

Research Article

Mixed Spline Smoothing and Kernel Estimator in Biresponse Nonparametric Regression

Dyah P. Rahmawati , **I. N. Budiantara** , **Dedy D. Prastyo** , and **Made A. D. Octavanny** 

Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

Correspondence should be addressed to I. N. Budiantara; i_nyoman_b@statistika.its.ac.id

Received 6 October 2020; Revised 9 February 2021; Accepted 17 February 2021; Published 12 March 2021

Academic Editor: Sergejs Solovjovs

Copyright © 2021 Dyah P. Rahmawati et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mixed estimators in nonparametric regression have been developed in models with one response. The biresponse cases with different patterns among predictor variables that tend to be mixed estimators are often encountered. Therefore, in this article, we propose a biresponse nonparametric regression model with mixed spline smoothing and kernel estimators. This mixed estimator is suitable for modeling biresponse data with several patterns (response vs. predictors) that tend to change at certain subintervals such as the spline smoothing pattern, and other patterns that tend to be random are commonly modeled using kernel regression. The mixed estimator is obtained through two-stage estimation, i.e., penalized weighted least square (PWLS) and weighted least square (WLS). Furthermore, the proposed biresponse modeling with mixed estimators is validated using simulation data. This estimator is also applied to the percentage of the poor population and human development index data. The results show that the proposed model can be appropriately implemented and gives satisfactory results.

1. Introduction

One of the most popular statistical methods often used for prediction is regression analysis. Regression analysis is commonly used to determine the functional relationship between independent variables (predictors) and dependent variables (responses) [1]. Functional relationships between predictor variables and response variables can have clear or unknown patterns; if these relationships have unknown patterns, the appropriate type of regression analysis is nonparametric regression [2]. In the nonparametric regression, the regression curve is assumed to be smooth. This regression has high flexibility because the data can drive to estimate its own regression curve without subjectivity from the researcher [3]. Researchers have proposed methods for estimating nonparametric regression functions such as spline, kernel, and Fourier series functions. The spline nonparametric regression has been developed by Eubank [3], Becher et al. [4], and Wang et al. [5]. Hall and Huang [6], Okumura and Naito [7], Du et al. [8], Chamidah and

Saifudin [9], and Erçelik and Nadar [10] developed kernel nonparametric regression. Bilodeau [11] and Amato et al. [12] estimated the nonparametric regression function with the Fourier series function.

In applying nonparametric regression modeling, researchers sometimes assume that each predictor variable has the same pattern. Although there are often real cases with different patterns between the response and each predictor, if the researcher still insists on applying one type of estimator to all predictor variables, the estimation results can be inaccurate and produce a large error. Researchers have begun to develop nonparametric regression with a mixed estimator, including Hidayat et al. [13], Mariati et al. [14], and Octavanny et al. [15]. These mixed estimators are formed by referring to the idea of semiparametric regression. The semiparametric regression model is an additive regression model that consists of a parametric component and a nonparametric component; see the work of Green and Yandell [16], Roozbeh and Arashi [17], and Roozbeh and Najarian [18]. In these

mixed estimators, the additive model concept in semi-parametric regression is adapted by modification using two different nonparametric components. However, these mixed estimators only use one response variable, even though some biresponse cases also have different patterns among the predictor variables. At present, nonparametric regression studies with mixed estimators for biresponse cases have never been developed. Therefore, this study develops a new theory about the mixed estimator of spline smoothing and kernel in biresponse nonparametric regression. This mixed estimator is the development of the mixed estimator proposed by Hidayat et al. [13] that the kernel estimator is considered to be fixed, while the kernel function in this paper is estimated. In addition, this new mixed estimator can be applied for biresponse cases. This mixed estimator is obtained through a two-stage estimation. The first stage of estimation uses the penalized weighted least square (PWLS) to obtain the spline smoothing component, followed by the second stage that employs the weighted least-square (WLS) estimation method to estimate the kernel component.

The spline smoothing estimator is very dependent on the smoothing parameter, while the kernel estimator is very dependent on the bandwidth parameter. These smoothing and bandwidth parameters are tuning parameters. The optimal value of these parameters will produce the best regression model. In nonparametric and semiparametric regression, there are several methods to determine the optimal parameter value to obtain the best regression model. Some of the popular methods are the cross-validation (CV) and generalized cross-validation (GCV) methods. The CV method is a method of selecting the best model based on the best predictive ability from all the different datasets. This method is widely used, but the calculation of this method will become more complex as the number of datasets increases. In addition, for partial linear models including mixed estimator models, the one-out crossover method tends to be time consuming even for moderate sample sizes [19]. Craven and Wahba [20] modified the CV method to make the calculation simpler, and the result of this modification is called the GCV method. This method is widely used by researchers because it has several advantages. The advantages of the GCV method include the following: simple and efficient in calculation, invariant to transformation, and does not require variant information. This method also has the advantage of optimal asymptotic properties over other methods [21, 22]. Some researchers develop specific GCV methods according to the model in their research such as the GCV for semiparametric ridge regression with kernel smoothing [23–25]; also, several types of GCV were developed for uniresponse nonparametric regression with mixed estimators including the mixed estimator of spline smoothing and kernel [13], mixed estimator of spline smoothing and Fourier series [14], and mixed estimator of truncated spline and fourier series for longitudinal data [15]. Therefore, in this study, the determination of the best model was carried out using the GCV method which was developed specifically for the mixed estimator of spline smoothing and kernel in biresponse nonparametric regression.

Next, the proposed mixed estimator is applied to the simulation data. The formula for generating data contains two different functions to represent two different patterns of predictor variables. This estimator is also implemented to model to the percentage of the poor population and human development index data in Papua Province, Indonesia. Empirical results indicate that the proposed mixed estimator performs very well for modeling the data with two different patterns. One pattern (response vs. predictors) tends to change at certain subintervals, and another pattern appears to be random, which are commonly modeled using kernel regression.

