means clustering [24], the joint asymptotic framework called joint Bahadur representation [25], the empirical likelihood estimation for non-functional high-dimension covariates [26], Sparse and penalized least-squares estimators [27] and the software for doing this analysis is available [28] .
- **Confidence Regions:** Some papers have some sections for calculating confidence regions ,and we do not repeat them. We emphasize the following articles: the empirical likelihood ratio with plug-in approach and its bias-corrected version [29] and the confidence bands for Partial functional linear regression [30].

- **Time Series and Panel Data**: The extensions that are related to the time series and forecasting are: the Semi-functional partial linear time series modeling for prediction [31], with autoregressive errors [32, 33], with time-varying parameters for latent parameter regimes [34], regularized forecasting via smooth-rough partitioning of the regression coefficients [35].
- **Bayesian:** The Bayesian estimation methods are present in some papers ,but we only mention these two papers in this part: the Bayesian bandwidth estimation and semi-metric selection for a functional partial linear model with unknown error density [36, 37].
- **Spatial:** The spatial variability is considered in many research articles such as The partial functional linear spatial regression autoregressive model with spatial dependence responses [38], with two-stage estimator based on quasi-maximum likelihood estimation (QMLE) method and local linear regression method [39], studying the asymptotic normality of the parametric component, and probability convergence with the rate of the nonparametric component [40], B-spline approximation for slope function and residual-based approach for pointwise confidence-intervals [41], the robust spatial autoregressive model with t-distribution error terms with an expectation-maximization algorithm [42].
- **Robust:** Existing outliers in the data or violations from distributional assumptions yield to the robust methods such as the sieve M-estimator for semi-functional linear model [43], with polynomial splines to approximate the slope parameter and resistance to heavy-tailed errors or outliers in the response [44], different estimators such as M-estimators with bi-square function, GM-estimator with Huber function, LMS-estimator and LTS- estimators [45], estimation based on exponential squared loss and FPCA [46], estimation based on the class of scale mixtures of normal (SMN) distributions for measurement errors and Bayesian framework with MCMC algorithm [47], Robust MM-estimators with B-Spline approximation [48], with modal regression [49] and a modified Huber's function with tail function with a data-driven procedure for selecting the tuning parameters [50].
- **Testing:** Different hypothesis and testing statistics are developed ,such as: testing the linear component [51, 52] with B-spline [53], functional covariates [54], densely and sparsely observed single and multiple functional covariates with four tests such as Wald, Score, likelihood ratio and F [55], Goodness-of-fit tests with wild bootstrap resampling, false discovery rate and independence test with generalized distance covariance or new metric, functional martingale difference divergence (FMDD), [56-58], series correlation test [59].
- **Quantile regression**: Some extensions consider quantile regression property such as: proposed functional partially linear quantile regression model (FPLQRM) that has the linear variables which may be categorical [60], estimating the slope function between a dependent variable and both vector and functional random variable with FPCA [61], and piecewise polynomial [62] and kNN quantile method [63], functional composite quantile regression (CQR) with simple partial quantile regression (SIMPQR) algorithm and partial quantile regression (PQR) basis [64], composite quantile estimation with strictly stationary process errors [65] and with polynomial splines [66], Hill estimator for extreme quantile estimation with heavy-tailed distributions [67], developed quantile rank score test for a parametric component of the model [68], varying-coefficient partially functional linear quan-

tile regression model [69] with quantile estimation [70]. And responses with missing-at-random (MAR) [71].

