

Article

The Curve Estimation of Combined Truncated Spline and Fourier Series Estimators for Multiresponse Nonparametric Regression

Helida Nurcahayani ^{1,2}, I Nyoman Budiantara ^{1,*} and Ismaini Zain ¹

¹ Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia; helida.nur@bps.go.id (H.N.); ismaini_z@statistika.its.ac.id (I.Z.)

² BPS—Statistics of Daerah Istimewa Yogyakarta Province, Bantul 55183, Indonesia

* Correspondence: nyomanbudiantara65@gmail.com

Abstract: Nonparametric regression becomes a potential solution if the parametric regression assumption is too restrictive while the regression curve is assumed to be known. In multivariable nonparametric regression, the pattern of each predictor variable's relationship with the response variable is not always the same; thus, a combined estimator is recommended. In addition, regression modeling sometimes involves more than one response, i.e., multiresponse situations. Therefore, we propose a new estimation method of performing multiresponse nonparametric regression with a combined estimator. The objective is to estimate the regression curve using combined truncated spline and Fourier series estimators for multiresponse nonparametric regression. The regression curve estimation of the proposed model is obtained via two-stage estimation: (1) penalized weighted least square and (2) weighted least square. Simulation data with sample size variation and different error variance were applied, where the best model satisfied the result through a large sample with small variance. Additionally, the application of the regression curve estimation to a real dataset of human development index indicators in East Java Province, Indonesia, showed that the proposed model had better performance than uncombined estimators. Moreover, an adequate coefficient of determination of the best model indicated that the proposed model successfully explained the data variation.

Keywords: combined estimator; Fourier series; multiresponse; nonparametric regression; truncated spline

Citation: Nurcahayani, H.; Budiantara, I.N.; Zain, I. The Curve Estimation of Combined Truncated Spline and Fourier Series Estimators for Multi-Response Nonparametric Regression. *Mathematics* **2021**, *9*, 1141. <https://doi.org/10.3390/math9101141>

Academic Editor: Lev Klebanov

Received: 11 April 2021

Accepted: 17 May 2021

Published: 18 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the renowned methods of regression analysis, parametric regression has been used for many years in various scientific fields. However, some parametric regression assumptions are too restrictive, such as identifying a regression curve's shape with some prespecified functional form (e.g., linear, quadratic, or cubic) [1]. In real datasets, not all regression curves have a visible pattern due to the absence of relationship information between response and predictor variables. Therefore, in such scenarios, nonparametric regression analysis is more recommended [2]. Nonparametric regression is able to degrade the misspecification risk because data inherently represent the real shape of a regression curve without interference due to the researcher's subjectivity [2].

To date, researchers have investigated various functions for estimating the regression curve in nonparametric regression, such as spline [3–6], Fourier series [7–9], kernel [10–12], polynomial [13–15], and wavelet [16–18] functions. The present study explores the nonparametric regression approach to analyzing spatial data, i.e., the use of the nonparametric truncated spline function in geographically weighted regression [19,20]. In nonparametric regression with more than one predictor variable

(multivariable), the pattern of the relationship between each predictor variable and the response variable is not always the same; therefore, in this study, we employ a nonparametric regression model with a combined estimator. Considering the concept of semiparametric regression, some studies have proposed combined estimator models to estimate the regression curve, such as spline and kernel functions [21–23]. Similarly, in [24], the researchers developed a combined estimator of kernel and Fourier series, whereas in [25–27], the researchers presented nonparametric regression with a mixed estimator of truncated spline and Fourier series. Recently, researchers have also been giving attention to a mixed model of longitudinal data, as carried out by [28].

So far, studies on nonparametric regression with combined estimators have not dealt with more than two responses, i.e., multiresponse nonparametric regression. Numerous studies on nonparametric regression have been applied to many scientific fields, such as sociodemographic analysis [23–26], finance [5,29], climatology [4], and economics [30], wherein some cases have involved real data with two or more correlated response variables. Therefore, this study makes a major contribution to the research on multiresponse nonparametric regression by using a combined estimator. Of the several types of regression functions for nonparametric regression, the truncated spline function has the major advantages of high flexibility, good visual interpretation, and the ability to handle smooth function characters and data that have changed behavior at certain subintervals [2]. On the other hand, the Fourier series function is used to estimate the regression curve if the data are smooth and follow a repeated pattern at specific intervals [9]. The Fourier series function is also a trigonometric polynomial that can adjust the data's local nature effectively. In this study, we adopt a cosine Fourier series function [9], which follows a trend line as the regression curve estimator. A Fourier series with a cosine function was employed since it is an even function, such that the second derivative usually mathematically obtains a scalar or a nonzero value. Furthermore, the penalties in the penalized least square method can be well defined. Considering some of the advantages of these two functions, as outlined above, this study highlights the combined use of truncated spline and Fourier series estimators for multiresponse nonparametric regression.

One of the most well-known tools for the estimation of regression is the ordinary least square (OLS) method. However, the OLS method cannot be directly used in nonparametric regression because of the unknown curve shape. By utilizing smooth, continuous, and differentiable properties, as well as other components, the OLS method was modified with conditional optimization to create the penalized least-square (PLS) regression method [31]. The combined use of estimators for multiresponse nonparametric regression leads to the use of an error variance–covariance matrix as a weighting matrix. Therefore, PLS optimization is still added with a weight; accordingly, the proposed model was obtained through penalized weighted least square (PWLS) optimization. To evaluate the proposed model performance, we employed a simulation study with three different sample sizes and three error variances and applied the proposed model to a real dataset.

Given the view above, the objective of this study is to obtain a curve estimation using combined truncated spline and Fourier series estimators for multiresponse nonparametric regression. This aim is followed by estimating an error variance–covariance matrix as a weighting matrix. To estimate the proposed model, a two-stage estimation method was used. Stage 1 was to estimate the Fourier series component using PWLS optimization. The following stage was to estimate the truncated spline component by utilizing weighted least square (WLS) optimization. In addition to curve estimation, we attempted to simulate and apply the proposed model to a dataset of the HDI in East Java Province, Indonesia. The paper is organized as follows: Section 1 provides an overview of the topic while identifying the knowledge gap and stating the aims of the research. Section 2 describes the truncated spline function and the Fourier series function, along with the PWLS method. Section 3 presents the findings of the regression curve estimation via the combined use of truncated spline and Fourier series estimators for multiresponse

nonparametric regression, along with an estimation of the variance–covariance matrix, smoothing parameter selection, and a simulation study following the application of a real dataset. Section 4 gives a brief conclusion and makes future recommendations.

2. Materials and Methods

2.1. Multiresponse Nonparametric Regression, Truncated Spline Function, and Fourier Series Function

Given paired data $(y_{1i}, y_{2i}, \dots, y_{ri}, x_{1i}, x_{2i}, \dots, x_{pi}, z_{1i}, z_{2i}, \dots, z_{qi})$, the relationship between response $(y_{1i}, y_{2i}, \dots, y_{ri})$ and predictor $(x_{1i}, x_{2i}, \dots, x_{pi}, z_{1i}, z_{2i}, \dots, z_{qi})$ variables is assumed to follow multiresponse nonparametric regression, as follows:

$$y_{hi} = \mu_{hi}(x_{1i}, x_{2i}, \dots, x_{pi}, z_{1i}, z_{2i}, \dots, z_{qi}) + \varepsilon_{hi}, \varepsilon_{hi} \sim N(0, \sigma_h^2) \tag{1}$$

where $h = 1, 2, \dots, r$ and $i = 1, 2, \dots, n$. In this study, r is defined as the number of response variables, and n refers to the number of observations for each response and predictor variable. For convenience, Equation (1) can be rewritten in the following matrix form:

$$\mathbf{y} = \boldsymbol{\mu}(\mathbf{x}, \mathbf{z}) + \boldsymbol{\varepsilon} \tag{2}$$

The regression curve in Equation (1) can be assumed to be an additive model.

$$\mu_{hi}(x_{1i}, x_{2i}, \dots, x_{pi}, z_{1i}, z_{2i}, \dots, z_{qi}) = \sum_{j=1}^p f_{hj}(x_{ji}) + \sum_{k=1}^q g_{hk}(z_{ki}). \tag{3}$$

In Equation (3), the regression curve $f_{hj}(x_{ji}); j = 1, 2, \dots, p$ is assumed to be smooth and is approached by a truncated spline function. Meanwhile, the regression curve $g_{hk}(z_{ki}); k = 1, 2, \dots, q$ is approached by a Fourier series function, which assumes that the regression curve is unknown and is contained in continuous space $C[0, \pi]$. Hence, the multiresponse nonparametric regression in Equation (3) can be written as

$$y_{hi} = \sum_{j=1}^p f_{hj}(x_{ji}) + \sum_{k=1}^q g_{hk}(z_{ki}) + \varepsilon_{hi}, \varepsilon_{hi} \sim N(0, \sigma_h^2). \tag{4}$$

Under the assumption $\varepsilon_{hi} \sim N(0, \sigma_h^2)$, as described in [32], $\text{corr}(\varepsilon_{hi}, \varepsilon_{\ell i}) = \rho$ for $h \neq \ell; h, \ell = 1, 2, \dots, r$. This term refers to situations in which the response variables y_{hi} and $y_{\ell i}$ are a pair, such that a correlation between the h -th response and the ℓ -th response yields $\text{corr}(\varepsilon_{hi}, \varepsilon_{\ell i}) = \rho$, and zero otherwise. The error correlation between responses is the same for every response, defined as $\rho = \frac{\text{cov}(\varepsilon_{hi}, \varepsilon_{\ell i})}{\sqrt{\sigma_{hh}\sigma_{\ell\ell}}}$; $i = 1, 2, \dots, n$; $h \neq \ell; h, \ell = 1, 2, \dots, r$.

Thus, the regression curve $f_{hj}(x_{ji})$ in Equation (4) is approached by a linear truncated spline function with knots $K_{hj1}, K_{hj2}, \dots, K_{hju}$, as follows:

$$f_{hj}(x_{ji}) = \alpha_{hj}x_{ji} + \sum_{s=1}^u \beta_{hjs}(x_{ji} - K_{hjs})_+, \tag{5}$$

with the truncated function

$$(x_{ji} - K_{hjs})_+ = \begin{cases} (x_{ji} - K_{hjs}), & x_{ji} \geq K_{hjs} \\ 0, & x_{ji} < K_{hjs} \end{cases}$$

Adopting the Fourier series function in [9], $g_{hk}(z_{ki})$ in Equation (4) is approached by a Fourier series cosine function following the trend line $(b_{hk}z_{ki})$, as given in Equation (6):

$$g_{hk}(z_{ki}) = b_{hk}z_{ki} + \frac{1}{2}a_{ohk} + \sum_{t=1}^T a_{thk} \cos tz_{ki}. \tag{6}$$

2.2. Penalized Weighted Least Square

Suppose that the dataset $(y_{1i}, y_{2i}, \dots, y_{ri}, x_{1i}, x_{2i}, \dots, x_{pi}, z_{1i}, z_{2i}, \dots, z_{qi})$ follows the multiresponse nonparametric regression in Equation (1) and is rewritten in matrix form as Equation (2). The random error $\boldsymbol{\varepsilon}$ in Equation (2) is normally distributed, with mean $\mathbf{0}$ and error variance–covariance matrix \mathbf{W} , such that it can be written as $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{W})$. In particular, \mathbf{W} as a weighting matrix plays an important role in accommodating the correlation between responses. If there is a correlation between responses, the correlation is defined as $\rho = \frac{\sigma_{h\ell}}{\sqrt{\sigma_{hh}\sigma_{\ell\ell}}}$, such that $\sigma_{h\ell} = \rho\sqrt{\sigma_{hh}\sigma_{\ell\ell}}$. Note that \mathbf{I} is the identity matrix. When all observations come in pairs, \mathbf{W} has $\sigma_{h,\ell}$ elements, as described below:

$$\begin{aligned} \mathbf{W} &= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1r} & \sigma_{2r} & \cdots & \sigma_{rr} \end{pmatrix} \otimes \mathbf{I} \\ &= \boldsymbol{\Sigma} \otimes \mathbf{I} \end{aligned}$$

The regression curve of the combined truncated spline and Fourier series estimators for multiresponse nonparametric regression was estimated by carrying out the following PWLS optimization:

$$\text{Min}_{g_k \in C[0, \pi]} \left\{ N^{-1} \sum_{h=1}^r \sum_{i=1}^n w_{hi} \left(y_{hi} - \sum_{j=1}^p f_{hj}(x_{ji}) - \sum_{k=1}^q g_{hk}(z_{ki}) \right)^2 + \sum_{k=1}^q \lambda_k \int_0^\pi \frac{2}{\pi} (g_k''(z_k))^2 dz_k \right\}, \tag{7}$$

where N refers to the number of observations for all response variables, alternatively written as $N = \sum_{h=1}^r n_h$. The first component in Equation (7) is a function that measures the goodness-of-fit (GoF), while the second component is the penalty. In addition, w_{hi} is a weighted component, and λ_k serves as a positive smoothing parameter that controls the balance between the GoF and the penalty. In this study, during the process of regression curve estimation, the parameters w_{hi} and λ_k are given. Note that the superscript T refers to transpose and the superscript “ $''$ ” refers to the second derivative of the function.