The rest of this paper is organized as follows: In Section 2, we present the materials and methods about the two-stage estimation method, i.e., the PWLS and followed by the WLS. The proposed mixed spline smoothing and kernel estimator in biresponse nonparametric regression is explained in Section 3.1. The selection of smoothing and bandwidth parameters using generalized cross validation (GCV) is described in Section 3.2. The simulation study and real data analysis are conducted to illustrate the performance of the proposed biresponse mixed estimator in Sections 3.3 and 3.4. The conclusions and further research are presented in the last section.

2. Materials and Methods

The paired data $(y_{1i}, y_{2i}, x_i, t_i)$ are given, where some predictor variables have a pattern that changes at certain subintervals, and the remaining predictors have a typically random pattern. This work adopts the idea of the semi-parametric regression model developed by Green and Yandell [16], which employed an additive model to combine parametric and nonparametric models. Thus, the additive model for nonparametric biresponse regression with two different estimators can be formulated as follows:

$$y_{hi} = \mu_{hi}(x_i, t_i) + \varepsilon_{hi}, \quad h = 1, 2; i = 1, 2, \dots, n, \quad (1)$$

where $\mu_{hi}(x_i, t_i)$ is a regression curve. In the biresponse cases, if y_{hi} and $y_{h'i}$ are in pairs, then there is a correlation between the error of h -response and error of h' -response ($h \neq h'$). Error correlation between responses can be defined as $\text{corr}(\varepsilon_{hi}, \varepsilon_{h'i}) = \rho$, where $\rho = \text{cov}(\varepsilon_{hi}, \varepsilon_{h'i}) / \sigma_h \sigma_{h'}$ $i = 1, 2, \dots, n; h = 1, 2; h \neq h'$ [5].

Each regression curve $\mu_{hi}(x_i, t_i)$ is assumed to be unknown and additive so that it can be written as

$$\mu_{hi}(x_i, t_i) = g_h(x_i) + f_h(t_i). \quad (2)$$

If equation (2) is substituted into equation (1), then equation (1) can be expressed in the following vector form:

$$\mathbf{y} = \mathbf{g} + \mathbf{f} + \boldsymbol{\varepsilon}. \quad (3)$$

The component of the regression curve \mathbf{g} is approximated by the spline smoothing function, which is assumed to be smooth and contained in the Sobolev space $\mathbf{g} \in W_2^m[a, b]$, while the component of the regression curve \mathbf{f} is approximated by the kernel function. Mixed spline smoothing and kernel estimator in biresponse

nonparametric regression can be obtained through the two-stage estimation method. The first stage is performed by estimating the spline smoothing component using the PWLS method, and the second stage is estimating the kernel component using the WLS method. To estimate the spline smoothing part, equation (3) is modified to the following form:

$$\text{Min}_{g_h \in W_2^m(a_h, b_h)} \left\{ (2n)^{-1} (\mathbf{z} - \mathbf{g})^T \mathbf{M}^{-1} (\mathbf{z} - \mathbf{g}) + \sum_{h=1}^2 \lambda_h \int_{a_h}^{b_h} [g_h^{(m)}(x)]^2 dx \right\}, \quad 0 < \lambda_h < \infty. \tag{5}$$

The estimation of the kernel component \hat{f} is obtained in the second stage estimation using WLS optimization in equation (6). Furthermore, the results of the two-stage estimations are substituted into equation (3) to obtain the mixed spline smoothing and kernel estimator in biresponse nonparametric regression.

$$\text{Min}_{\alpha_h} \{ (\mathbf{y} - \hat{\mathbf{g}} - \mathbf{f})^T \mathbf{M}^{-1} \mathbf{K}_\alpha(t_0) (\mathbf{y} - \hat{\mathbf{g}} - \mathbf{f}) \}, \tag{6}$$

where $\mathbf{K}_\alpha(t_0)$ is a weighting matrix for the kernel estimator and \mathbf{M}^{-1} in equations (5) and (6) is a weighting matrix for the biresponse nonparametric regression like in the work of Wang et al. [5].

3. Results and Discussion

3.1. Mixed Spline Smoothing and Kernel Estimator in Biresponse Nonparametric Regression. Before conducting the estimation to obtain biresponse mixed spline smoothing and kernel estimator, it is necessary to obtain the function form for each component of this mixed estimator. The structure of the spline smoothing component is explained in Lemma 1, while the kernel component is explained in Lemma 2.

Lemma 1. *If the regression curve \mathbf{g} is assumed to be smooth and contained in the Sobolev space $\mathbf{g} \in W_2^m[a_h, b_h]$ [22], then the function form of the spline smoothing component in the biresponse nonparametric regression can be stated as*

$$\mathbf{g} = \mathbf{U} \mathbf{d} + \mathbf{V} \mathbf{c}, \tag{7}$$

with

$$\begin{aligned} \mathbf{U} &= \begin{pmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 \end{pmatrix}, \\ \mathbf{d} &= \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix}, \\ \mathbf{V} &= \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{pmatrix}, \\ \mathbf{c} &= \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}, \end{aligned} \tag{8}$$

$$\mathbf{z} = \mathbf{g} + \boldsymbol{\varepsilon}, \tag{4}$$

where $\mathbf{z} = \mathbf{y} - \mathbf{f}$.