- **Varying Coefficient Models:** There are some papers with extensions of varying coefficient models ,but we don't repeat them ,and we only select the following extensions: the partially varying coefficient models stratified by a functional covariate [72], varying coefficient partially functional linear regression model (VCPFLM) [73], partially functional linear varying coefficient model (PFLVCM) with a hypothesis and bootstrap [74], and the robust estimation based on the rank-based estimation [75].
- **Variable Selection:** These papers are related to the variable selections methodologies: the variable selection with nonconcave-penalized least square in a high-dimensional partial linear regression model and with penalized composite quantile regression method [76, 77], simultaneously consider multiple functional and scalar predictors and identify the important features [78], estimation and variable selection based on penalized regression estimators [79].
- **Single-index model :** The SFPLR with single-index models are: functional partial linear single-index model (FPLSIM) [80], with B-spline approximations [81] , with profile least-squares estimation (PLSE) for slope [82] and Partially Linear Generalized Single Index Models for Functional Data (PLGSIMF) [83]. The systemic review of semiparametric regression (single functional index regression (SFIR)) model is available [84].
- **Measurement error:** There are some extensions that variables have measurement error, such as the model with error-in-response and FPCA estimation [85], non-functional covariate with error, its test and with corrected profile least-squares based estimation [86, 87], both scalar and functional covariate measured with additive error [88].
- **Multiple Functions:** In addition to the variable selection papers, there are some papers for multiple functions extensions: the multi-functional partially linear model (MFPLR) with multiple functional covariates with proposed for variable selection (PVS) [89], the FPCA estimation for each component, linear hypothesis and confidence bands [90].
- **Missing values:** In some cases with missing values, there are some papers such as the empirical likelihood estimation and least squares estimator (LSE) for models with missing at random (MAR) responses with their confidence intervals [91-93].
- **Rank Method:** the nonparametric estimation methods are extended ,such as the rank estimation for partial functional linear regression models with FPCA [94], hypothesis test for the parametric component based on the rank score function [95].
- **Others:** Some important papers are: the functional partial linear model that combines the parametric and nonparametric approaches with functional regression [96]. The error variance estimation and confidence region construction are presented [97]. Naïve and wild bootstrap procedures are for kernel-based estimators [98]. The partial functional linear model with skew-normal errors and homogeneity test is proposed [99]. The generalized partial functional linear additive models (GPFLAM) are approximated by polynomial splines ,and FPCA and asymptotic normality of the estimator is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

obtained [100, 101]. The time-to-event response in the presence of random right censoring is modeled with a synthetic response by transforming it with three different types of transformation (L), (KSV), and a more general class of transformation (FG). Then it models with functional linear regression model [102]. The two-sample functional linear model with functional responses is a general model for the partial functional linear model ,and it has been studied recently [103]. An example of partial functional linear regression in reinforced risk prediction with electronic health records is simulated [104].

4. Applications

We present some different applications and datasets with these extension models. They are in the following categories:

- **Tecator Dataset:** The Tecator Infratec Food and Feed Analyzer use near infrared transmission (NIT) principle. This dataset has the following wavelength from 850 to 1050 nm. The variables are fat, protein and moisture contents. [4, 17, 20, 27, 29, 30, 36, 37, 43, 45, 46, 48, 51, 56, 58, 63, 66, 69, 70, 72, 74, 75, 80-83, 85-87, 89, 91, 92, 97, 99-101].
- **Air Quality:** Different air quality indices in different cities. [23, 26, 29, 31, 40, 47, 78]
- **Electricity:** The electricity consumptions and demand. [33, 48, 59, 65, 98, 105-109]
- **Berkeley growth study data:** The dataset for functional data analysis with male and female growth datasets. [11, 15, 52, 53, 61, 68, 94, 95]
- **Environment and epidemiology (Temperature, sunspot, corn yield, barley plants, AIDS):** Different datasets with the main topics in environment and epidemiology [22, 24, 35, 42, 54, 56, 57, 103]
- **Neuroimaging:** Different neuroimaging dataset such Attention deficit hyperactivity disorder (ADHD) and Diffusion Tensor Imaging (DTI) [20, 55, 62, 64, 67, 90, 102]
- **Economics (Auction, stock market, real estate):** Different datasets with economics subjects. [13, 34, 35, 55, 79]
- **Biscuit dough data:** [49, 77]
- **Yogurt:** [19]

5. Discussion

The hybrid data [2] is widespread, in the FDA problems ,and three main scalar-on-functions regression models are introduced [4, 5, 96] ,and their extensions are the main focus of this article. But there is not limited to only these methods ,and we discuss some other frameworks and models with different structures.

5.1 Other Models

Among the various models, we select the following methods:

Dimension reduction methods: the principal component analysis for hybrid or mixed data was introduced by [2] with some examples in Canadian weather stations

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

that consider average vectors along the temperature curves. The other models are three dimension hybrid PCA with an application in electroencephalography [110] and hybrid functional and vector data (HFV-PCA) [111].