As stated in the Introduction section, we estimated the regression curve of the proposed model in two stages. The first stage involves completing the estimation of the regression curve $g_{hk}(z_{ki})$ through PWLS optimization, which results in Theorem 1 (see the Section 3.1). Subsequently, the second stage involves completing the estimated regression curve $f_{hj}(x_{ji})$ using WLS optimization, which results in Theorem 2 (see the Section 3.1).

3. Results

3.1. Curve Estimation of Combined Truncated Spline and Fourier Series Estimators for Multiresponse Nonparametric Regression

As previously stated, PWLS and WLS optimization were used to obtain combined truncated spline and Fourier series estimators for multiresponse nonparametric regression. Therefore, some lemmas and theorems were required. Whenever Lemma 1 presents the GoF, Lemma 2 presents the penalty of the PWLS optimization. By using the results of Lemma 1 and Lemma 2, the first stage estimation is obtained, which is formulated in Theorem 1. The second stage involves estimating the regression curve using WLS optimization, as presented in Lemma 3, with the main result presented in Theorem 2. All the proofs of the lemmas and theorems are presented in Appendix A to Appendix E.

Lemma 1. *If the multiresponse nonparametric regression model is written as Equation (4) and the regression curve $g_{hk}(z_{ki})$ is as given in Equation (6), the GoF can be formulated as follows:*

$$R(\mathbf{g}_1, \dots, \mathbf{g}_r) = N^{-1}(\mathbf{v} - \mathbf{Z}\mathbf{a})^T \mathbf{W}(\mathbf{v} - \mathbf{Z}\mathbf{a}),$$

where \mathbf{W} is the $nh \times nh$ weighting matrix,

$$\mathbf{v} = [\mathbf{v}_1^T \quad \mathbf{v}_2^T \quad \dots \quad \mathbf{v}_r^T]^T; \mathbf{v}_h = [v_{h1} \quad v_{h2} \quad \dots \quad v_{hn}]^T; v_{hi} = y_{hi} - \sum_{j=1}^p f_{hj}(x_{ji})$$

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{Z}_r \end{bmatrix}, \mathbf{a} = [a_1^T \quad a_2^T \quad \dots \quad a_r^T]^T,$$

$$\mathbf{z}_h = \begin{bmatrix} z_{11} & 1/2 & \cos 1z_{11} & \cos 2z_{11} & \dots & \cos Tz_{11} & \dots & z_{q1} & 1/2 & \cos 1z_{q1} & \cos 2z_{q1} & \dots & \cos Tz_{q1} \\ z_{12} & 1/2 & \cos 1z_{12} & \cos 2z_{12} & \dots & \cos Tz_{12} & \dots & z_{q2} & 1/2 & \cos 1z_{q2} & \cos 2z_{q2} & \dots & \cos Tz_{q2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{1n} & 1/2 & \cos 1z_{1n} & \cos 2z_{1n} & \dots & \cos Tz_{1n} & \dots & z_{qn} & 1/2 & \cos 1z_{qn} & \cos 2z_{qn} & \dots & \cos Tz_{qn} \end{bmatrix},$$

$$\mathbf{a}_h = [b_{1h} \quad a_{01h} \quad a_{11h} \quad a_{21h} \quad \dots \quad a_{T1h} \quad \dots \quad b_{qh} \quad a_{0qh} \quad a_{1qh} \quad a_{2qh} \quad \dots \quad a_{Tqh}]^T.$$

The proof of Lemma 1 is given in Appendix A.

Lemma 2. *If the multiresponse nonparametric regression model is written as Equation (4) and the regression curve $g_{hk}(z_{ki})$ is as presented in Equation (6), the penalty component is as follows:*

$$P(\lambda_1, \lambda_2, \dots, \lambda_q) = \mathbf{a}^T \mathbf{D}(\lambda) \mathbf{a}$$

where

$$\mathbf{D}(\lambda) = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \lambda_1 1^4 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_1 2^4 & \dots & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda_1 T^4 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \lambda_q 1^4 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \lambda_q 2^4 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 & \dots & \lambda_q T^4 \end{bmatrix}$$

A complete description of the proof of Lemma 2 can be found in Appendix B.

Having discussed how to construct the GoF and the penalty, Theorem 1 addresses potential solutions to the first-stage estimation using PWLS optimization in Equation (7).

Theorem 1. *If the GoF and the penalty of the model are given in Lemma 1 and Lemma 2, the curve estimation for multiresponse nonparametric regression can be attained from PWLS optimization, as follows:*

$$\hat{\mathbf{g}}_{(k,\lambda,T)}(\mathbf{x}, \mathbf{z}) = (\mathbf{Z}[\mathbf{Z}^T \mathbf{W} \mathbf{Z} + N\mathbf{D}(\lambda)]^{-1} \mathbf{Z}^T \mathbf{W}) \mathbf{v}.$$

The evidence of Theorem 1 is given in Appendix C. The second stage involves estimating the regression curve using WLS optimization, as described in Lemma 3, with the result in Theorem 2.

Lemma 3. *If the regression curve $f_{hj}(x_{ji})$ is as described in Equation (5), WLS optimization can be obtained as*

$$[(\mathbf{I} - \mathbf{H})(\mathbf{y} - \mathbf{J}\boldsymbol{\gamma})]^T \mathbf{W} [(\mathbf{I} - \mathbf{H})(\mathbf{y} - \mathbf{J}\boldsymbol{\gamma})],$$

where $\mathbf{H} = (\mathbf{Z}[\mathbf{Z}^T \mathbf{W} \mathbf{Z} + N\mathbf{D}(\lambda)]^{-1} \mathbf{Z}^T \mathbf{W})$,

$$\mathbf{s}_i = \begin{bmatrix} (x_{11} - K_{h11})_+ & (x_{11} - K_{h12})_+ & \cdots & (x_{11} - K_{h1u})_+ & \cdots & (x_{p1} - K_{hp1})_+ & (x_{11} - K_{hp2})_+ & \cdots & (x_{11} - K_{hpu})_+ \\ (x_{12} - K_{h11})_+ & (x_{12} - K_{h12})_+ & \cdots & (x_{12} - K_{h1u})_+ & \cdots & (x_{p2} - K_{hp1})_+ & (x_{12} - K_{hp2})_+ & \cdots & (x_{12} - K_{hpu})_+ \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ (x_{1n} - K_{h11})_+ & (x_{1n} - K_{h12})_+ & \cdots & (x_{1n} - K_{h1u})_+ & \cdots & (x_{pn} - K_{hp1})_+ & (x_{1n} - K_{hp2})_+ & \cdots & (x_{1n} - K_{hpu})_+ \end{bmatrix}, \mathbf{y} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_r \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_r \end{bmatrix}$$

$$\alpha_h = [\alpha_{1h} \ \alpha_{2h} \ \cdots \ \alpha_{ph}]^T, \beta_h = [\beta_{h11} \ \beta_{h12} \ \cdots \ \beta_{h1u} \ \cdots \ \beta_{hp1} \ \beta_{hp2} \ \cdots \ \beta_{hpu}]^T.$$

As a note, Lemma 3’s proof is provided in Appendix D.

Theorem 2. *If WLS optimization is given by Lemma 3, the regression curve estimation for multiresponse nonparametric regression can be attained from WLS optimization such that*

$$\hat{\mathbf{f}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) = \mathbf{J}[\mathbf{J}^T(\mathbf{I} - 2\mathbf{H}^T)\mathbf{W}\mathbf{J} + \mathbf{J}^T\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{J}]^{-1}[\mathbf{J}^T(\mathbf{H}^T - \mathbf{I})\mathbf{W}(\mathbf{H} - \mathbf{I})]\mathbf{y}. \tag{8}$$

$$= \mathbf{J}\mathbf{K}^{-1}\mathbf{L}\mathbf{y}, \tag{9}$$

where $\mathbf{K} = [\mathbf{J}^T(\mathbf{I} - 2\mathbf{H}^T)\mathbf{W}\mathbf{J} + \mathbf{J}^T\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{J}]$ and $\mathbf{L} = [\mathbf{J}^T(\mathbf{H}^T - \mathbf{I})\mathbf{W}(\mathbf{H} - \mathbf{I})]$.

The proof of Theorem 2 is given in Appendix E.

After obtaining $\hat{\mathbf{f}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z})$ in Theorem 2, the result of $\hat{\mathbf{g}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z})$ in Theorem 1 can be rewritten as Equation (12). First, substituting Equation (9) into Equation (A9) gives Equation (10):

$$\begin{aligned} \hat{\mathbf{a}} &= [\mathbf{Z}^T\mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1}\mathbf{Z}^T\mathbf{W}\mathbf{v} \\ &= [\mathbf{Z}^T\mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1}\mathbf{Z}^T\mathbf{W}(\mathbf{y} - \mathbf{f}) \\ &= [\mathbf{Z}^T\mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1}\mathbf{Z}^T\mathbf{W}(\mathbf{y} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L}\mathbf{y}) \\ &= [\mathbf{Z}^T\mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1}\mathbf{Z}^T\mathbf{W}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})\mathbf{y}. \end{aligned} \tag{10}$$

If the result of Equation (10) is substituted into Equation (A10), the regression curve estimation of the Fourier series can be rewritten as Equation (12). As a note, Equation (A9) and Equation (A10) can be found in Appendix C.

$$\begin{aligned} \hat{\mathbf{g}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) &= \mathbf{Z}\hat{\mathbf{a}} \\ &= \mathbf{Z}([\mathbf{Z}^T\mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1}\mathbf{Z}^T\mathbf{W}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})\mathbf{y}) \end{aligned} \tag{11}$$

$$= \mathbf{H}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})\mathbf{y}. \tag{12}$$

Another main finding of this study is the regression curve of the combined truncated spline and Fourier series estimators for multiresponse nonparametric regression, as presented in Corollary 1.