Estimation of the spline smoothing component $\hat{\mathbf{g}}$ can be performed by applying PWLS optimization to equation (5), and then, the estimation results in the first stage are substituted into equation (3).

where

$$\begin{aligned} \mathbf{U}_h &= \begin{pmatrix} \langle \eta_{h1}, \tau_{h1} \rangle & \langle \eta_{h1}, \tau_{h2} \rangle & \cdots & \langle \eta_{h1}, \tau_{hm} \rangle \\ \langle \eta_{h2}, \tau_{h1} \rangle & \langle \eta_{h2}, \tau_{h2} \rangle & \cdots & \langle \eta_{h2}, \tau_{hm} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_{hm}, \tau_{h1} \rangle & \langle \eta_{hm}, \tau_{h2} \rangle & \cdots & \langle \eta_{hm}, \tau_{hm} \rangle \end{pmatrix}, \\ \mathbf{d}_h &= \begin{pmatrix} d_{h1} \\ d_{h2} \\ \vdots \\ d_{hm} \end{pmatrix}, \\ \mathbf{V}_h &= \begin{pmatrix} \langle \eta_{h1}, \beta_{h1} \rangle & \langle \eta_{h1}, \beta_{h2} \rangle & \cdots & \langle \eta_{h1}, \beta_{hm} \rangle \\ \langle \eta_{h2}, \beta_{h1} \rangle & \langle \eta_{h2}, \beta_{h2} \rangle & \cdots & \langle \eta_{h2}, \beta_{hm} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_{hm}, \beta_{h1} \rangle & \langle \eta_{hm}, \beta_{h2} \rangle & \cdots & \langle \eta_{hm}, \beta_{hm} \rangle \end{pmatrix}, \\ \mathbf{c}_h &= \begin{pmatrix} c_{h1} \\ c_{h2} \\ \vdots \\ c_{hm} \end{pmatrix}, \end{aligned} \tag{9}$$

for $h = 1, 2$.

$\{\tau_{h1}, \tau_{h2}, \dots, \tau_{hm}\}$ and $\{\beta_{h1}, \beta_{h2}, \dots, \beta_{hm}\}$ are bases of spaces in the Sobolev space.

Proof. If $g_h; h = 1, 2$ is a function lying in Hilbert space H , the H space can be decomposed into a direct sum of two spaces H_0 and H_1 where $H = H_0 \oplus H_1$ and $H_0 \perp H_1$. If $\{\tau_{h1}, \tau_{h2}, \dots, \tau_{hm}\}$ is the basis in H_0 and $\{\beta_{h1}, \beta_{h2}, \dots, \beta_{hm}\}$ is the basis in H_1 , according to Wahba [22], for each function $g_h \in H$ with $u_h \in H_0$ and $v_h \in H_1$, we obtain

$$g_h = u_h + v_h = \sum_{l=1}^m d_{hl} \tau_{hl} + \sum_{i=1}^n c_{hi} \beta_{hi} = \boldsymbol{\tau}_h^T \mathbf{d}_h + \boldsymbol{\beta}_h^T \mathbf{c}_h, \tag{10}$$

where

$$\begin{aligned}
 \boldsymbol{\tau}_h &= \begin{pmatrix} \tau_{h1} \\ \tau_{h2} \\ \vdots \\ \tau_{hm} \end{pmatrix}, \\
 \mathbf{d}_h &= \begin{pmatrix} d_{h1} \\ d_{h2} \\ \vdots \\ d_{hm} \end{pmatrix}, \\
 \boldsymbol{\beta}_h &= \begin{pmatrix} \beta_{h1} \\ \beta_{h2} \\ \vdots \\ \beta_{hm} \end{pmatrix}, \\
 \mathbf{c}_h &= \begin{pmatrix} c_{h1} \\ c_{h2} \\ \vdots \\ c_{hm} \end{pmatrix}.
 \end{aligned} \tag{11}$$

Equation (10) is a limited linear function in H space and $g_h \in H$; therefore, equation (10) can be stated as

$$\mathcal{L}_x g_h = \langle \eta_{hi}, g_h \rangle = g_h(x_i), \tag{12}$$

and using inner product properties, equation (12) can be written as

$$\begin{aligned}
 g_h(x_i) &= \langle \eta_{hi}, g_h \rangle, \\
 &= \langle \eta_{hi}, \boldsymbol{\tau}_h^T \mathbf{d}_h \rangle + \langle \eta_{hi}, \boldsymbol{\beta}_h^T \mathbf{c}_h \rangle, \\
 &= d_{h1} \langle \eta_{hi}, \tau_{h1} \rangle + \dots + d_{hm} \langle \eta_{hi}, \tau_{hm} \rangle + c_{h1} \langle \eta_{hi}, \beta_{h1} \rangle \\
 &\quad + \dots + c_{hm} \langle \eta_{hi}, \beta_{hm} \rangle,
 \end{aligned} \tag{13}$$

and for each response with all observations ($i = 1, 2, 3, \dots, n$), we can obtain the following vector \mathbf{g}_h :

$$\begin{aligned}
 \mathbf{g}_h &= (g_h(x_1) \ g_h(x_2) \ \dots \ g_h(x_n))^T, \\
 \mathbf{g}_h &= \begin{pmatrix} d_{h1} \langle \eta_{h1}, \tau_{h1} \rangle + \dots + d_{hm} \langle \eta_{h1}, \tau_{hm} \rangle + c_{h1} \langle \eta_{h1}, \beta_{h1} \rangle + \dots + c_{hm} \langle \eta_{h1}, \beta_{hm} \rangle \\ d_{h1} \langle \eta_{h2}, \tau_{h1} \rangle + \dots + d_{hm} \langle \eta_{h2}, \tau_{hm} \rangle + c_{h1} \langle \eta_{h2}, \beta_{h1} \rangle + \dots + c_{hm} \langle \eta_{h2}, \beta_{hm} \rangle \\ \vdots \\ d_{h1} \langle \eta_{hm}, \tau_{h1} \rangle + \dots + d_{hm} \langle \eta_{hm}, \tau_{hm} \rangle + c_{h1} \langle \eta_{hm}, \beta_{h1} \rangle + \dots + c_{hm} \langle \eta_{hm}, \beta_{hm} \rangle \end{pmatrix}, \\
 \mathbf{g}_h &= \mathbf{U}_h \mathbf{d}_h + \mathbf{V}_h \mathbf{c}_h,
 \end{aligned} \tag{14}$$