Regression Models: The scalar-on-function regression with FACE, Penalized functional regression and scalar-on-additive models are counted as functional regressions[1, 112].Other examples are the general additive regression model and variable selection with scalar response and mixed (scalar, functional, directional,…) covariates [113], function-on-both functional and scalar covariates with signal compression [114], covariate-adjusted generalized functional linear model [115].

5.2 Final Remarks

The extensions of semi-functional partial linear regression models published in more than 40 ISI-indexed journals, mainly in statistical journals. The first four journals from several articles are "Communications in Statistics-Theory and Methods", "Journal of Statistical Planning and Inference", "Journal of Multivariate Analysis", and "Metrika" with 9,8,7 and 6 published articles, respectively. The number of published articles from 2006 to 2021 is increasing ,which is about 26 for 2020 and 20 for 2021 at the end of September.

5. Conclusions

The semi-functional partial linear regression model and their extensions for different methods and situations are for examples time-series, quintile regression, varying coefficient model, statistical testing, robust estimation, Bayesian estimation, multi-functional covariates, variable selection, confidence bands, and prediction intervals, missing data, errors in variable and others and different tests for both parametric and non-parametric components of the model.

And there are also different applications such as in spectroscopy, in air-pollution and related topics, child growth study, neuroimaging, electricity demand and price, and others. Most of them published in statistical journals but some of them published in neuroscience, energy, and mathematics journal. But they are other methods for mixed and hybrid data that we discussed.

With exploiting of big data and availability of different kinds of data types such as scalar, functions, time-series, spatial points, missing data, directional, survival analysis, images, etc. in various fields of research such as genetics, pharmaceuticals, neuroimaging, movement, and mobile health (mhealth) monitoring, the need for models that can use the most information of them are vital. Among them, Semi-Functional Partial Linear Regression models are developed and used widely.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

PRISMA 2020 flow diagram for updated systematic reviews which included searches of databases and registers only

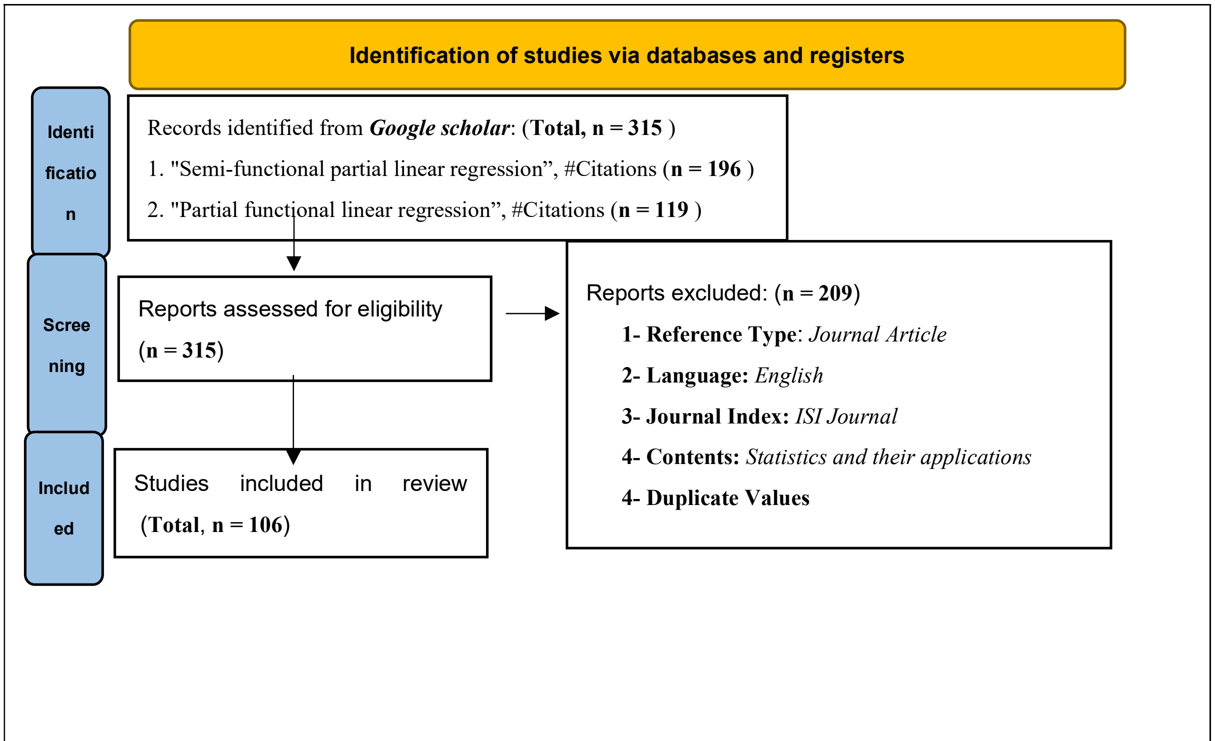

Figure 1. The flow diagram of systematic review based on PRISMA 2020 standard. (Since 10/20/2021.)