Corollary 1. *According to the main result in Equation (8) and Equation (11), the regression curve estimation of the combined truncated spline and Fourier series estimators for the additive model is presented below:*

$$\hat{\boldsymbol{\mu}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) = [\mathbf{J}\mathbf{K}^{-1}\mathbf{L} + \mathbf{H}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})]\mathbf{y}.$$

Proof. By using an additive model in Equation (3), the regression curve estimation of the combined estimator for multiresponse nonparametric regression can be written in the following matrix form:

$$\hat{\boldsymbol{\mu}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) = \hat{\mathbf{f}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) + \hat{\mathbf{g}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}).$$

Based on the result of $\hat{\mathbf{f}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z})$ in Equation (8) and $\hat{\mathbf{g}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z})$ in Equation (11), thus $\hat{\boldsymbol{\mu}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z})$ can be drawn as

$$\begin{aligned} \hat{\boldsymbol{\mu}}_{(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})}(\mathbf{x}, \mathbf{z}) &= \{[\mathbf{J}(\mathbf{J}^T(\mathbf{I} - 2\mathbf{H}^T)\mathbf{W}\mathbf{J} + \mathbf{J}^T\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{J})^{-1}]\mathbf{J}^T(\mathbf{H}^T - \mathbf{I})\mathbf{W}(\mathbf{H} - \mathbf{I})\mathbf{y}\} \\ &\quad + \{\mathbf{Z}([\mathbf{Z}^T\mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\boldsymbol{\lambda})]^{-1}\mathbf{Z}^T\mathbf{W}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})\mathbf{y})\}. \end{aligned} \tag{13}$$

$$= \mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})\mathbf{y}$$

For simplification, by substituting Equation (9) and Equation (12) into Equation (13), $\hat{\boldsymbol{\mu}}_{(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})}(\mathbf{x}, \mathbf{z})$ can also be expressed as

$$\hat{\boldsymbol{\mu}}_{(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})}(\mathbf{x}, \mathbf{z}) = [\mathbf{J}\mathbf{K}^{-1}\mathbf{L} + \mathbf{H}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})]\mathbf{y}. \quad \square$$

3.2. Estimation of Error Variance–Covariance Matrix

The result in Equation (13) shows that curve estimation of the proposed model leads to the use of an error variance–covariance matrix as a weighting matrix. Hence, estimation of the error variance–covariance matrix is presented in Theorem 3.

Theorem 3. *If the regression curve estimation is given by Corollary 1 and random error $\boldsymbol{\varepsilon}$ is normally distributed with mean $\mathbf{0}$ and error variance–covariance matrix \mathbf{W} such that it can be written as $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{W})$, the error variance–covariance matrix estimation is as follows:*

$$\hat{\mathbf{W}} = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \cdots & \hat{\sigma}_{1r} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} & \cdots & \hat{\sigma}_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{1r} & \hat{\sigma}_{2r} & \cdots & \hat{\sigma}_{rr} \end{pmatrix} \otimes \mathbf{I}$$

where

$$\begin{aligned} \hat{\sigma}_{11} &= \frac{[\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]^T [\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]}{n}, \quad \hat{\sigma}_{22} = \frac{[\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]^T [\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]}{n}, \quad \hat{\sigma}_{rr} = \frac{[\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]^T [\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]}{n}, \\ \hat{\sigma}_{12} &= \frac{[\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]^T [\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]}{n}, \quad \hat{\sigma}_{1r} = \frac{[\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]^T [\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]}{n}, \quad \hat{\sigma}_{2r} = \frac{[\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]^T [\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]}{n}, \end{aligned}$$

$$\mathbf{F} = (([\mathbf{I} - (\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{J}]^{-1}(\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T) - ((\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T)),$$

$$\mathbf{G} = (([\mathbf{I} - (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{J}(\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T\mathbf{Z}]^{-1}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T) - ((\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{J}(\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T)).$$

The proof of Theorem 3 is given in Appendix F.

3.3. Smoothing Parameter Selection

Another critical step in nonparametric regression modeling is selecting the optimal knot, oscillation parameter, and smoothing parameter. A large smoothing parameter produces a smoother estimator function but constrains the capability mapping of data. In contrast, if the smoothing parameter is too small, the function estimator is rougher [33]. We concluded that a suitable method was required to determine the optimal smoothing parameters. The smoothing parameter selection in this study was based on the GCV method, as conducted by [33]; the detailed procedure and optimal properties of this method have been carried out by [34]. The modified GCV method for combined truncated spline and Fourier series estimators in the multiresponse nonparametric regression model is as follows:

$$GCV(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T}) = \frac{MSE(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})}{[N^{-1}\text{trace}(\mathbf{I} - \mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T}))]^2} \tag{14}$$

where

$$\begin{aligned} MSE(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T}) &= N^{-1}(\mathbf{y} - \hat{\boldsymbol{\mu}})^T(\mathbf{y} - \hat{\boldsymbol{\mu}}) = N^{-1}(\mathbf{y} - \mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})\mathbf{y})^T(\mathbf{y} - \mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})\mathbf{y}) \\ &= N^{-1}\|\mathbf{I} - \mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})\mathbf{y}\|^2 \end{aligned} \tag{15}$$

By substituting Equation (15) into Equation (14), the optimum smoothing parameter is obtained by taking the minimum of modified GCV for the proposed model, as presented below

$$GCV(\mathbf{K}_{opt}, \boldsymbol{\lambda}_{opt}, \mathbf{T}_{opt}) = \underset{\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T}}{\text{Min}} \left\{ \frac{N^{-1} \|\mathbf{I} - \mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})\mathbf{y}\|^2}{[N^{-1} \text{trace}(\mathbf{I} - \mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T}))]^2} \right\}$$

where $\mathbf{C}(\mathbf{K}, \boldsymbol{\lambda}, \mathbf{T})$ is the regression curve estimation of the proposed model, as in Equation (13).

3.4. Simulation Study

Simulations with 100 replications were carried out to verify and validate the theoretical result of the proposed model. Using three data sample sizes ($n = 20, 50,$ and 100) and three different error variances ($\sigma^2 = 0.1, 0.5,$ and 1), random data were generated for multiresponse nonparametric regression. As shown in Figure 1, each predictor variable describes different functions, i.e., a polynomial function represents the truncated spline (x_i), while a trigonometry function with a trend represents the Fourier series function (z_i). As a note, Figure 1 only presents a partial scatterplot of the numerical example function for $n = 100$ and $\sigma^2 = 0.1$. The simulation study followed the model in Equation (4), so the polynomial and trigonometric functions used in this numerical example were obtained from the following function:

$$y_{1i} = 15(x_i - 1)(1 - x_i)^2 - 10.5z_i - 4 \cos(2\pi z_i) + \varepsilon_{1i}$$

$$y_{2i} = 14(x_i - 1)(1 - x_i)^2 - 7.5z_i - 4 \cos(2\pi z_i) + \varepsilon_{2i}$$

$$y_{3i} = 13(x_i - 1)(1 - x_i)^2 - 4.5z_i - 4 \cos(2\pi z_i) + \varepsilon_{3i}$$

In particular, the predictor variables x_i and z_i are generated from a Uniform (0,1) distribution, while random errors ε_{hi} are generated from a multivariate normal distribution. In addition, $\varepsilon_{1i}, \varepsilon_{2i},$ and ε_{3i} are correlated with $\text{corr}(\varepsilon_{1i}, \varepsilon_{2i}) = \text{corr}(\varepsilon_{1i}, \varepsilon_{3i}) = \text{corr}(\varepsilon_{2i}, \varepsilon_{3i}) = \rho = 0.9$. In order for ease of the computational process and based on the results of the partial scatterplot identification from Figure 1, the simulation study was carried out with the combination of three knots ($K = 1, 2, 3$) and three oscillation parameters ($T = 1, 2, 3$). To determine the optimal smoothing parameter, the minimum GCV criteria were used.

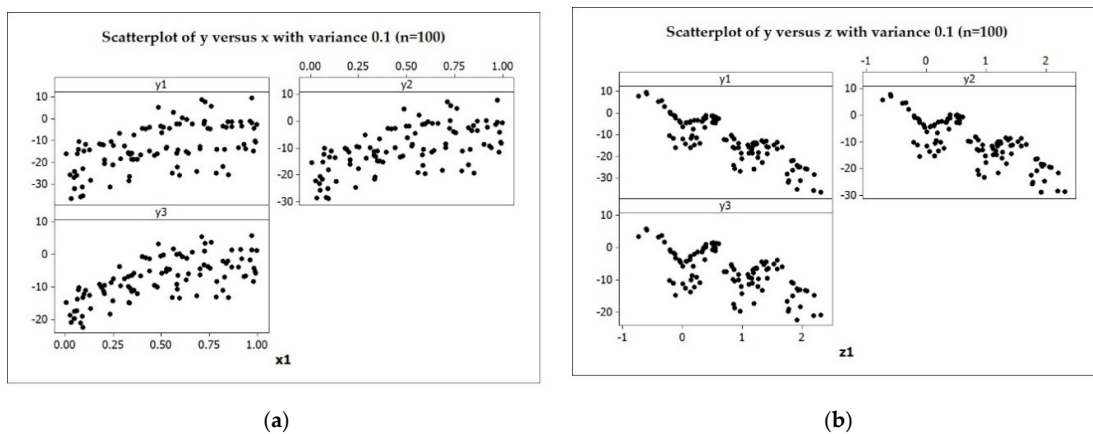


Figure 1. Partial scatterplot between the response variable and each predictor variable, which represents (a) a truncated spline function and (b) a Fourier series function.

Table 1 provides a comparison of the statistical results for $n = 100$ with the error variances, the number of knots, and oscillation parameters. A complete summary of the

statistical results for $n = 20$ and 50 is provided in Appendix G. Table 1 shows that the smallest GCV (6.098) occurs for variance $\sigma^2 = 0.1$ with the combination of the three knots–three oscillation model. This model yields adequate results, with coefficient of determination (R^2) = 93,202 and mean-square-error (MSE) = 5.675. Surprisingly, the same result is seen for the other sample sizes ($n = 20$ and 50), where the smallest GCV is obtained for variance $\sigma^2 = 0.1$. In addition, a comparison between sample sizes gives the result in which the smallest GCV is gained from the simulation with the largest sample number ($n = 100$). Thus, a smaller variance with a large sample size will produce a minimum GCV, which implies that the regression curve will be better estimated as well. These findings are consistent with [23], who developed mixed estimators in nonparametric regression for cross-sectional data.

Table 1. Summary of statistical results for $n = 100$ with error variances, the number of knots, and oscillation parameters

Variance	Number of Oscillations	Number of Knots	Generalized Cross-Validation (GCV)	R^2	Mean Square Error (MSE)
0.1	1	1	7.050	91.866	6.817
		2	6.996	92.100	6.719
		3	6.525	92.991	5.853
	2	1	6.413	92.964	5.875
		2	6.391	93.072	5.784
		3	6.386	93.167	5.704
	3	1	7.017	91.876	6.785
		2	6.435	93.066	5.789
		3	6.098 *	93.202	5.675
0.5	1	1	6.915	92.454	6.686
		2	6.759	92.716	6.492
		3	6.634	93.023	6.328
	2	1	8.183	91.121	7.913
		2	6.720	92.757	6.454
		3	6.383	93.671	5.640
	3	1	6.145	93.645	5.663
		2	6.260	93.660	5.650
		3	6.381	93.673	5.638
1.0	1	1	6.408	94.856	5.905
		2	6.479	94.906	5.847
		3	6.582	94.933	5.816
	2	1	6.358	94.895	5.860
		2	6.417	94.955	5.791
		3	6.490	95.004	5.734
	3	1	6.358	94.895	5.860
		2	6.417	94.955	5.791
		3	6.490	95.004	5.734

* The smallest value of GCV.

3.5. Data Application

This section attempts to provide the application of the proposed model to a real dataset. The proposed model was applied to four indicators of HDI data in 38 regencies/cities across East Java Province, Indonesia, in 2018. To measure human development in some regions, the UNDP developed the HDI in 1990. In the 1990 Human Development Report, three basic dimensions of human development were defined in the HDI: income (economy), health, and education [35]. From these dimensions, four indicators of HDI data were derived: life expectancy rate, expected years of schooling, mean years of schooling, and adjusted per capita expenditure [35]. In this study, these four indicators are the response variables, $y_1, y_2, y_3,$ and $y_4,$ respectively.

The issue of the HDI has received considerable critical attention in every province/regency in Indonesia as it is one of the determining components of the general allocation fund (DAU) [36]. As one of the largest provinces in Indonesia, East Java Province contributes significantly to Indonesia’s HDI. East Java Province’s HDI in 2018 was 70.77 [37], slightly lower than the national achievement. In addition, it was the lowest among the five other provinces of Java island in the last three years. Therefore, further studies on East Java Province’s HDI are to be potentially discussed.

To determine whether the correlation matrix has an identity matrix (homogeneity of variances), we used Bartlett’s test of sphericity with statistical significance $\alpha = 0.05$. The results showed a statistical value of $\chi^2 = 121.993$ with p -value = 0.00, which means that the decision was to reject the null hypothesis as p -value < α . According to this result, we can infer that there is a significant correlation among the response variables; hence, multiresponse nonparametric regression analysis can be used.