where

$$\begin{aligned}
 \mathbf{U}_h &= \begin{pmatrix} \langle \eta_{h1}, \tau_{h1} \rangle & \langle \eta_{h1}, \tau_{h2} \rangle & \dots & \langle \eta_{h1}, \tau_{hm} \rangle \\ \langle \eta_{h2}, \tau_{h1} \rangle & \langle \eta_{h2}, \tau_{h2} \rangle & \dots & \langle \eta_{h2}, \tau_{hm} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_{hm}, \tau_{h1} \rangle & \langle \eta_{hm}, \tau_{h2} \rangle & \dots & \langle \eta_{hm}, \tau_{hm} \rangle \end{pmatrix}, \\
 \mathbf{d}_h &= \begin{pmatrix} d_{h1} \\ d_{h2} \\ \vdots \\ d_{hm} \end{pmatrix}, \\
 \mathbf{V}_h &= \begin{pmatrix} \langle \eta_{h1}, \beta_{h1} \rangle & \langle \eta_{h1}, \beta_{h2} \rangle & \dots & \langle \eta_{h1}, \beta_{hm} \rangle \\ \langle \eta_{h2}, \beta_{h1} \rangle & \langle \eta_{h2}, \beta_{h2} \rangle & \dots & \langle \eta_{h2}, \beta_{hm} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_{hm}, \beta_{h1} \rangle & \langle \eta_{hm}, \beta_{h2} \rangle & \dots & \langle \eta_{hm}, \beta_{hm} \rangle \end{pmatrix}, \\
 \mathbf{c}_h &= \begin{pmatrix} c_{h1} \\ c_{h2} \\ \vdots \\ c_{hm} \end{pmatrix},
 \end{aligned} \tag{15}$$

where $\langle \eta_{hi}, \tau_{hs} \rangle = x_i^{s-1} / (s-1)!$ and $\langle \eta_{hi}, \beta_{hi'} \rangle$ are equal to $\langle \beta_{hi}, \beta_{hi'} \rangle$ with $\langle \beta_{hi}, \beta_{hi'} \rangle = \int_a^b ((x_i - u)_+^{m-1} (x_i - u)_+^{m-1}) / ((m-1)!)^2 du$ $i = 1, 2, \dots, n; i' = 1, 2, \dots, n; s = 1, 2, \dots, m; h = 1, 2$

Therefore, for all responses ($h = 1, 2$), we can obtain the function form of the spline smoothing component in the biresponse regression curve as follows:

$$\begin{aligned} \mathbf{g} &= \begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{pmatrix}, \\ \mathbf{g} &= \begin{pmatrix} \mathbf{U}_1 \mathbf{d}_1 + \mathbf{V}_1 \mathbf{c}_1 \\ \mathbf{U}_2 \mathbf{d}_2 + \mathbf{V}_2 \mathbf{c}_2 \end{pmatrix}, \\ \mathbf{g} &= \begin{pmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 \end{pmatrix} \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{pmatrix} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}, \\ \mathbf{g} &= \mathbf{Ud} + \mathbf{Vc}. \end{aligned} \tag{16}$$

□

Lemma 2. *If the regression curve \mathbf{f} is approached by the kernel function, then the function form of kernel component in the biresponse nonparametric regression can be expressed as*

$$f_h(t) \approx f_h(t_0) + \frac{(t-t_0)^1 f_h^1(t_0)}{1!} + \frac{(t-t_0)^2 f_h^2(t_0)}{2!} + \dots + \frac{(t-t_0)^{m_h} f_h^{m_h}(t_0)}{m_h!}, \tag{20}$$

and if $\omega_{h,d}(t_0) = f_h^d(t_0)/d!$, $d = 0, 1, \dots, m_h; h = 1, 2$, then equation (20) can be stated as

$$f_h(t) \approx \omega_{h,0}(t_0) + (t-t_0)^1 \omega_{h,1}(t_0) + (t-t_0)^2 \omega_{h,2}(t_0) + \dots + (t-t_0)^{m_h} \omega_{h,m_h}(t_0), \tag{21}$$

where $t \in (t_0 - \alpha, t_0 + \alpha)$, t_0 is the predictor value for prediction.

The kernel estimator can be obtained when the polynomial order $m_h = 0; h = 1, 2$. Therefore, the function form for each response involving all observations can be stated as

$$\begin{aligned} \mathbf{f}_h(\mathbf{t}) &= (f_h(t_1) \ f_h(t_2) \ \dots \ f_h(t_n))^T, \\ \mathbf{f}_h(\mathbf{t}) &= (\omega_{h,0}(t_{0(1)}) \ \omega_{h,0}(t_{0(2)}) \ \dots \ \omega_{h,0}(t_{0(n)}))^T. \end{aligned} \tag{22}$$

Then, we can obtain the function form of component \mathbf{f} in biresponse nonparametric regression as follows:

$$\mathbf{f} = \omega_0(t_0), \tag{23}$$

$$\mathbf{f} = \omega_0(t_0), \tag{17}$$

where

$$\begin{aligned} \omega_0(t_0) &= \left((\omega_{1,0}(t_0))^T \ (\omega_{2,0}(t_0))^T \right)^T, \\ \mathbf{f} &= (\mathbf{f}_1(\mathbf{t}) \ \mathbf{f}_2(\mathbf{t}))^T, \end{aligned} \tag{18}$$

with

$$\mathbf{f}_h(\mathbf{t}) = (f_h(t_1) \ f_h(t_2) \ \dots \ f_h(t_n))^T, \quad h = 1, 2. \tag{19}$$

The function $f_h(t_i); h = 1, 2; i = 1, 2, \dots, n$ is approached by the Taylor series with t around t_0 .