Supplementary Materials: Some visualization is available in Shiny web applications: https://mohammadfayaz.shinyapps.io/IWFOS_2021/ . The research article is checked for Plagiarism with iThenticate (www.ithenticate.com) and it is checked for correctness, clarity and writing tips with Grammarly Premium (www.grammarly.com). This paper is prepared for a special issue functional data analysis (FDA) of the Stats.

Funding: "This research received no external funding".

Acknowledgments: The 5th International Workshop on Functional and Operatorial Statistics IWFOS 2021 (https://iwfos2021.sci.muni.cz/) that some part of this research is presented.

Conflicts of Interest: "The authors declare no conflict of interest."

References

1. Wood, S.N., *Generalized additive models: an introduction with R*. 2017: CRC Press.
2. Ramsay, J.S., Bernard;, *Functional Data Analysis*. 2 ed. Springer Series in Statistics. 2005: Springer-Verlag New York.
3. Ferraty, F. and P. Vieu, *Nonparametric functional data analysis: theory and practice*. 2006: Springer Science & Business Media.
4. Aneiros-Pérez, G. and P. Vieu, *Semi-functional partial linear regression.* Statistics & Probability Letters, 2006. **76**(11): p. 1102-1110.
5. Shin, H., *Partial functional linear regression.* Journal of Statistical Planning and Inference, 2009. **139**(10): p. 3405-3418.
6. Page, M.J., et al., *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.* BMJ, 2021. **372**.
7. Aneiros-Pérez, G. and P. Vieu, *Automatic estimation procedure in partial linear model with functional data.* Statistical Papers, 2011. **52**(4): p. 751-771.
8. Chokri, K. and D. Louani, *Asymptotic results for the linear parameter estimate in partially linear additive regression model.* Comptes Rendus Mathematique, 2011. **349**(19-20): p. 1105-1109.