Several studies have highlighted the potential factors associated with Indonesia’s HDI, such as the percentage of people living in poverty [25,38], the unemployment rate [25,38,39], population density [25,38], and the per capita gross regional domestic product (GRDP) [25,39,40]. Based on these, the predictor variables used in this study were population density and the percentage of people living in poverty. In addition, the relationship between the response variable and each predictor variable was identified by a partial scatterplot, as shown in Figure 2. The partial scatter plot between the four indicators of the HDI and population density (x) showed changes at a particular subinterval that is fit for the truncated spline function. Meanwhile, the partial scatterplot between the response variable and the percentage of people living in poverty (z) followed a pattern that repeated at a certain interval, with a particular trend; thus, this variable was approached by the Fourier series function.

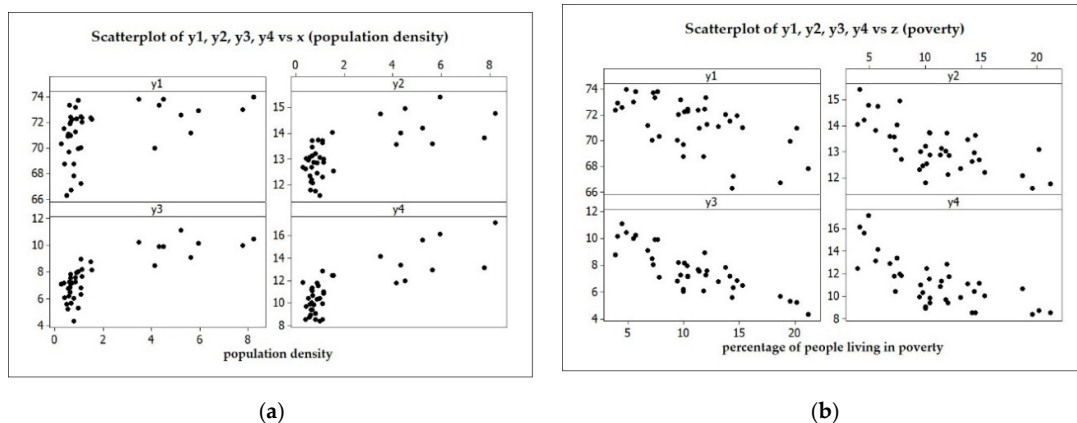


Figure 2. Partial scatterplot between the four indicators of the Human Development Index (HDI) with two predictor variables: (a) population density and (b) percentage of people living in poverty.

As shown in Table 2, some scenarios using a combined estimator (Model 1) and uncombined estimators (Model 2 and Model 3) were performed to compare the effectiveness of the proposed model. By using the modified GCV formula in Section 3.3, we selected the best model according to the minimum GCV as the criterion. In this study, a comparison of these three models revealed that the proposed model (Model 1) had the smallest GCV compared to the other uncombined estimator models for multiresponse nonparametric regression. In summary, these results indicate that the proposed model is better recommended for modeling the 2018 HDI data of East Java Province.

Table 2. Summary of statistical results for the case study.

Model 1 Combined Truncated Spline and Fourier Series Estimator for Multiresponse Nonparametric Regression				
	Number of Oscillations	Number of Knots	GCV	MSE
Model 1	1	1	1.43056	1.16177
		2	1.45686	1.14882
		3	1.41989	1.07056
	2	1	1.43464	1.11886
		2	1.46877	1.07960
		3	1.43883	1.07192
	3	1	2.10749	1.91785
		2	1.44040	1.13633
		3	1.39189 *	1.08377
Model 2 Truncated Spline Estimator for Multiresponse Nonparametric Regression				
	Number of Knots		GCV	MSE
Model 2	1		1.478036	1.114668
	2		1.522833	1.013464
	3		1.586664	1.055945
Model 3 Fourier Series Estimator for Multiresponse Nonparametric Regression				
	Number of Oscillations		GCV	MSE
Model 3	1		1.524463	1.149681
	2		1.536081	1.089298
	3		1.578724	1.05066

* The smallest value of GCV.

Another important result from Table 2 is the comparison of the number of knots and oscillation combinations from the proposed model. Similar to the simulation studies, we used a combination of three knot and oscillation parameters each. According to the minimum GCV as the criterion, the leading model was a three knots–three oscillations model with a GCV equal to 1.39189 and MSE = 1.08377. In addition, this model yields a coefficient of determination (R^2) of 99.84%. Due to the satisfactory result of R^2 , it can be said that the proposed model is able to describe the variance of the response variable through the predictor variables exceptionally well. Another interesting result was the number of knot and oscillation combinations of the best model in the simulation study, in line with those of data application.

Appendix H presents the results of parameter estimation from the best model obtained from the minimum GCV in Table 2. As such, we obtained a three knots–three oscillations model. Based on the parameter estimation results for each response variable for the 2018 HDI data of East Java Province, Indonesia, the multiresponse nonparametric regression model with a combined truncated spline and Fourier series estimator can be written as

$$\begin{aligned} \hat{y}_1 &= -0.932 + 0.869x_{1i} - 3057.409 (x_{1i} - 3.718)_+ + 6107.433 (x_{1i} - 3.933)_+ - 3050.878 (x_{1i} - 4.148)_+ - 0.677 z_{1i} + \\ &\quad 1/2 (1355.724) - 935.001 \cos z_{1i} + 423.443 \cos 2z_{1i} - 92.348 \cos 3z_{1i}. \\ &= 676.929 + 0.869x_{1i} - 3057.409 (x_{1i} - 3.718)_+ + 6107.433 (x_{1i} - 3.933)_+ - 3050.878 (x_{1i} - 4.148)_+ - 0.677 z_{1i} - \\ &\quad 935.001 \cos z_{1i} + 423.443 \cos 2z_{1i} - 92.348 \cos 3z_{1i}. \\ \hat{y}_2 &= -0.747 + 0.393x_{1i} - 619.153 (x_{1i} - 3.718)_+ + 1235.608 (x_{1i} - 3.933)_+ - 616.769 (x_{1i} - 4.148)_+ - 0.180 z_{1i} + \\ &\quad 1/2 (274.146) - 189.489 \cos z_{1i} + 85.224 \cos 2z_{1i} - 18.760 \cos 3z_{1i}. \\ &= 136.326 + 0.393x_{1i} - 619.153 (x_{1i} - 3.718)_+ + 1235.608 (x_{1i} - 3.933)_+ - 616.769 (x_{1i} - 4.148)_+ - 0.180 z_{1i} - \\ &\quad 189.489 \cos z_{1i} + 85.224 \cos 2z_{1i} - 18.760 \cos 3z_{1i}. \end{aligned}$$

$$\hat{y}_3 = -0.102 + 0.865x_{1i} - 510.292 (x_{1i} - 3.718)_+ + 1015.208 (x_{1i} - 3.933)_+ - 505.678 (x_{1i} - 4.148)_+ - 0.411 z_{1i} + 1/2 (224.955) - 156.440 \cos z_{1i} + 69.027 \cos 2z_{1i} - 15.588 \cos 3z_{1i}.$$

$$= 112.375 + 0.865x_{1i} - 510.292 (x_{1i} - 3.718)_+ + 1015.208 (x_{1i} - 3.933)_+ - 505.678 (x_{1i} - 4.148)_+ - 0.411 z_{1i} - 156.440 \cos z_{1i} + 69.027 \cos 2z_{1i} - 15.588 \cos 3z_{1i}.$$

$$\hat{y}_4 = -0.482 + 0.969x_{1i} - 684.712 (x_{1i} - 3.718)_+ + 1363.397 (x_{1i} - 3.933)_+ - 679.163 (x_{1i} - 4.148)_+ - 0.456 z_{1i} + 1/2 (302.146) - 209.894 \cos z_{1i} + 92.927 \cos 2z_{1i} - 20.891 \cos 3z_{1i}.$$

$$= 150.591 + 0.969x_{1i} - 684.712 (x_{1i} - 3.718)_+ + 1363.397 (x_{1i} - 3.933)_+ - 679.163 (x_{1i} - 4.148)_+ - 0.456 z_{1i} - 209.894 \cos z_{1i} + 92.927 \cos 2z_{1i} - 20.891 \cos 3z_{1i}.$$

Figure 3 presents a comparison between the response variable and the fitted values using the proposed model. From the graph, we can see that all fitted values resemble the pattern of the real data; although some do not, there are only a few deviations. Thus, the proposed model can produce a prediction model.

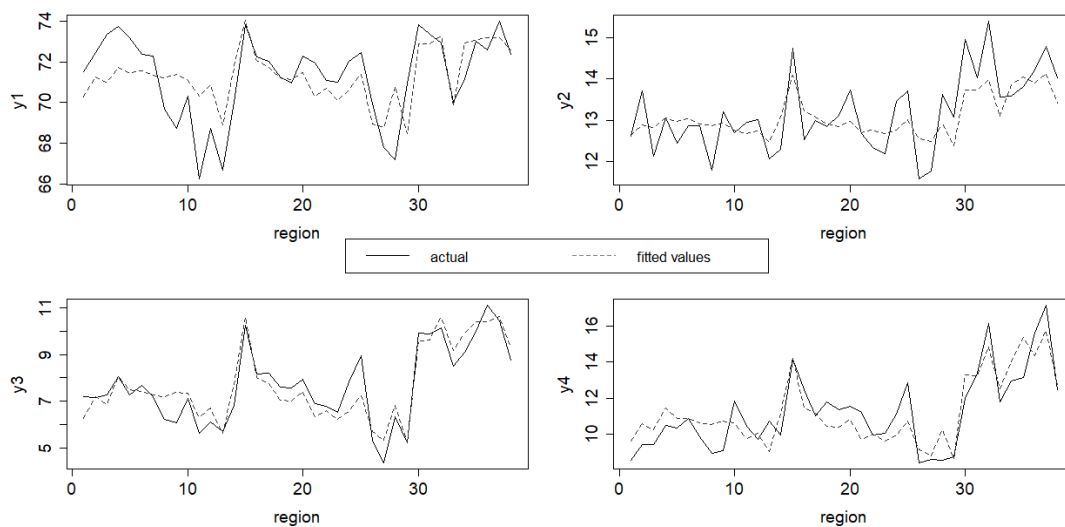


Figure 3. Comparison between actual and fitted values.

4. Discussion and Conclusions

This study presents the major findings from regression curve estimation with a combined truncated spline and Fourier series estimator for an additive model in multiresponse nonparametric regression using PWLS and WLS optimization, as shown below:

$$\hat{\mathbf{f}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) = \mathbf{J}[\mathbf{J}^T(\mathbf{I} - 2\mathbf{H}^T)\mathbf{W}\mathbf{J} + \mathbf{J}^T\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{J}]^{-1}[\mathbf{J}^T(\mathbf{H}^T - \mathbf{I})\mathbf{W}(\mathbf{H} - \mathbf{I})\mathbf{y}],$$

$$\hat{\mathbf{g}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) = \mathbf{Z}([\mathbf{Z}^T\mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1}\mathbf{Z}^T\mathbf{W}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})\mathbf{y}),$$

$$\hat{\boldsymbol{\mu}}_{(K,\lambda,T)}(\mathbf{x}, \mathbf{z}) = [\mathbf{J}\mathbf{K}^{-1}\mathbf{L} + \mathbf{H}(\mathbf{I} - \mathbf{J}\mathbf{K}^{-1}\mathbf{L})]\mathbf{y}$$

Furthermore, the result of error variance–covariance matrix estimation is as follows:

$$\hat{\mathbf{W}} = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \cdots & \hat{\sigma}_{1r} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} & \cdots & \hat{\sigma}_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{1r} & \hat{\sigma}_{2r} & \cdots & \hat{\sigma}_{rr} \end{pmatrix} \otimes \mathbf{I}$$

From the simulation study, the minimum GCV obtained a satisfactory outcome through a large sample with small variance. In addition, data application to a real dataset

of the 2018 HDI in East Java revealed that the proposed model had a better result than the uncombined model. An adequate coefficient of determination (R^2) from the best model indicates that the proposed model can explain the data variation remarkably well. Interestingly, the number of knots and oscillations of the best model in this simulation study and data application is consistent; that is a combination of three knots–three oscillations. In summary, these findings have significant implications for the understanding of regression curve estimation when using combined estimators for multiresponse nonparametric regression. Although the focus of the research was on combined truncated spline and Fourier series estimators, the procedure in this study can be applied to other combinations of estimators for multiresponse nonparametric regression.