Proof. The form of the regression function \mathbf{f} is derived from the component $f_h(t_i)$ in equation (2). The form of the function is unknown and approached using the kernel estimator. The function $f_h(t_i)$ for $h = 1, 2$ can be approached by the Taylor series with t around t_0 as follows [9]:

$$\begin{aligned} \mathbf{f} &= (\mathbf{f}_1(\mathbf{t}) \ \mathbf{f}_2(\mathbf{t}))^T; \\ \omega_0(t_0) &= \left((\omega_{1,0}(t_0))^T \ (\omega_{2,0}(t_0))^T \right)^T. \end{aligned} \tag{24}$$

Theorem 1. *The biresponse nonparametric regression model is given in equation (1), where each component of the regression curve is additive as stated in equation (2). The function form of the \mathbf{g} component is presented in Lemma 1, and the function form of the \mathbf{f} component is presented in Lemma 2. Using PWLS in the first-stage estimation and WLS in the second-stage estimation, the mixed spline smoothing and kernel estimator in biresponse nonparametric regression is obtained as follows:*

$$\hat{\boldsymbol{\mu}} = \mathbf{H}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y}, \tag{24}$$

where

$$\begin{aligned} \mathbf{H}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) &= (\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) + \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha})), \\ \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) &= \mathbf{U}(\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1} + \mathbf{V}\mathbf{R}^{-1}\mathbf{M}^{-1}\left(\mathbf{I} - \mathbf{U}\left((\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1}\right)\right), \\ \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) &= (\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})^T(\mathbf{D}\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) - 2\mathbf{D}) + \mathbf{D})^{-1}(\mathbf{I} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})^T)\mathbf{D}(\mathbf{I} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})), \end{aligned} \quad (25)$$

while $\widehat{\mathbf{g}} = \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\widehat{\mathbf{f}}$ is obtained from the first-stage estimation and $\widehat{\mathbf{f}} = \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y}$ is obtained from the second-stage estimation.

Proof. The first-stage estimation on the mixed spline smoothing and kernel estimator in biresponse nonparametric regression is performed by estimating the spline smoothing component using the PWLS in equation (5). The penalty component $\sum_{h=1}^2 \lambda_h \int_{a_h}^{b_h} [g_h^{(m)}(x)]^2 dx$ in equation (5) can be obtained through the following decomposition [26]:

$$\begin{aligned} \int_{a_h}^{b_h} [g_h^{(m)}(x)]^2 dx &= \|P_1 g_h\|^2, \\ &= \langle P_1 g_h, P_1 g_h \rangle, \\ &= \langle P_1(\boldsymbol{\tau}_h^T \mathbf{d}_h + \boldsymbol{\beta}_h^T \mathbf{c}_h), P_1(\boldsymbol{\tau}_h^T \mathbf{d}_h + \boldsymbol{\beta}_h^T \mathbf{c}_h) \rangle, \\ &= \langle \boldsymbol{\beta}_h^T \mathbf{c}_h, \boldsymbol{\beta}_h^T \mathbf{c}_h \rangle, \\ &= (\boldsymbol{\beta}_h^T \mathbf{c}_h)^T (\boldsymbol{\beta}_h^T \mathbf{c}_h), \\ &= \mathbf{c}_h^T (\boldsymbol{\beta}_h \boldsymbol{\beta}_h^T) \mathbf{c}_h, \\ &= \mathbf{c}_h^T \mathbf{V}_h \mathbf{c}_h, \end{aligned} \quad (26)$$

where P_1 is the orthogonal projection of g_h to H_1 in H space. By substituting equation (26) into the penalty component, we obtain

$$\begin{aligned} \sum_{h=1}^2 \lambda_h \int_{a_h}^{b_h} [g_h^{(m)}(x)]^2 dx &= \sum_{h=1}^2 \lambda_h \mathbf{c}_h^T \mathbf{V}_h \mathbf{c}_h, \\ &= (\mathbf{c}_1 \ \mathbf{c}_2) \begin{pmatrix} \lambda_1 \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \lambda_2 \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{pmatrix} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}, \\ &= \mathbf{c}^T \boldsymbol{\Lambda} \mathbf{V} \mathbf{c}, \quad [26], \end{aligned} \quad (27)$$

and furthermore, the PWLS optimization in equation (5) can be written in matrix notation as follows:

$$\text{Min}_{g_h \in W_2^m(a_h, b_h)} \{ (2n)^{-1} (\mathbf{z} - \mathbf{U}\mathbf{d} - \mathbf{V}\mathbf{c})^T \mathbf{M}^{-1} (\mathbf{z} - \mathbf{U}\mathbf{d} - \mathbf{V}\mathbf{c}) + \mathbf{c}^T \boldsymbol{\Lambda} \mathbf{V} \mathbf{c} \} = \text{Min}_{g_h \in W_2^m(a_h, b_h)} \{ Q(\mathbf{c}, \mathbf{d}) \}. \quad (28)$$

The solution for the PWLS optimization can be obtained from the partial derivative $Q(\mathbf{c}, \mathbf{d})$ by \mathbf{c} and \mathbf{d} . The partial derivative $Q(\mathbf{c}, \mathbf{d})$ by \mathbf{c} results as follows:

$$\begin{aligned} \frac{\partial Q(\mathbf{c}, \mathbf{d})}{\partial \mathbf{c}} &= -2\mathbf{V}^T \mathbf{M}^{-1} \mathbf{z} + 2\mathbf{V}^T \mathbf{M}^{-1} \mathbf{U} \mathbf{d} + 2\mathbf{V}^T \mathbf{M}^{-1} \mathbf{V} \widehat{\mathbf{c}} + 2(2n)\boldsymbol{\Lambda} \mathbf{V} \widehat{\mathbf{c}} = \mathbf{0}, \\ 2\mathbf{V}^T [-\mathbf{M}^{-1} \mathbf{z} + \mathbf{M}^{-1} \mathbf{U} \mathbf{d} + (\mathbf{M}^{-1} \mathbf{V} + 2n\boldsymbol{\Lambda} \mathbf{I}) \widehat{\mathbf{c}}] &= \mathbf{0}, \\ -\mathbf{M}^{-1} \mathbf{z} + \mathbf{M}^{-1} \mathbf{U} \mathbf{d} + (\mathbf{M}^{-1} \mathbf{V} + 2n\boldsymbol{\Lambda}) \widehat{\mathbf{c}} &= \mathbf{0}, \end{aligned} \quad (29)$$

if $\mathbf{R} = \mathbf{M}^{-1}\mathbf{V} + 2n\Lambda$; thus, we can obtain

$$-\mathbf{M}^{-1}\mathbf{z} + \mathbf{M}^{-1}\mathbf{U}\mathbf{d} + \mathbf{R}\hat{\mathbf{c}} = 0, \tag{30}$$

$$\hat{\mathbf{c}} = \mathbf{R}^{-1}(\mathbf{M}^{-1}\mathbf{z} - \mathbf{M}^{-1}\mathbf{U}\mathbf{d}).$$

The partial derivative of $Q(\mathbf{c}, \mathbf{d})$ by \mathbf{d} gives the result

$$\frac{\partial Q(\mathbf{c}, \mathbf{d})}{\partial \mathbf{d}} = -2\mathbf{U}^T\mathbf{M}^{-1}\mathbf{z} + 2\mathbf{U}^T\mathbf{M}^{-1}\mathbf{U}\hat{\mathbf{d}} + 2\mathbf{U}^T\mathbf{M}^{-1}\mathbf{V}\hat{\mathbf{c}} = \mathbf{0}, \tag{31}$$

and by substituting equation (30) into equation (31), we can obtain

$$-2\mathbf{U}^T\mathbf{M}^{-1}\mathbf{z} + 2\mathbf{U}^T\mathbf{M}^{-1}\mathbf{U}\hat{\mathbf{c}} + 2\mathbf{U}^T\mathbf{M}^{-1}\mathbf{V}[\mathbf{R}^{-1}(\mathbf{M}^{-1}\mathbf{z} - \mathbf{M}^{-1}\mathbf{U}\hat{\mathbf{d}})] = \mathbf{0}, \tag{32}$$

considering that $\mathbf{R} = \mathbf{M}^{-1}\mathbf{V} + 2n\Lambda$ so matrix \mathbf{V} can be stated as $\mathbf{V} = \mathbf{M}(\mathbf{R} - 2n\Lambda)$; then, we can modify

$$\begin{aligned} \hat{\mathbf{g}} &= \mathbf{U}\hat{\mathbf{d}} + \mathbf{V}\hat{\mathbf{c}}, \\ &= \mathbf{U}(\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{z} + \mathbf{V}\mathbf{R}^{-1}\mathbf{M}^{-1}(\mathbf{I} - \mathbf{U}((\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1})), \\ &= (\mathbf{U}(\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1} + \mathbf{V}\mathbf{R}^{-1}\mathbf{M}^{-1}(\mathbf{I} - \mathbf{U}((\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1}))), \\ &= \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{z}, \end{aligned} \tag{36}$$

where

$$\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) = \mathbf{U}(\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1} + \mathbf{V}\mathbf{R}^{-1}\mathbf{M}^{-1}(\mathbf{I} - \mathbf{U}((\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1})). \tag{37}$$

Because $\mathbf{z} = \mathbf{y} - \mathbf{f}$, the first-stage estimation results can be stated as

$$\hat{\mathbf{g}} = \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{f}, \tag{38}$$

and remember that $\mathbf{z} = \mathbf{g} + \boldsymbol{\varepsilon}$; then, the model in equation (4) can be written as

$$\text{Min}_{\alpha_h} \{ \boldsymbol{\varepsilon}^T \mathbf{M}^{-1} \mathbf{K}_{\boldsymbol{\alpha}}(t_0) \boldsymbol{\varepsilon} \} = \text{Min}_{\alpha_h} \{ (\mathbf{y} - \hat{\mathbf{g}} - \mathbf{f})^T \mathbf{M}^{-1} \mathbf{K}_{\boldsymbol{\alpha}}(t_0) (\mathbf{y} - \hat{\mathbf{g}} - \mathbf{f}) \}, \tag{40}$$

where \mathbf{M}^{-1} is a weighting matrix for biresponse nonparametric regression and $\mathbf{K}_{\boldsymbol{\alpha}}(t_0)$ is a weighting for the kernel estimator with the following structure:

$$\mathbf{K}_{\boldsymbol{\alpha}}(t_0) = \text{diag}(\mathbf{K}_{\alpha_1}(t_i - t_0), \mathbf{K}_{\alpha_2}(t_i - t_0)), \tag{41}$$

where $\mathbf{K}_{\alpha_h}(t_i - t_0) = (K_{\alpha_h}(t_1 - t_0) \ K_{\alpha_h}(t_2 - t_0) \ \dots \ K_{\alpha_h}(t_n - t_0))^T$ and $K_{\alpha_h}(\cdot)$ is the kernel function.