9. Zhang, T. and Q. Wang, *Semiparametric partially linear regression models for functional data.* Journal of Statistical Planning and Inference, 2012. **142**(9): p. 2518-2529.
10. Zhou, J. and M. Chen, *Spline estimators for semi-functional linear model.* Statistics & Probability Letters, 2012. **82**(3): p. 505-513.
11. Shin, H. and M.H. Lee, *On prediction rate in partial functional linear regression.* Journal of Multivariate Analysis, 2012. **103**(1): p. 93-106.
12. Zhang, F. and H. Lian, *Partially functional linear regression with quadratic regularization.* Inverse Problems, 2019. **35**(10): p. 105002.
13. Qingguo, T. and B. Minjie, *Estimation for functional linear semiparametric model.* Statistical Papers, 2020: p. 1-25.
14. Zhang, T., *Asymptotic properties in semiparametric partially linear regression models for functional data.* Acta Mathematicae Applicatae Sinica, English Series, 2013. **29**(3): p. 631-644.
15. Qingguo, T., *Estimation for semi-functional linear regression.* Statistics, 2015. **49**(6): p. 1262-1278.
16. Zhou, J., Z. Chen, and Q. Peng, *Polynomial spline estimation for partial functional linear regression models.* Computational Statistics, 2016. **31**(3): p. 1107-1129.
17. Ling, N., G. Aneiros, and P. Vieu, *k NN estimation in functional partial linear modeling.* Statistical Papers, 2020. **61**(1): p. 423-444.
18. Novo, S., G. Aneiros, and P. Vieu, *A kNN procedure in semiparametric functional data analysis.* Statistics & Probability Letters, 2021. **171**: p. 109028.
19. Almanjahie, I.M., et al., *Computational aspects of the k NN local linear smoothing for some conditional models in high dimensional statistics.* Communications in Statistics-Simulation and Computation, 2021: p. 1-32.
20. Sang, P., R.A. Lockhart, and J. Cao, *Sparse estimation for functional semiparametric additive models.* Journal of Multivariate Analysis, 2018. **168**: p. 105-118.
21. Yang, G., H. Lin, and H. Lian, *On double-index dimension reduction for partially functional data.* Journal of Nonparametric Statistics, 2019. **31**(3): p. 761-768.
22. Cui, X., H. Lin, and H. Lian, *Partially functional linear regression in reproducing kernel Hilbert spaces.* Computational Statistics & Data Analysis, 2020. **150**: p. 106978.
23. Jiang, R., L. Wang, and Y. Bai, *Optimal model averaging estimator for semi-functional partially linear models.* Metrika, 2021. **84**(2): p. 167-194.
24. Wang, W., Y. Sun, and H.J. Wang, *Latent group detection in functional partially linear regression models.* Biometrics, 2021.
25. Cheng, G. and Z. Shang, *Joint asymptotics for semi-nonparametric regression models with partially linear structure.* Annals of Statistics, 2015. **43**(3): p. 1351-1390.
26. Jiang, Z., Z. Huang, and G. Fan, *Empirical likelihood for high-dimensional partially functional linear model.* Random Matrices: Theory and Applications, 2020. **9**(04): p. 2050017.
27. Novo, S., G. Aneiros, and P. Vieu, *Sparse semiparametric regression when predictors are mixture of functional and high-dimensional variables.* TEST, 2021. **30**(2): p. 481-504.
28. Febrero-Bande, M. and M.O. de la Fuente, *Statistical Computing in Functional Data Analysis: The R Package fda. usc.* Journal of Statistical Software, 2012. **51**(i04).
29. Lian, H., *Empirical likelihood confidence intervals for nonparametric functional data analysis.* Journal of Statistical Planning and Inference, 2012. **142**(7): p. 1669-1677.
30. Imaizumi, M. and K. Kato, *A simple method to construct confidence bands in functional linear regression.* Statistica Sinica, 2019. **29**(4): p. 2055-2081.
31. Aneiros-Perez, G. and P. Vieu, *Nonparametric time series prediction: A semi-functional partial linear modeling.* Journal of Multivariate Analysis, 2008. **99**(5): p. 834-857.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