The scope of this study was limited in terms of curve estimation and estimation of the variance–covariance matrix; thus, the major limitation was the absence of hypothesis testing and confidence intervals testing. Considering its necessity for good statistical practice, there is a commitment to perform such a goodness-of-fit test (model checking). Therefore, further research will be conducted to achieve hypothesis testing and confidence intervals testing. Another possible area is that a further simulation study could be performed using another function, such as an exponential or logarithmic function, to validate the capability of the proposed model. An additional variation of ρ in a simulation study could be used to gain more insights into the performance evaluation of the proposed model.

Author Contributions: Conceptualization, H.N., IN.B., and I.Z.; methodology, H.N. and IN.B.; writing—original draft preparation, H.N.; writing—review and editing, H.N., IN.B., and I.Z. All authors have read and agreed to the published version of the manuscript.

Funding: H.N. thanks BPS—Statistics Indonesia, which has bestowed a doctoral scholarship through APBN BPS. This research was funded by Deputy Bidang Penguatan Riset dan Pengembangan, Ministry of Research and Technology/National Research and Innovation Agency (Kemenristek or RISTEK-BRIN), the Republic of Indonesia, via grant Penelitian Disertasi Doktor (PDD) in 2021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: BPS—Statistics of Jawa Timur Province, Indonesia, <https://jatim.bps.go.id>, accessed on 1 March 2021.

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

Abbreviations

GoF	goodness-of-fit
GCV	generalized cross-validation
MSE	mean square error
MLE	maximum likelihood estimation
OLS	ordinary least square
PLS	penalized least square
PWLS	penalized weighted least square
WLS	weighted least square

Appendix A

Lemma 1 can be proven by the completed $R(g_{hk}(z_{k1}), g_{hk}(z_{k2}), \dots, g_{hk}(z_{kn}))$ as the GoF of PWLS optimization in Equation (7), as shown below:

$$R(g_{hk}(z_{k1}), g_{hk}(z_{k2}), \dots, g_{hk}(z_{kn})) = N^{-1} \sum_{h=1}^r \sum_{i=1}^n w_{hi} \left(y_{hi} - \sum_{j=1}^p f_{hj}(x_{ji}) - \sum_{k=1}^q g_{hk}(z_{ki}) \right)^2. \tag{A1}$$

If $v_{hi} = y_{hi} - \sum_{j=1}^p f_{hj}(x_{ji})$, then Equation (A1) can be written as

$$R(g_{hk}(z_{k1}), g_{hk}(z_{k2}), \dots, g_{hk}(z_{kn})) = N^{-1} \sum_{h=1}^r \sum_{i=1}^n w_{hi} \left(v_{hi} - \sum_{k=1}^q g_{hk}(z_{ki}) \right)^2, \tag{A2}$$

where $g_{hk}(z_{ki})$ is the Fourier series function, as shown in Equation (6). Furthermore, the function $g_{hk}(z_{ki})$ with one predictor variable, symbolized as k , can be written in the following matrix form:

$$\begin{aligned} \mathbf{g}_h &= \begin{bmatrix} g_{hk}(z_{k1}) \\ g_{hk}(z_{k2}) \\ \vdots \\ g_{hk}(z_{kn}) \end{bmatrix} = \begin{bmatrix} b_{hk}z_{k1} + \frac{1}{2}a_{0hk} + a_{1hk}\cos 1z_{k1} + a_{2hk}\cos 2z_{k1} + \dots + a_{Thk}\cos Tz_{k1} \\ b_{hk}z_{k2} + \frac{1}{2}a_{0hk} + a_{1hk}\cos 1z_{k2} + a_{2hk}\cos 2z_{k2} + \dots + a_{Thk}\cos Tz_{k2} \\ \vdots \\ b_{hk}z_{kn} + \frac{1}{2}a_{0hk} + a_{1hk}\cos 1z_{kn} + a_{2hk}\cos 2z_{kn} + \dots + a_{Thk}\cos Tz_{kn} \end{bmatrix} \\ &= \begin{bmatrix} z_{k1} & 1/2 & \cos 1z_{k1} & \cos 2z_{k1} & \dots & \cos Tz_{k1} \\ z_{k2} & 1/2 & \cos 1z_{k2} & \cos 2z_{k2} & \dots & \cos Tz_{k2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{kn} & 1/2 & \cos 1z_{kn} & \cos 2z_{kn} & \dots & \cos Tz_{kn} \end{bmatrix} \begin{bmatrix} b_{hk} \\ a_{0hk} \\ a_{1hk} \\ a_{2hk} \\ \vdots \\ a_{Thk} \end{bmatrix} \\ &= \mathbf{Z}_k \mathbf{a}_{hk} \end{aligned} \tag{A3}$$

Similarly, for $k = 1, 2, \dots, q$ predictors, the Fourier series function for the multiresponse nonparametric regression presented in Equation (A3) can be written as

$$\begin{aligned} \mathbf{g}_h &= \mathbf{Z}_1 \mathbf{a}_{h1} + \mathbf{Z}_2 \mathbf{a}_{h2} + \dots + \mathbf{Z}_q \mathbf{a}_{hq} \\ &= \sum_{k=1}^q \mathbf{Z}_k \mathbf{a}_{hk} = \mathbf{Z}_h \mathbf{a}_h, \end{aligned}$$

such that

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_r^T \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 \mathbf{a}_1 \\ \mathbf{Z}_2 \mathbf{a}_2 \\ \vdots \\ \mathbf{Z}_r \mathbf{a}_r \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{Z}_r \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_r \end{bmatrix} = \mathbf{Za} \tag{A4}$$

Thus, the GoF in Equation (7), where $g_{hk}(z_{ki})$ is the Fourier series function, can be expressed as

$$\begin{aligned} R(g_{hk}(z_{k1}), g_{hk}(z_{k2}), \dots, g_{hk}(z_{kn})) &= N^{-1} \sum_{h=1}^r \sum_{i=1}^n w_{hi} \left(v_{hi} - \sum_{k=1}^q g_{hk}(z_{ki}) \right)^2 \\ &= N^{-1} \sum_{h=1}^r \sum_{i=1}^n w_{hi} \left(v_{hi} - \sum_{k=1}^q \left(b_{hk}z_{ki} + \frac{1}{2}a_{0hk} + \sum_{t=1}^T a_{thk}\cos tz_{ki} \right) \right)^2. \end{aligned} \tag{A5}$$

As a result, the GoF component in Equation (A5) can be drawn in the following matrix form:

$$R(\mathbf{g}_1, \dots, \mathbf{g}_r) = N^{-1} \mathbf{W}(\mathbf{v} - \mathbf{Za})^2$$

$$= N^{-1}(\mathbf{v} - \mathbf{Za})^T \mathbf{W}(\mathbf{v} - \mathbf{Za}).$$

Appendix B

The penalty component of the PWLS optimization in Equation (7) is as follows:

$$\sum_{k=1}^q \lambda_k \int_0^\pi \frac{2}{\pi} (g_k''(z_k))^2 dz_k, \tag{A6}$$

where $g_{hk}(z_{ki})$ is presented in Equation (6). To prove Lemma 2, let us begin with solving the second derivative in Equation (A6), as follows:

$$\begin{aligned} \int_0^\pi \frac{2}{\pi} (g_k''(z_k))^2 dz_k &= \int_0^\pi \frac{2}{\pi} \left(\frac{d}{dz_k} \left[\frac{d}{dz_k} \left(b_{hk} z_{ki} + \frac{1}{2} a_{0hk} + \sum_{t=1}^T a_{thk} \cos tz_{ki} \right) \right] \right)^2 dz_k \\ &= \frac{2}{\pi} \int_0^\pi \left(\sum_{t=1}^T t^2 a_{thk} \cos tz_{ki} \right)^2 dz_k \\ &= \frac{2}{\pi} \int_0^\pi \left\{ \sum_{t=1}^T (t^2 a_{thk} \cos tz_{ki})^2 + 2 \sum_{t < u} (t^2 a_{thk} \cos tz_{ki})(u^2 a_{thk} \cos uz_{ki}) \right\} dz_k \\ &= \frac{2}{\pi} \sum_{t=1}^T \int_0^\pi (t^4 a_{thk}^2 \cos^2 tz_{ki}) dz_k + \frac{4}{\pi} \int_0^\pi \sum_{t < u} ((tu)^2 a_{thk} \cos tz_{ki} a_{uhk} \cos uz_{ki}) dz_k \\ &= \sum_{t=1}^T t^4 a_{thk}^2. \tag{A7} \end{aligned}$$

After obtaining the result in Equation (A7), the solution to the penalty component in Equation (A6) can be written as

$$\begin{aligned} P(\lambda_1, \lambda_2, \dots, \lambda_q) &= \sum_{k=1}^q \lambda_k \int_0^\pi \frac{2}{\pi} (g_k''(z_k))^2 dz_k \\ &= \sum_{k=1}^q \left(\lambda_k \sum_{t=1}^T t^4 a_{thk}^2 \right) \\ &= \mathbf{a}^T \mathbf{D}(\lambda) \mathbf{a}. \end{aligned}$$

Appendix C

According to the result of the GoF in Lemma 1 and the penalty in Lemma 2, the PWLS optimization in Equation (7) can be written as

$$\text{Min}_{g_k \in C[0, \pi]} \left\{ N^{-1} \sum_{h=1}^r \sum_{i=1}^n w_{hi} \left(y_{hi} - \sum_{j=1}^p f_{hj}(x_{ji}) - \sum_{k=1}^q g_{hk}(z_{ki}) \right)^2 + \sum_{k=1}^q \lambda_k \int_0^\pi \frac{2}{\pi} (g_k''(z_k))^2 dz_k \right\}$$

$$= \text{Min}_{\mathbf{a} \in R^{(2+T)q \times 1}} \{ N^{-1}(\mathbf{v} - \mathbf{Za})^T \mathbf{W}(\mathbf{v} - \mathbf{Za}) + \mathbf{a}^T \mathbf{D}(\lambda) \mathbf{a} \}$$

If $\mathbf{F}(\mathbf{a}) = N^{-1}(\mathbf{v} - \mathbf{Za})^T \mathbf{W}(\mathbf{v} - \mathbf{Za}) + \mathbf{a}^T \mathbf{D}(\lambda) \mathbf{a}$, then

$$\mathbf{F}(\mathbf{a}) = N^{-1} \mathbf{v}^T \mathbf{W} \mathbf{v} - 2N^{-1} \mathbf{a}^T \mathbf{Z}^T \mathbf{W} \mathbf{v} + N^{-1} \mathbf{a}^T \mathbf{Z}^T \mathbf{W} \mathbf{Z} \mathbf{a} + \mathbf{a}^T \mathbf{D}(\lambda) \mathbf{a},$$

such that

$$\begin{aligned} & \underset{\mathbf{a} \in \mathbb{R}^{(2+T)q \times 1}}{\text{Min}} \{N^{-1}(\mathbf{v} - \mathbf{Za})^T \mathbf{W}(\mathbf{v} - \mathbf{Za}) + \mathbf{a}^T \mathbf{D}(\lambda)\mathbf{a}\} \\ = & \underset{\mathbf{a} \in \mathbb{R}^{(2+T)q \times 1}}{\text{Min}} \{N^{-1}\mathbf{v}^T \mathbf{W}\mathbf{v} - 2N^{-1}\mathbf{a}^T \mathbf{Z}^T \mathbf{W}\mathbf{v} + N^{-1}\mathbf{a}^T \mathbf{Z}^T \mathbf{W}\mathbf{Za} + \mathbf{a}^T \mathbf{D}(\lambda)\mathbf{a}\}. \end{aligned} \tag{A8}$$