By substituting the results of the first-stage estimation (38) and the function form \mathbf{f} (Lemma 2) into the

$$\mathbf{M}^{-1}\mathbf{V}\mathbf{R}^{-1} = \mathbf{M}^{-1}(\mathbf{M}(\mathbf{R} - 2n\Lambda))\mathbf{R}^{-1} = \mathbf{I} - 2n\Lambda\mathbf{R}^{-1}. \tag{33}$$

Equation (33) is substituted into equation (32), and then, we solve it and get the following result:

$$\hat{\mathbf{d}} = (\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{z}. \tag{34}$$

Furthermore, by substituting equation (34) to equation (30), we obtain

$$\hat{\mathbf{c}} = \mathbf{R}^{-1}\mathbf{M}^{-1}(\mathbf{I} - \mathbf{U}((\mathbf{R}^{-1}\mathbf{M}^{-1}\mathbf{U})^{-1}\mathbf{R}^{-1}\mathbf{M}^{-1}))\mathbf{z}. \tag{35}$$

$\hat{\mathbf{c}}$ and $\hat{\mathbf{d}}$ are substituted into the function form of the spline smoothing component in equation (16), and then, the following spline smoothing estimator component in the biresponse nonparametric regression model is obtained:

$$\begin{aligned} \hat{\mathbf{z}} &= \hat{\mathbf{g}}, \\ \hat{\mathbf{z}} &= \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{f}. \end{aligned} \tag{39}$$

In the second stage of estimation, the function \mathbf{f} as the kernel component on a mixed spline smoothing and kernel estimator in biresponse nonparametric regression is estimated using the WLS method with the following formula:

model of the mixed estimator (equation (3)), the error of this mixed estimator model can be written as follows:

$$\begin{aligned} \boldsymbol{\varepsilon} &= \mathbf{y} - \hat{\mathbf{g}} - \mathbf{f}, \\ \boldsymbol{\varepsilon} &= \mathbf{y} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y} + \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\boldsymbol{\omega}_0(t_0) - \boldsymbol{\omega}_0(t_0). \end{aligned} \tag{42}$$

Furthermore, by supposing $\mathbf{D} = \mathbf{M}^{-1}\mathbf{K}_{\boldsymbol{\alpha}}(t_0)$ and substituting equation (42) into equation (40), we can get the following equation for WLS optimization:

$$\text{Min}_{\alpha_h} \{ (\mathbf{y} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y} + \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\boldsymbol{\omega}_0(t_0) - \boldsymbol{\omega}_0(t_0))^T \mathbf{D}(\mathbf{y} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y} + \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\boldsymbol{\omega}_0(t_0) - \boldsymbol{\omega}_0(t_0)) \} = \text{Min}_{\alpha_h} \{ Q(t_0) \}. \quad (43)$$

The solution for the WLS optimization can be obtained from the partial derivative $Q(t_0)$ by $\boldsymbol{\omega}_0(t_0)$. The optimization result is obtained as follows:

$$\hat{\boldsymbol{\omega}}_0(t_0) = (\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})^T (\mathbf{D}\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) - 2\mathbf{D}) + \mathbf{D})^{-1} (\mathbf{I} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})^T) \mathbf{D} (\mathbf{I} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})) \mathbf{y}. \quad (44)$$

Therefore, the estimation for kernel estimator component $\hat{\mathbf{f}}$ can be written as

$$\begin{aligned} \hat{\mathbf{f}} &= \hat{\boldsymbol{\omega}}_0(t_0), \\ \hat{\mathbf{f}} &= \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y}, \end{aligned} \quad (45)$$

where

$$\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) = (\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})^T (\mathbf{D}\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) - 2\mathbf{D}) + \mathbf{D})^{-1} (\mathbf{I} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})^T) \mathbf{D} (\mathbf{I} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})). \quad (46)$$

Based on the first-stage estimation results in equation (38) and the second-stage estimation results in equation (45), the estimation of the additive regression curve in equation (2) with the mixed spline smoothing and kernel estimator in biresponse nonparametric regression can be stated as

$$\begin{aligned} \hat{\boldsymbol{\mu}} &= \hat{\mathbf{g}} + \hat{\mathbf{f}}, \\ \hat{\boldsymbol{\mu}} &= \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y} - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\hat{\mathbf{f}} + \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y}, \\ \hat{\boldsymbol{\mu}} &= (\mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) + \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha}))\mathbf{y}, \end{aligned} \quad (47)$$

and we can write equation (47) as

$$\hat{\boldsymbol{\mu}} = \mathbf{H}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{y}, \quad (48)$$

where $\mathbf{H}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) = \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) - \mathbf{A}(\boldsymbol{\lambda}, \boldsymbol{\alpha})\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) + \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{\alpha})$. \square

3.2. Selection of Smoothing and Bandwidth Parameters. The best model for the biresponse mixed spline smoothing and kernel estimator depends on the optimal smoothing parameters (λ_{opt}) and optimal bandwidth parameters (α_{opt}), where λ and α are tuning parameters. These optimal parameters can be obtained from the model with the smallest generalized cross-validation (GCV) value, see [13–15, 23–25]. The GCV criteria are one of the methods to

determine the best model in nonparametric regression [20]. The GCV formula for the mixed spline smoothing and kernel estimator in biresponse nonparametric regression (48) can be stated as follows:

$$\text{GCV}(\boldsymbol{\lambda}, \boldsymbol{\alpha}) = \frac{(2n)^{-1} \|(\mathbf{I} - \mathbf{H}(\boldsymbol{\lambda}, \boldsymbol{\alpha}))\mathbf{y}\|^2}{((2n)^{-1} \text{trace}(\mathbf{I} - \mathbf{H}(\boldsymbol{\lambda}, \boldsymbol{\alpha})))^2}. \quad (49)$$

3.3. Simulation Study. In this simulation, the mixed spline smoothing and kernel estimator in biresponse nonparametric regression, as formulated in equation (48), is applied to the simulation data. The simulation data are generated from a formula that contains two different functions, i.e., a polynomial function and an exponential function, to present two different patterns for each predictor. The polynomial function is used to generate spline smoothing-like patterns, and the exponential function is used to generate patterns in the data such as kernel. Using two response variables and two predictor variables, the formula for generating data is defined as follows:

$$y_{hi} = f_{hi}(x_i, t_i) + \varepsilon_{hi}, \quad h = 1, 2; i = 1, 2, \dots, n, \quad (50)$$

with

$$\begin{aligned} f_{1i}(x_i, t_i) &= 9.75(x_i - 1)(1 - x_i)^2 + 1.35(e^{-4.05t_i} - 4e^{-7.12t_i} + 3e^{-9.74t_i}), \\ f_{2i}(x_i, t_i) &= 6.89(x_i - 1)(1 - x_i)^2 + 0.23(e^{-2.35t_i} - 4e^{-5.72t_i} + 3e^{-7.25t_i}). \end{aligned} \quad (51)$$