32. Dabo-Niang, S. and S. Guillas, *Functional semiparametric partially linear model with autoregressive errors.* Journal of multivariate analysis, 2010. **101**(2): p. 307-315.
33. Xiao, P. and G. Wang, *Partial functional linear regression with autoregressive errors.* Communications in Statistics-Theory and Methods, 2020: p. 1-22.
34. Liebl, D. and F. Walders, *Parameter regimes in partial functional panel regression.* Econometrics and Statistics, 2019. **11**: p. 105-115.
35. Maeng, H. and P. Fryzlewicz, *Regularised forecasting via smooth-rough partitioning of the regression coefficients.* Electronic Journal of Statistics, 2019. **13**(1): p. 2093-2120.
36. Shang, H.L., *Bayesian bandwidth estimation for a semi-functional partial linear regression model with unknown error density.* Computational Statistics, 2014. **29**(3): p. 829-848.
37. Shang, H.L., *Bayesian bandwidth estimation and semi-metric selection for a functional partial linear model with unknown error density.* Journal of Applied Statistics, 2021. **48**(4): p. 583-604.
38. Hu, Y., et al., *Estimation in Partial Functional Linear Spatial Autoregressive Model.* Mathematics, 2020. **8**(10): p. 1680.
39. Li, Y. and C. Ying, *Semi-functional partial linear spatial autoregressive model.* Communications in Statistics-Theory and Methods, 2020: p. 1-14.
40. Benallou, M., et al., *Asymptotic results of semi-functional partial linear regression estimate under functional spatial dependency.* Communications in Statistics-Theory and Methods, 2021: p. 1-21.
41. Liu, G. and Y. Bai, *Statistical inference in functional semiparametric spatial autoregressive model.* AIMS Mathematics, 2021. **6**(10): p. 10890-10906.
42. Huang, T., et al., *A robust spatial autoregressive scalar-on-function regression with t-distribution.* Advances in Data Analysis and Classification, 2021. **15**(1): p. 57-81.
43. Huang, L., et al., *Sieve M-estimator for a semi-functional linear model.* Science China Mathematics, 2015. **58**(11): p. 2421-2434.
44. Zhou, J., J. Du, and Z. Sun, *M-Estimation for partially functional linear regression model based on splines.* Communications in Statistics-Theory and Methods, 2016. **45**(21): p. 6436-6446.
45. Boente, G. and A. Vahnovan, *Robust estimators in semi-functional partial linear regression models.* Journal of Multivariate Analysis, 2017. **154**: p. 59-84.
46. Yu, P., Z. Zhu, and Z. Zhang, *Robust exponential squared loss-based estimation in semi-functional linear regression models.* Computational Statistics, 2019. **34**(2): p. 503-525.
47. Shan, G., Y. Hou, and B. Liu, *Bayesian robust estimation of partially functional linear regression models using heavy-tailed distributions.* Computational Statistics, 2020: p. 1-16.
48. Boente, G., M. Salibian-Barrera, and P. Vena, *Robust estimation for semi-functional linear regression models.* Computational Statistics & Data Analysis, 2020. **152**: p. 107041.
49. Yu, P., et al., *Robust Estimation for Partial Functional Linear Regression Model Based on Modal Regression.* Journal of Systems Science and Complexity, 2020. **33**: p. 527-544.
50. Cai, X., L. Xue, and F. Lu, *Robust estimation with a modified Huber's loss for partial functional linear models based on splines.* Journal of the Korean Statistical Society, 2020: p. 1-24.
51. Aneiros-Pérez, G. and P. Vieu, *Testing linearity in semi-parametric functional data analysis.* Computational statistics, 2013. **28**(2): p. 413-434.
52. Yu, P., Z. Zhang, and J. Du, *A test of linearity in partial functional linear regression.* Metrika, 2016. **79**(8): p. 953-969.
53. Yuan, M. and Y. Zhang, *Test for the parametric part in partial functional linear regression based on B-spline.* Communications in Statistics-Simulation and Computation, 2021. **50**(1): p. 1-15.
54. Maistre, S. and V. Patilea, *Testing for the significance of functional covariates.* Journal of Multivariate Analysis, 2020. **179**: p. 104648.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

55. Kong, D., A.-M. Staicu, and A. Maity, *Classical testing in functional linear models.* Journal of nonparametric statistics, 2016. **28**(4): p. 813-838.
56. Cuesta-Albertos, J.A., et al., *Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes.* Annals of Statistics, 2019. **47**(1): p. 439-467.
57. Lai, T., Z. Zhang, and Y. Wang, *Testing independence and goodness-of-fit jointly for functional linear models.* Journal of the Korean Statistical Society, 2021. **50**(2): p. 380-402.
58. Lee, C., X. Zhang, and X. Shao, *Testing conditional mean independence for functional data.* Biometrika, 2020. **107**(2): p. 331-346.
59. Li, Q., X. Tan, and L. Wang, *Testing for error correlation in partially functional linear regression models.* Communications in Statistics-Theory and Methods, 2021. **50**(3): p. 747-761.
60. Lu, Y., J. Du, and Z. Sun, *Functional partially linear quantile regression model.* Metrika, 2014. **77**(2): p. 317-332.
61. Tang, Q. and L. Cheng, *Partial functional linear quantile regression.* Science China Mathematics, 2014. **57**(12): p. 2589-2608.
62. Qingguo, T. and L. Kong, *Quantile regression in functional linear semiparametric model.* Statistics, 2017. **51**(6): p. 1342-1358.
63. Ding, H., et al., *Semi-functional partial linear quantile regression.* Statistics & Probability Letters, 2018. **142**: p. 92-101.
64. Yu, D., L. Kong, and I. Mizera, *Partial functional linear quantile regression for neuroimaging data analysis.* Neurocomputing, 2016. **195**: p. 74-87.
65. Yu, P., et al., *Composite quantile estimation in partial functional linear regression model with dependent errors.* Metrika, 2019. **82**(6): p. 633-656.
66. Yu, P., et al., *Composite Quantile Estimation in Partial Functional Linear Regression Model Based on Polynomial Spline.* Acta Mathematica Sinica, English Series, 2021. **37**(10): p. 1627-1644.
67. Zhu, H., et al., *Extreme quantile estimation for partial functional linear regression models with heavy‐tailed distributions.* Canadian Journal of Statistics, 2021.
68. Yu, P., J. Du, and Z. Zhang, *Testing linearity in partial functional linear quantile regression model based on regression rank scores.* Journal of the Korean Statistical Society, 2021. **50**: p. 214-232.
69. Yu, P., J. Du, and Z. Zhang, *Varying-coefficient partially functional linear quantile regression models.* Journal of the Korean Statistical Society, 2017. **46**(3): p. 462-475.
70. Ding, H., R. Zhang, and J. Zhang, *Quantile estimation for a hybrid model of functional and varying coefficient regressions.* Journal of Statistical Planning and Inference, 2018. **196**: p. 1-18.
71. Xu, D. and J. Du, *Nonparametric quantile regression estimation for functional data with responses missing at random.* Metrika, 2020. **83**(8): p. 977-990.
72. Maity, A. and J.Z. Huang, *Partially linear varying coefficient models stratified by a functional covariate.* Statistics & probability letters, 2012. **82**(10): p. 1807-1814.
73. Peng, Q.-Y., J.-J. Zhou, and N.-S. Tang, *Varying coefficient partially functional linear regression models.* Statistical Papers, 2016. **57**(3): p. 827-841.
74. Feng, S. and L. Xue, *Partially functional linear varying coefficient model.* Statistics, 2016. **50**(4): p. 717-732.
75. Liu, W., J. Sun, and J. Yang, *Rank-based estimation in varying coefficient partially functional linear regression models.* Communications in Statistics-Theory and Methods, 2020: p. 1-14.
76. Aneiros, G., F. Ferraty, and P. Vieu, *Variable selection in partial linear regression with functional covariate.* Statistics, 2015. **49**(6): p. 1322-1347.
77. Du, J., D. Xu, and R. Cao, *Estimation and variable selection for partially functional linear models.* Journal of the Korean Statistical Society, 2018. **47**(4): p. 436-449.
78. Kong, D., et al., *Partially functional linear regression in high dimensions.* Biometrika, 2016. **103**(1): p. 147-159.
79. Tang, Q. and P. Jin, *Estimation and variable selection for partial functional linear regression.* AStA Advances in Statistical Analysis, 2019. **103**(4): p. 475-501.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