To solve the optimization in Equation (A8), the estimator $\hat{\mathbf{a}}$ can be obtained by deriving the partial Equation (A8) against \mathbf{a} and equating the result with $\mathbf{0}$, as follows:

$$\frac{\partial \mathbf{F}(\mathbf{a})}{\partial \mathbf{a}} = \frac{\partial [N^{-1}\mathbf{v}^T \mathbf{W}\mathbf{v} - 2N^{-1}\mathbf{a}^T \mathbf{Z}^T \mathbf{W}\mathbf{v} + N^{-1}\mathbf{a}^T \mathbf{Z}^T \mathbf{W}\mathbf{Za} + \mathbf{a}^T \mathbf{D}(\lambda)\mathbf{a}]}{\partial \mathbf{a}}$$

$$\hat{\mathbf{a}} = [\mathbf{Z}^T \mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1} \mathbf{Z}^T \mathbf{W}\mathbf{v}. \tag{A9}$$

The subsequent step is to substitute $\hat{\mathbf{a}}$ in Equation (A9) into Equation (A4), such that the estimator of the truncated spline function can be written as

$$\hat{\mathbf{g}}_{(\kappa, \lambda, T)}(\mathbf{x}, \mathbf{z}) = \mathbf{Z}\hat{\mathbf{a}} \tag{A10}$$

$$= (\mathbf{Z}[\mathbf{Z}^T \mathbf{W}\mathbf{Z} + \mathbf{N}\mathbf{D}(\lambda)]^{-1} \mathbf{Z}^T \mathbf{W})\mathbf{v}$$

$$= \mathbf{H}\mathbf{v}. \tag{A11}$$

Appendix D

The solution in Theorem 1 still contains an \mathbf{f} component, where \mathbf{f} is a truncated linear spline function, as described in Equation (5). Hence, to complete the PWLS optimization in Equation (7), it is necessary to estimate the regression curve of $f_{hj}(x_{ji})$. First, the multiresponse nonparametric regression curve of the truncated spline component in Equation (5), which involves only one predictor, is written in the following matrix form:

$$\begin{aligned} \mathbf{f}_h &= \begin{bmatrix} f_{hj}(x_{j1}) \\ f_{hj}(x_{j2}) \\ \vdots \\ f_{hj}(x_{jn}) \end{bmatrix} = \begin{bmatrix} \alpha_{hj}x_{j1} + \sum_{s=1}^u \beta_{hjs}(x_{j1} - K_{hjs})_+ \\ \alpha_{hj}x_{j2} + \sum_{s=1}^u \beta_{hjs}(x_{j2} - K_{hjs})_+ \\ \vdots \\ \alpha_{hj}x_{jn} + \sum_{s=1}^u \beta_{hjs}(x_{jn} - K_{hjs})_+ \end{bmatrix} \\ &= \begin{bmatrix} x_{j1} \\ x_{j2} \\ \vdots \\ x_{jn} \end{bmatrix} \alpha_{hj} + \begin{bmatrix} (x_{j1} - K_{hj1})_+ & (x_{j1} - K_{hj2})_+ & \cdots & (x_{j1} - K_{hju})_+ \\ (x_{j2} - K_{hj1})_+ & (x_{j2} - K_{hj2})_+ & \cdots & (x_{j2} - K_{hju})_+ \\ \vdots & \vdots & \ddots & \vdots \\ (x_{jn} - K_{hj1})_+ & (x_{jn} - K_{hj2})_+ & \cdots & (x_{jn} - K_{hju})_+ \end{bmatrix} \begin{bmatrix} \beta_{hj1} \\ \beta_{hj2} \\ \vdots \\ \beta_{hju} \end{bmatrix} \\ &= \mathbf{x}_{ji}\alpha_{hj} + \mathbf{S}_{ji}\boldsymbol{\beta}_{hj} \end{aligned} \tag{A12}$$

Consequently, \mathbf{f}_h with the $j = 1, 2, \dots, p$ predictor can be written as follows:

$$\begin{aligned} \mathbf{f}_h &= \mathbf{x}_{1i}\alpha_{1h} + \mathbf{S}_{1i}\boldsymbol{\beta}_{1h} + \mathbf{x}_{2i}\alpha_{2h} + \mathbf{S}_{2i}\boldsymbol{\beta}_{2h} + \cdots + \mathbf{x}_{pi}\alpha_{ph} + \mathbf{S}_{pi}\boldsymbol{\beta}_{ph} \\ &= \mathbf{X}_i\boldsymbol{\alpha}_h + \mathbf{S}_i\boldsymbol{\beta}_h, \end{aligned} \tag{A13}$$

According to Equation (A13), the truncated spline function for multiresponse nonparametric regression can be expressed in matrix form, as follows:

$$\begin{aligned}
 \mathbf{f} &= \begin{bmatrix} \mathbf{f}_1^T \\ \mathbf{f}_2^T \\ \vdots \\ \mathbf{f}_r^T \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \boldsymbol{\alpha}_1 \\ \mathbf{X}_2 \boldsymbol{\alpha}_2 \\ \vdots \\ \mathbf{X}_r \boldsymbol{\alpha}_r \end{bmatrix} + \begin{bmatrix} \mathbf{S}_1 \boldsymbol{\beta}_1 \\ \mathbf{S}_2 \boldsymbol{\beta}_2 \\ \vdots \\ \mathbf{S}_r \boldsymbol{\beta}_r \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{X}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{X}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_r \end{bmatrix} + \begin{bmatrix} \mathbf{S}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{S}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{S}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_r \end{bmatrix} \\
 &= \begin{bmatrix} \mathbf{X}_1 & 0 & \cdots & 0 & \mathbf{S}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{X}_2 & \cdots & 0 & 0 & \mathbf{S}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{X}_r & 0 & 0 & \cdots & \mathbf{S}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_r \\ \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_r \end{bmatrix} \\
 &= [\mathbf{X}|\mathbf{S}] \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix} = \mathbf{J}\boldsymbol{\gamma}. \tag{A14}
 \end{aligned}$$

To solve the regression curve estimation of the truncated spline and Fourier series estimators for multiresponse nonparametric regression with PWLS optimization, substitute Equations (A11) and (A14) into Equation (4), as follows:

$$\begin{aligned}
 \mathbf{y} &= \mathbf{f} + \mathbf{g} + \boldsymbol{\varepsilon}; \quad \mathbf{g} = \mathbf{H}\mathbf{v} \\
 \boldsymbol{\varepsilon} &= \mathbf{y} - \mathbf{f} - \mathbf{H}\mathbf{v}; \quad \mathbf{v} = \mathbf{y} - \mathbf{f} \\
 &= \mathbf{y} - \mathbf{f} - \mathbf{H}(\mathbf{y} - \mathbf{f}); \quad \mathbf{f} = \mathbf{J}\boldsymbol{\gamma} \\
 &= (\mathbf{I} - \mathbf{H})(\mathbf{y} - \mathbf{J}\boldsymbol{\gamma}). \tag{A15}
 \end{aligned}$$

The subsequent step is to solve Equation (A15) with WLS optimization. Therefore, the estimator can be obtained by completing the following WLS:

$$\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = [(\mathbf{I} - \mathbf{H})(\mathbf{y} - \mathbf{J}\boldsymbol{\gamma})]^T \mathbf{W} [(\mathbf{I} - \mathbf{H})(\mathbf{y} - \mathbf{J}\boldsymbol{\gamma})]$$

Appendix E

The second stage of the proposed method is performed using WLS optimization, as follows:

$$\text{Min}_{\boldsymbol{\gamma} \in R^{(1+u)nr}} \{[(\mathbf{I} - \mathbf{H})(\mathbf{y} - \mathbf{J}\boldsymbol{\gamma})]^T \mathbf{W} [(\mathbf{I} - \mathbf{H})(\mathbf{y} - \mathbf{J}\boldsymbol{\gamma})]\}$$

If $\mathbf{T}(\boldsymbol{\gamma}) = [\mathbf{y} - \mathbf{H}\mathbf{y} - \mathbf{J}\boldsymbol{\gamma} + \mathbf{H}\mathbf{J}\boldsymbol{\gamma}]^T \mathbf{W} [\mathbf{y} - \mathbf{H}\mathbf{y} - \mathbf{J}\boldsymbol{\gamma} + \mathbf{H}\mathbf{J}\boldsymbol{\gamma}]$, then we perform the multiplication in parentheses to obtain,

$$\begin{aligned}
 \mathbf{T}(\boldsymbol{\gamma}) &= \mathbf{y}^T \mathbf{W} \mathbf{y} + \mathbf{y}^T \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{y} + \boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{W} \mathbf{J} \boldsymbol{\gamma} + \boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{J} \boldsymbol{\gamma} - 2\mathbf{y}^T \mathbf{H}^T \mathbf{W} \mathbf{y} - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{W} \mathbf{y} \\
 &\quad + 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{y} + 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{W} \mathbf{H} \mathbf{y} - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{y} - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{J} \boldsymbol{\gamma}.
 \end{aligned}$$

WLS optimization can be obtained by partial derivation of $\mathbf{T}(\boldsymbol{\gamma})$ against $\boldsymbol{\gamma}$. When the expression is equal to 0, the result is Equation (A16):

$$\begin{aligned}
 \frac{\partial \mathbf{T}(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} &= \frac{\partial}{\partial \boldsymbol{\gamma}} (\mathbf{y}^T \mathbf{W} \mathbf{y} + \mathbf{y}^T \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{y} + \boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{W} \mathbf{J} \boldsymbol{\gamma} + \boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{J} \boldsymbol{\gamma} - 2\mathbf{y}^T \mathbf{H}^T \mathbf{W} \mathbf{y} \\
 &\quad - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{W} \mathbf{y} + 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{y} + 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{W} \mathbf{H} \mathbf{y} - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{y} - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{J} \boldsymbol{\gamma}) = 0,
 \end{aligned}$$

$$\hat{\boldsymbol{\gamma}} = [\mathbf{J}^T (\mathbf{I} - 2\mathbf{H}^T) \mathbf{W} \mathbf{J} + \mathbf{J}^T \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{J}]^{-1} [\mathbf{J}^T (\mathbf{H}^T - \mathbf{I}) \mathbf{W} (\mathbf{H} - \mathbf{I}) \mathbf{y}]. \tag{A16}$$

Substituting Equation (A16) into Equation (A14) yields

$$\hat{\mathbf{f}}_{(\kappa, \lambda, \mathbf{T})}(\mathbf{x}, \mathbf{z}) = \mathbf{J}\hat{\boldsymbol{\gamma}} \tag{A17}$$

$$\begin{aligned}
 &= \mathbf{J}[\mathbf{J}^T(\mathbf{I} - 2\mathbf{H}^T)\mathbf{W}\mathbf{J} + \mathbf{J}^T\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{J}]^{-1}[\mathbf{J}^T(\mathbf{H}^T - \mathbf{I})\mathbf{W}(\mathbf{H} - \mathbf{I})]\mathbf{y} \\
 &= \mathbf{J}\mathbf{K}^{-1}\mathbf{L}\mathbf{y}.
 \end{aligned}
 \tag{A18}$$