80. Wang, G., X.N. Feng, and M. Chen, *Functional partial linear single‐index model.* Scandinavian Journal of Statistics, 2016. **43**(1): p. 261-274.
81. Yu, P., J. Du, and Z. Zhang, *Single-index partially functional linear regression model.* Statistical Papers, 2020. **61**(3): p. 1107-1123.
82. Feng, S., et al., *Estimation in functional single-index varying coefficient model.* Journal of Statistical Planning and Inference, 2021. **214**: p. 62-75.
83. Alahiane, M., et al., *Partially Linear Generalized Single Index Models for Functional Data (PLGSIMF).* Stats, 2021. **4**(4): p. 793-813.
84. Ling, N. and P. Vieu, *On semiparametric regression in functional data analysis.* Wiley Interdisciplinary Reviews: Computational Statistics, 2021: p. e1538.
85. Zhang, T., J. Meng, and B. Wang, *Partially function linear error-in-response models with validation data.* Journal of Systems Science and Complexity, 2017. **30**(3): p. 734-750.
86. Zhu, H., et al., *Estimation and testing for partially functional linear errors-in-variables models.* Journal of Multivariate Analysis, 2019. **170**: p. 296-314.
87. Zhu, H., R. Zhang, and G. Zhu, *Estimation and Inference in Semi-Functional Partially Linear Measurement Error Models.* Journal of Systems Science and Complexity, 2020. **33**(4): p. 1179-1199.
88. Zhu, H., R. Zhang, and H. Li, *Estimation on semi-functional linear errors-in-variables models.* Communications in Statistics-Theory and Methods, 2019. **48**(17): p. 4380-4393.
89. Aneiros, G. and P. Vieu, *Partial linear modelling with multi-functional covariates.* Computational Statistics, 2015. **30**(3): p. 647-671.
90. Xu, W., et al., *Estimation and inference in partially functional linear regression with multiple functional covariates.* Journal of Statistical Planning and Inference, 2020. **209**: p. 44-61.
91. Hu, Y., L. Xue, and S. Feng, *Empirical likelihood inference for partial functional linear model with missing responses.* Communications in Statistics-Theory and Methods, 2018. **47**(19): p. 4673-4691.
92. Ling, N., et al., *Semi-functional partially linear regression model with responses missing at random.* Metrika, 2019. **82**(1): p. 39-70.
93. Zhou, J. and Q. Peng, *Estimation for functional partial linear models with missing responses.* Statistics & Probability Letters, 2020. **156**: p. 108598.
94. Cao, R., T. Xie, and P. Yu, *Rank method for partial functional linear regression models.* Journal of the Korean Statistical Society, 2020: p. 1-26.
95. Xie, T., R. Cao, and P. Yu, *Rank-Based Test for Partial Functional Linear Regression Models.* Journal of Systems Science and Complexity, 2020. **33**(5): p. 1571-1584.
96. Lian, H., *Functional partial linear model.* Journal of Nonparametric Statistics, 2011. **23**(1): p. 115-128.
97. Aneiros, G., N. Ling, and P. Vieu, *Error variance estimation in semi-functional partially linear regression models.* Journal of Nonparametric Statistics, 2015. **27**(3): p. 316-330.
98. Aneiros, G., et al., *Bootstrap in semi-functional partial linear regression under dependence.* Test, 2018. **27**(3): p. 659-679.
99. Hu, Y., et al., *Skew-normal partial functional linear model and homogeneity test.* Journal of Statistical Planning and Inference, 2020. **204**: p. 116-127.
100. Du, J., et al., *Estimation for generalized partially functional linear additive regression model.* Journal of Applied Statistics, 2019. **46**(5): p. 914-925.
101. Cao, R., et al., *FPCA-based estimation for generalized functional partially linear models.* Statistical Papers, 2020. **61**(6): p. 2715-2735.
102. Yang, S.J., et al., *Functional linear regression model with randomly censored data: Predicting conversion time to Alzheimer's disease.* Computational Statistics & Data Analysis, 2020. **150**: p. 107009.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 November 2021 doi:10.20944/preprints202111.0310.v1