Appendix F

Assume random error $\boldsymbol{\varepsilon}$ in Equation (2) is normally distributed with mean $\mathbf{0}$ and error variance–covariance matrix \mathbf{W} such that it can be written as $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{W})$. Thus, matrix \mathbf{W} is outlined as follows

$$\begin{aligned}
 \mathbf{W} &= \text{Var}(\boldsymbol{\varepsilon}) = E\left[(\boldsymbol{\varepsilon} - E(\boldsymbol{\varepsilon}))(\boldsymbol{\varepsilon} - E(\boldsymbol{\varepsilon}))^T\right] \\
 &= E\left[(\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n}, \dots, \varepsilon_{r1}, \varepsilon_{r2}, \dots, \varepsilon_{rn})^T (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2n}, \dots, \varepsilon_{r1}, \varepsilon_{r2}, \dots, \varepsilon_{rn})\right] \\
 &= E\left[\begin{pmatrix} \varepsilon_{11}^2 & \varepsilon_{11}\varepsilon_{12} & \dots & \varepsilon_{11}\varepsilon_{1n} & \varepsilon_{11}\varepsilon_{21} & \varepsilon_{11}\varepsilon_{22} & \dots & \varepsilon_{11}\varepsilon_{2n} & \dots & \varepsilon_{11}\varepsilon_{r1} & \varepsilon_{11}\varepsilon_{r2} & \dots & \varepsilon_{11}\varepsilon_{rn} \\ \varepsilon_{12}\varepsilon_{11} & \varepsilon_{12}^2 & \dots & \varepsilon_{12}\varepsilon_{1n} & \varepsilon_{12}\varepsilon_{21} & \varepsilon_{12}\varepsilon_{22} & \dots & \varepsilon_{12}\varepsilon_{2n} & \dots & \varepsilon_{12}\varepsilon_{r1} & \varepsilon_{12}\varepsilon_{r2} & \dots & \varepsilon_{12}\varepsilon_{rn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{1n}\varepsilon_{11} & \varepsilon_{1n}\varepsilon_{12} & \dots & \varepsilon_{1n}^2 & \varepsilon_{1n}\varepsilon_{21} & \varepsilon_{1n}\varepsilon_{22} & \dots & \varepsilon_{1n}\varepsilon_{2n} & \dots & \varepsilon_{1n}\varepsilon_{r1} & \varepsilon_{1n}\varepsilon_{r2} & \dots & \varepsilon_{1n}\varepsilon_{rn} \\ \varepsilon_{21}\varepsilon_{11} & \varepsilon_{21}\varepsilon_{12} & \dots & \varepsilon_{21}\varepsilon_{1n} & \varepsilon_{21}^2 & \varepsilon_{21}\varepsilon_{22} & \dots & \varepsilon_{21}\varepsilon_{2n} & \dots & \varepsilon_{21}\varepsilon_{r1} & \varepsilon_{21}\varepsilon_{r2} & \dots & \varepsilon_{21}\varepsilon_{rn} \\ \varepsilon_{22}\varepsilon_{11} & \varepsilon_{22}\varepsilon_{12} & \dots & \varepsilon_{22}\varepsilon_{1n} & \varepsilon_{22}\varepsilon_{21} & \varepsilon_{22}^2 & \dots & \varepsilon_{22}\varepsilon_{2n} & \dots & \varepsilon_{22}\varepsilon_{r1} & \varepsilon_{22}\varepsilon_{r2} & \dots & \varepsilon_{22}\varepsilon_{rn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{2n}\varepsilon_{11} & \varepsilon_{2n}\varepsilon_{12} & \dots & \varepsilon_{2n}\varepsilon_{1n} & \varepsilon_{2n}\varepsilon_{21} & \varepsilon_{2n}\varepsilon_{22} & \dots & \varepsilon_{2n}^2 & \dots & \varepsilon_{2n}\varepsilon_{r1} & \varepsilon_{2n}\varepsilon_{r2} & \dots & \varepsilon_{2n}\varepsilon_{rn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{r1}\varepsilon_{11} & \varepsilon_{r1}\varepsilon_{12} & \dots & \varepsilon_{r1}\varepsilon_{1n} & \varepsilon_{r1}\varepsilon_{21} & \varepsilon_{r1}\varepsilon_{22} & \dots & \varepsilon_{r1}\varepsilon_{2n} & \dots & \varepsilon_{r1}^2 & \varepsilon_{r1}\varepsilon_{r2} & \dots & \varepsilon_{r1}\varepsilon_{rn} \\ \varepsilon_{r2}\varepsilon_{11} & \varepsilon_{r2}\varepsilon_{12} & \dots & \varepsilon_{r2}\varepsilon_{1n} & \varepsilon_{r2}\varepsilon_{21} & \varepsilon_{r2}\varepsilon_{22} & \dots & \varepsilon_{r2}\varepsilon_{2n} & \dots & \varepsilon_{r2}\varepsilon_{r1} & \varepsilon_{r2}^2 & \dots & \varepsilon_{r2}\varepsilon_{rn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{rn}\varepsilon_{11} & \varepsilon_{rn}\varepsilon_{12} & \dots & \varepsilon_{rn}\varepsilon_{1n} & \varepsilon_{rn}\varepsilon_{21} & \varepsilon_{rn}\varepsilon_{22} & \dots & \varepsilon_{rn}\varepsilon_{2n} & \dots & \varepsilon_{rn}\varepsilon_{r1} & \varepsilon_{rn}\varepsilon_{r2} & \dots & \varepsilon_{rn}^2 \end{pmatrix} \right].
 \end{aligned}
 \tag{A19}$$

The correlation between response variables, define as $\text{corr}(\varepsilon_{hi}, \varepsilon_{\ell i})$, is assumed to be constant as ρ for all response variables and succeeded by the other assumptions, as follows:

$$\begin{aligned}
 E(\varepsilon_{hi}\varepsilon_{h' i'}) &= \begin{cases} \sigma_{hh}, & \text{if } i = i' \\ 0, & \text{if } i \neq i' \end{cases} \text{ where } h = 1, 2, \dots, r \\
 E(\varepsilon_{hi}\varepsilon_{\ell i'}) &= \begin{cases} \sigma_{h\ell}, & \text{if } i = i' \\ 0, & \text{if } i \neq i' \end{cases} \text{ with } \sigma_{h\ell} = \rho\sqrt{\sigma_{hh}\sigma_{\ell\ell}}
 \end{aligned}$$

According to the above assumption, Equation (A19) can be written as

$$\begin{aligned}
 \mathbf{W} &= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1r} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1r} & \sigma_{2r} & \dots & \sigma_{rr} \end{pmatrix} \otimes \mathbf{I} \\
 &= \boldsymbol{\Sigma} \otimes \mathbf{I}
 \end{aligned}
 \tag{A20}$$

The estimation of matrix \mathbf{W} is obtained by the maximum likelihood estimation (MLE) method, as in Equation (A20). Furthermore, substitute Equation (A10) and Equation (A17) into $\boldsymbol{\mu}(\mathbf{x}, \mathbf{z})$; then, the likelihood presented in Equation (A22) is as follows:

$$L(\mathbf{W}|\boldsymbol{\mu}, \mathbf{y}) = (2\pi)^{-rn/2} |\mathbf{W}|^{-n/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n \left((\mathbf{y} - \boldsymbol{\mu}(\mathbf{x}, \mathbf{z}))^T \mathbf{W}^{-1} (\mathbf{y} - \boldsymbol{\mu}(\mathbf{x}, \mathbf{z})) \right)\right]
 \tag{A21}$$

$$= (2\pi)^{-rn/2} |\mathbf{W}|^{-n/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n \left((\mathbf{y} - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a}))^T \mathbf{W}^{-1} (\mathbf{y} - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})) \right)\right].
 \tag{A22}$$

By taking the natural logarithm from the likelihood function and substituting \mathbf{W} , as in Equation (A20), the result is as follows:

$$\ln L(\Sigma) = -\frac{rn}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma \otimes \mathbf{I}| - \frac{1}{2} \sum_{i=1}^n \left[(\mathbf{y} - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a}))^T (\Sigma \otimes \mathbf{I})^{-1} (\mathbf{y} - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})) \right] \tag{A23}$$

Point $(\mathbf{y} - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a}))$ into the form of $vec(\mathbf{C})$

$$(\mathbf{y} - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a}))^T (\Sigma \otimes \mathbf{I})^{-1} (\mathbf{y} - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})) = (vec(\mathbf{C}))^T (\Sigma \otimes \mathbf{I})^{-1} (vec(\mathbf{C})) = tr(\Sigma^{-1} \mathbf{C}^T \mathbf{C})$$

such that Equation (A23) can be written as

$$\ln L(\Sigma|\boldsymbol{\gamma}, \mathbf{a}, \mathbf{y}) = -\frac{rn}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| - \frac{1}{2} tr(\Sigma^{-1} \mathbf{C}^T \mathbf{C}) \tag{A24}$$

The maximum value of the likelihood function is obtained by partially deriving Equation (A24) against $\sigma_{h\ell}$ and equating the result with 0, as follows:

$$\frac{\partial}{\partial \sigma_{h\ell}} \ln L(\Sigma|\boldsymbol{\gamma}, \mathbf{a}, \mathbf{y}) = -\frac{n}{2} \left(\frac{\partial}{\partial \sigma_{h\ell}} \ln|\Sigma| \right) - \frac{1}{2} \frac{\partial}{\partial \sigma_{h\ell}} tr(\Sigma^{-1} \mathbf{C}^T \mathbf{C}) = 0$$

such that

$$\hat{\Sigma} = \frac{1}{n} (\mathbf{C}^T \mathbf{C})$$

$$= \begin{pmatrix} \frac{[\mathbf{y}_1 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_1 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} & \frac{[\mathbf{y}_1 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_2 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} & \dots & \frac{[\mathbf{y}_1 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_r - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} \\ \frac{[\mathbf{y}_1 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_2 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} & \frac{[\mathbf{y}_2 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_2 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} & \dots & \frac{[\mathbf{y}_2 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_r - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{[\mathbf{y}_1 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_r - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} & \frac{[\mathbf{y}_2 - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_r - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} & \dots & \frac{[\mathbf{y}_r - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]^T [\mathbf{y}_r - (\mathbf{J}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{a})]}{n} \end{pmatrix} \tag{A25}$$

The subsequent step is to estimate $\boldsymbol{\gamma}$ and \mathbf{a} using the ordinary least square (OLS) method, with the result as follows:

$$\frac{\partial \mathbb{Q}}{\partial \boldsymbol{\gamma}} = \frac{\partial [\mathbf{y}^T \mathbf{y} - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{y} - 2\mathbf{a}^T \mathbf{Z}^T \mathbf{y} + 2\mathbf{a}^T \mathbf{Z}^T \mathbf{J} \boldsymbol{\gamma} + \boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{J} \boldsymbol{\gamma} + \mathbf{a}^T \mathbf{Z}^T \mathbf{Z} \mathbf{a}]}{\partial \boldsymbol{\gamma}}$$

$$\hat{\boldsymbol{\gamma}}_{\ell s} = (([\mathbf{I} - (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{J}]^{-1} (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T) - ((\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T)) \mathbf{y} = \mathbf{F} \mathbf{y} \tag{A26}$$

$$\frac{\partial \mathbb{Q}}{\partial \mathbf{a}} = \frac{\partial [\mathbf{y}^T \mathbf{y} - 2\boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{y} - 2\mathbf{a}^T \mathbf{Z}^T \mathbf{y} + 2\mathbf{a}^T \mathbf{Z}^T \mathbf{J} \boldsymbol{\gamma} + \boldsymbol{\gamma}^T \mathbf{J}^T \mathbf{J} \boldsymbol{\gamma} + \mathbf{a}^T \mathbf{Z}^T \mathbf{Z} \mathbf{a}]}{\partial \mathbf{a}}$$

$$\hat{\mathbf{a}}_{\ell s} = (([\mathbf{I} - (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{J} (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{Z}]^{-1} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T) - ((\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{J} (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T)) \mathbf{y} = \mathbf{G} \mathbf{y} \tag{A27}$$

Substituting the result of $\hat{\boldsymbol{\gamma}}_{\ell s}$ and $\hat{\mathbf{a}}_{\ell s}$ above into Equation (A25), error variance-covariance matrix \mathbf{W} in Equation (A20) can be written as follows:

$$\hat{\mathbf{W}} = \hat{\Sigma} \otimes \mathbf{I},$$

where

$$\hat{\sigma}_{11} = \frac{[\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]^T [\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]}{n}, \hat{\sigma}_{22} = \frac{[\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]^T [\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]}{n}, \hat{\sigma}_{rr} = \frac{[\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]^T [\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]}{n},$$

$$\hat{\sigma}_{12} = \frac{[\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]^T [\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]}{n}, \hat{\sigma}_{1r} = \frac{[\mathbf{y}_1 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_1]^T [\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]}{n}, \hat{\sigma}_{2r} = \frac{[\mathbf{y}_2 - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_2]^T [\mathbf{y}_r - (\mathbf{J}\mathbf{F} + \mathbf{Z}\mathbf{G})\mathbf{y}_r]}{n}$$