103. Xu, W., et al., *Two-sample functional linear models with functional responses.* Journal of Statistical Planning and Inference, 2021.
104. Pan, Y., et al., *Reinforced risk prediction with budget constraint using irregularly measured data from electronic health records.* Journal of the American Statistical Association, 2021(just-accepted): p. 1-24.
105. Vilar, J.M., R. Cao, and G. Aneiros, *Forecasting next-day electricity demand and price using nonparametric functional methods.* International Journal of Electrical Power & Energy Systems, 2012. **39**(1): p. 48-55.
106. Aneiros, G., et al., *Functional prediction for the residual demand in electricity spot markets.* IEEE Transactions on Power Systems, 2013. **28**(4): p. 4201-4208.
107. Wang, G., Y. Su, and L. Shu, *One-day-ahead daily power forecasting of photovoltaic systems based on partial functional linear regression models.* Renewable Energy, 2016. **96**: p. 469-478.
108. Vilar, J., G. Aneiros, and P. Raña, *Prediction intervals for electricity demand and price using functional data.* International Journal of Electrical Power & Energy Systems, 2018. **96**: p. 457-472.
109. Pani, A.K. and N. Nayak, *Photovoltaic Power Forecasting by Evolutionary Algorithm-Based Improved Extreme Learning Machine*, in *Advances in Electrical Control and Signal Systems*. 2020, Springer. p. 109-129.
110. Scheffler, A., et al., *Hybrid principal components analysis for region-referenced longitudinal functional EEG data.* Biostatistics, 2020. **21**(1): p. 139-157.
111. Jang, J.H., *Principal component analysis of hybrid functional and vector data.* Statistics in Medicine, 2021.
112. Harezlak, J., D. Ruppert, and M.P. Wand, *Semiparametric regression with R*. 2018: Springer.
113. Febrero-Bande, M., W. González-Manteiga, and M.O. de la Fuente, *Variable selection in functional additive regression models.* Computational Statistics, 2019. **34**(2): p. 469-487.
114. Luo, R. and X. Qi, *Function-on-function linear regression by signal compression.* Journal of the American Statistical Association, 2017. **112**(518): p. 690-705.
115. Scheffler, A.W., et al., *Covariate‐adjusted region‐referenced generalized functional linear model for EEG data.* Statistics in medicine, 2019. **38**(30): p. 5587-5602.