Appendix G

Table A1. Summary of statistical results for $n = 20$ and 50 with error variances, the number of knots, and oscillation parameters

Variance	Number of Oscillations	Number of Knots	n = 20			n = 50		
			GCV	R ²	MSE	GCV	R ²	MSE
0.1	1	1	7.023	88.665	4.527	7.207	90.170	6.735
		2	6.859	89.866	4.080	6.848	90.607	6.312
		3	6.470	91.672	3.336	6.685 *	91.172	6.075
	2	1	6.777	87.649	5.036	7.074	91.133	5.960
		2	6.386 *	88.288	4.712	6.972	91.609	5.638
		3	6.388	89.881	4.085	7.165	91.835	5.489
	3	1	6.535	86.398	5.294	7.074	91.133	5.959
		2	6.601	88.112	5.151	6.972	91.610	5.638
		3	6.617	87.929	4.970	7.166	91.833	5.489
0.5	1	1	9.185	93.280	5.667	9.217	90.952	7.783
		2	10.010	93.303	5.655	7.141	93.293	5.769
		3	10.139	93.824	5.219	7.265	93.384	5.692
	2	1	7.970	92.470	6.455	9.215	90.933	7.797
		2	8.193	92.538	6.393	7.133	93.284	5.775
		3	8.012	92.896	6.018	7.293	93.436	5.645
	3	1	10.772	92.210	6.599	9.215	90.930	7.799
		2	9.263	93.328	5.655	7.132	93.283	5.776
		3	8.396	92.557	6.306	7.293	93.433	5.647
1.0	1	1	9.256	92.935	5.719	7.559	94.734	7.064
		2	10.136	93.218	5.491	7.319	94.972	6.746
		3	10.241	93.751	5.058	6.890	95.332	6.262
	2	1	8.019	91.988	6.495	7.585	95.227	6.401
		2	7.838	92.518	6.116	7.461	95.508	6.024
		3	7.330	93.445	5.506	7.380	95.753	5.698
	3	1	10.593	91.666	6.763	7.490	94.862	6.903
		2	7.668	92.629	5.983	7.175	95.155	6.521
		3	8.697	93.833	5.013	6.968	95.351	6.245

* The smallest value of GCV.

Appendix H

Table A2. Results of parameter estimation of the best model.

Response variable	Parameter	Estimation	Response Variable	Parameter	Estimation
y_1	α_{01}	-0.932	y_3	α_{03}	-0.102
	α_{11}	0.869		α_{31}	0.865
	β_{111}	-3057.409		β_{311}	-510.292
	β_{112}	6107.433		β_{312}	1015.208
	β_{113}	-3050.878		β_{313}	-505.678
	b_{11}	-0.677		b_{31}	-0.411
	a_{011}	1355.724		a_{031}	224.955
	a_{111}	-935.001		a_{311}	-156.440
	a_{112}	423.443		a_{312}	69.027
a_{113}	-92.348	a_{313}	-15.588		
y_2	α_{02}	-0.747	y_4	α_{04}	-0.482
	α_{21}	0.393		α_{41}	0.969
	β_{211}	-619.153		β_{411}	-684.712

β_{212}	1235.608	β_{412}	1363.397
β_{213}	-616.769	β_{413}	-679.163
b_{21}	-0.180	b_{41}	-0.456
a_{021}	274.146	a_{041}	302.146
a_{211}	-189.489	a_{411}	-209.894
a_{212}	85.224	a_{412}	92.927
a_{213}	-18.760	a_{413}	-20.891

References

- Härdle, W. *Applied Nonparametric Regression*; Cambridge University Press: Cambridge, UK, 1990.
- Eubank, R.L. *Nonparametric Regression and Spline Smoothing*, 2nd ed; Marcel Dekker, Inc.: New York, NY, USA, 1999.
- Budakçı, G.; Oruç, H. Further Properties of Quantum Spline Spaces. *Mathematics* **2020**, *8*, 1770.
- Du, R.; Yamada, H. Principle of Duality in Cubic Smoothing Spline. *Mathematics* **2020**, *8*, 1839.
- Lapshin, V. A nonparametric approach to bond portfolio immunization. *Mathematics* **2019**, *7*, 1121.
- Nurcahayani, H.; Budiantara, I.N.; Zain, I. Nonparametric Truncated Spline Regression on Modelling Mean Years Schooling of Regencies in Java. *AIP Conf. Proc.* **2019**, *2194*, 020073, doi:10.1063/1.5139805.
- Yu, W.; Yong, Y.; Guan, G.; Huang, Y.; Su, W.; Cui, C. Valuing guaranteed minimum death benefits by cosine series expansion. *Mathematics* **2019**, *7*, 1–15.
- Kim, J.; Hart, J.D. A Change-Point Estimator Using Local Fourier series. *J. Nonparametr. Stat.* **2011**, *23*, 83–98.
- Bilodeau, M. Fourier Smoother and Additive Models. *Can. J. Stat.* **1992**, *20*, 257–269.
- Yang, Y.; Pilanci, M.; Wainwright, M.J. Randomized Sketches for Kernels: Fast and Optimal Nonparametric Regression. *Ann. Stat.* **2017**, *45*, 991–1023.
- Zhao, G.; Ma, Y. Robust Nonparametric Kernel Regression Estimator. *Stat. Probab. Lett.* **2016**, *116*, 72–79.
- Kayri, M.; Zirhlioğlu, G. Kernel Smoothing Function and Choosing Bandwidth for Non-Parametric Regression Methods. *Ozean J. Appl. Sci.* **2009**, *2*, 49–54.
- Syengo, C.K.; Pyeye, S.; Orwa, G.O.; Odhiambo, R.O. Local Polynomial Regression Estimator of the Finite Population Total under Stratified Random Sampling: A Model-Based Approach. *Open J. Stat.* **2016**, *6*, 1085–1097.
- Chamidah, N.; Budiantara, I.N.; Sunaryo, S.; Zain, I. Designing of Child Growth Chart Based on Multi-Response Local Polynomial Modeling. *J. Math. Stat.* **2012**, *8*, 342–347.
- Opsomer, J.D.; Ruppert, D. Fitting a Bivariate Additive Model by Local Polynomial Regression. *Ann. Stat.* **1997**, *25*, 186–211.
- Maronge, J.M.; Zhai, Y.; Wiens, D.P.; Fang, Z. Optimal designs for spline wavelet regression models. *J. Stat. Plan. Inference* **2017**, *184*, 94–104.
- Antoniadis, A.; Bigot, J.; Sapatinas, T. Wavelet estimators in nonparametric regression: A comparative simulation study. *J. Stat. Softw.* **2001**, *6*, 1–83.
- Antoniadis, A.; Leblanc, F. Nonparametric Wavelet Regression for Binary Response. *Statistics* **2000**, *34*, 183–213.
- Sifriyani; Budiantara, I.N.; Kartiko, S.H.; Gunardi. Evaluation of Factors Affecting Increased Unemployment in East Java Using NGWR-TS Method. *Int. J. Sci. Basic Appl. Res.* **2019**, *46*, 123–142.
- Sifriyani; Kartiko, S.H.; Budiantara, I.N.; Gunardi. Development of Nonparametric Geographically Weighted Regression using Truncated Spline Approach. *Songklanakarin J. Sci. Technol.* **2018**, *40*, 909–920.
- Budiantara, I.N.; Ratnasari, V.; Ratna, M.; Zain, I. The Combination of Spline and Kernel Estimator for Nonparametric Regression and Its Properties. *Appl. Math. Sci.* **2015**, *9*, 6083–6094.
- Ratnasari, V.; Budiantara, I.N.; Ratna, M.; Zain, I. Estimation of Nonparametric Regression Curve using Mixed Estimator of Multivariable Truncated Spline and Multivariable Kernel. *Glob. J. Pure Appl. Math.* **2016**, *12*, 5047–5057.
- Hidayat, R.; Budiantara, I.N.; Otok, B.W.; Ratnasari, V. The regression curve estimation by using mixed smoothing spline and kernel (MsS-K) model. *Commun. Stat. Theory Methods* **2020**, doi:10.1080/03610926.2019.1710201.
- Budiantara, I.N.; Ratnasari, V.; Ratna, M.; Wibowo, W.; Afifah, N.; Rahmawati, D.P.; Octavanny, M.A.D. Modeling Percentage of Poor People In Indonesia Using Kernel And Fourier Series Mixed Estimator In Nonparametric Regression. *Rev. Investig. Operacional.* **2019**, *40*, 538–550.
- Nurcahayani, H.; Budiantara, I.N.; Zain, I. The Semiparametric Regression Curve Estimation by Using Mixed Truncated Spline and Fourier Series Model. *AIP Conf. Proc.* **2021**, *2329*, 060025, doi:10.1063/5.0042870.
- Mariati, N.P.A.M.; Budiantara, I.N.; Ratnasari, V. Combination Estimation of Smoothing Spline and Fourier Series in Nonparametric Regression. *J. Math.* **2020**, *2020*, 1–10.
- Sudiarsa, I.W.; Budiantara, I.N.; Suhartono; Purnami, S.W. Combined Estimator Fourier Series and Spline Truncated in Multivariable Nonparametric Regression. *Appl. Math. Sci.* **2015**, *9*, 4997–5010.
- Octavanny, M.A.D.; Budiantara, I.N.; Kuswanto, H.; Rahmawati, D.P. Nonparametric Regression Model for Longitudinal Data with Mixed Truncated Spline and Fourier Series. *Abstr. Appl. Anal.* **2020**, *2020*, 1–11, doi:10.1155/2020/4710745.
- Caldeira, J.F.; Gupta, R.; Torrent, H.S. Forecasting U.S. Aggregate Stock Market Excess Return: Do Functional Data Analysis Add Economic Value? *Mathematics* **2020**, *8*, 2042.

30. Correa-Quezada, R.; Cueva-Rodríguez, L.; Álvarez-García, J.; del Río-Rama, M.d.I.C. Application of the Kernel Density Function for the Analysis of Regional Growth and Convergence in the Service Sector through Productivity. *Mathematics* **2020**, *8*, 1234.
31. Green, P.J.; Silverman, B.W. *Nonparametric Regression and Generalized Linear Models*, 1st ed.; Chapman and Hall/CRC: New York, NY, USA, 1993.
32. Wang, Y.; Guo, W.; Brown, M.B. Spline smoothing for bivariate data with applications to association between hormones. *Stat. Sin.* **2000**, *10*, 377–397.
33. Wahba, G. *Spline Models for Observational Data*; SIAM Society For Industrial And Applied Mathematics: Philadelphia, PA, USA, 1990.
34. Craven, P., Wahba, G. Smoothing noisy data with spline functions—Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numer. Math.* **1978**, *31*, 377–403.
35. United Nations Development Programme. *Human Development Report 1990*; Oxford University Press: Jericho, NY, USA, 1990.
36. BPS—Statistics Indonesia. *Indeks Pembangunan Manusia 2018*; BPS—Statistics Indonesia: Jakarta, Indonesia, 2019.
37. BPS—Statistics of Jawa Timur Province. *Provinsi Jawa Timur Dalam Angka 2019*; BPS—Statistics of Jawa Timur Province: Jawa Timur, Indonesia, 2019.
38. Rahayu, A.; Purhadi; Sutikno; Prastyo, D.D. Multivariate Gamma Regression: Parameter Estimation, Hypothesis Testing, and Its Application. *Symmetry* **2020**, *12*, 813.
39. Sofilda, E.; Hermiyanti, P.; Hamzah, M.Z. Determinant Variable Analysis of Human Development Index in Indonesia (Case For High And Low Index At Period 2004–2013). *OIDA Int. J. Sustain. Dev.* **2015**, *8*, 11–28.
40. Dewanti, P.; Budiantara, I.N.; Rumiati, A.T. Modelling of SDG's Achievement in East Java Using Bi-responses Nonparametric Regression with Mixed Estimator Spline Truncated and Kernel. *J. Phys. Conf. Ser.* **2020**, *1562*, 012016.