The predictors are generated from $x_i \sim U(0, 2)$ and $t_i \sim U(0, 2)$ with the sample size $n = 100$, and the random errors ε_{ii} are generated from bivariate normal distributions with $\mu_1 = 0, \mu_2 = 0, \sigma_1^2 = 0.5, \sigma_2^2 = 0.6$, and $\rho = 0.6$. The scatterplots of the simulated data are shown in Figure 1. It can be seen that the pattern between y_1, y_2 against x tends to change at certain subintervals such as the spline smoothing pattern, while the scatterplot between y_1, y_2 against t tends to have a random pattern that is commonly modeled with kernel regression.

In this simulation, the Gaussian kernel was employed. Based on the empirical results from the two-stage estimation, we obtain combination values of the smoothing parameters (λ_1, λ_2) and bandwidth parameters (α_1, α_2) around the optimal values (Table 1). The exhibited results are a few of all combinations due to the limited space. The best model is chosen based on the smallest GCV value resulted from optimal smoothing parameters $\lambda_{1(\text{opt})} = 0.0002413$ and $\lambda_{2(\text{opt})} = 0.0000899$ along with optimal bandwidth parameters $\alpha_{1(\text{opt})} = 2.396$ and $\alpha_{2(\text{opt})} = 2.416$. This model produces the lowest GCV = 3.239 with $R^2 = 99.91\%$ and RMSE = 0.1045.

The result of modeling simulation data using the mixed spline smoothing and kernel estimator is compared with modeling using either a spline smoothing estimator or kernel estimator only. The results of these models are presented in Table 2. From Table 2, we can find out the best model that gives the smallest GCV value is the model with the mixed spline smoothing and kernel estimator. Besides, this model has the largest R^2 and the lowest MSE value.

The plot between the estimation results and the original simulation data is presented in Figure 2, where the estimated values (red triangles) are very close to the original data (black squares). Thus, the proposed model and estimation procedure can be used to make a prediction correctly. Furthermore, on the left side of Figure 3, the surface plots are formed using equation (50), which is the equation for generating simulation data, whereas on the right side of Figure 3, there are two surface plots for each response which are formed from equation (48) where its parameters are estimated using two-stage estimation, i.e., the PWLS and WLS. The two sides of Figure 3 show that the plots appear to have similar surface shape. This evidence indicates that the estimation procedure proposed in equation (48) can be used appropriately to estimate the function generated from the simulation.

3.4. Data Application. The biresponse mixed spline smoothing and kernel estimator proposed in this paper is applied to regress the percentage of the poor population (PPP), as the first response, and human development index (HDI), as the second response, on several predictors. These two response variables are important because they are indicators of the success level of a country's development. The adoption of biresponse modeling for the two variables considers the initial study that there is a negative correlation between PPP and HDI. If the PPP in a region is getting lower, the HDI in that region will be higher [27, 28].

Several variables that typically affect the two response variables are the gross regional domestic product (GRDP) and the population growth rate. Some researchers have pointed out several factors that can affect the PPP and HDI, including Grubaugh [29], who stated the variables that influence the growth of HDI in developing countries are population, population growth, and the initial level of the gross domestic product (GDP). Meanwhile, Mallick and Ghani [30] found that high population growth is the cause of poverty in Pakistan. While in North Sumatra Province, Indonesia, the GRDP and education level to university have a positive and significant influence in reducing the PPP [31]. Based on Malthus's theory, poverty is considered as the impact of high population growth rates [32]. Also, additional life support needs are considered slower than population growth. A high increase in population growth will have an impact on decreasing the quality of natural resources and reducing the opportunity for people to access life-support facilities. This situation can reduce the quality of human life, and people will be challenged to live in prosperity. Rapid economic growth is one way to alleviate poverty [33]. The GDP or GRDP has a close relationship with economic growth because the economic growth of a region is related to an increase in production or an increase in income per capita. Besides, if the GRDP is higher, the income per capita in the region will increase and have an impact on increasing the ability of the community to meet their needs and improve their quality of life.

The application of this mixed estimator model on the PPP and HDI in this study is made after the authors conducted a preliminary study. Based on information from the initial research, it is known that the GRDP has a changing pattern at certain subintervals such as the spline pattern. In contrast, the population growth rate has a random pattern that is usually modeled by the kernel. Therefore, the data of PPP and HDI of Papua Province in 2017 are used as response variables, while the predictor variables are the GRDP and the population growth rate. The data were obtained from Statistics Indonesia (Badan Pusat Statistik-BPS), Papua Province, with 29 regencies/cities as the observation unit. The biresponse mixed spline smoothing and kernel estimator in equation (12) was applied to the data. This modeling produces the minimum GCV value of 62.75, the R^2 of 96.54%, and the RMSE of 3.166. Based on the R^2 value, it is known that the model can describe the relationship between predictor variables and response up to 96.54%. This finding shows that the biresponse mixed spline smoothing and kernel estimator is suitable for modeling the PPP and HDI in Papua Province. Also, the bar chart in Figures 4 and 5 show that the estimated values of PPP and HDI in Papua Province are close to their actual values.

Furthermore, to determine the predictive ability of this modeling, we use the model that has been obtained from the data in 2017 to predict the PPP and HDI values of Papua Province in 2018 and 2019. This prediction is carried out by applying the model that has been obtained to the GRDP and the

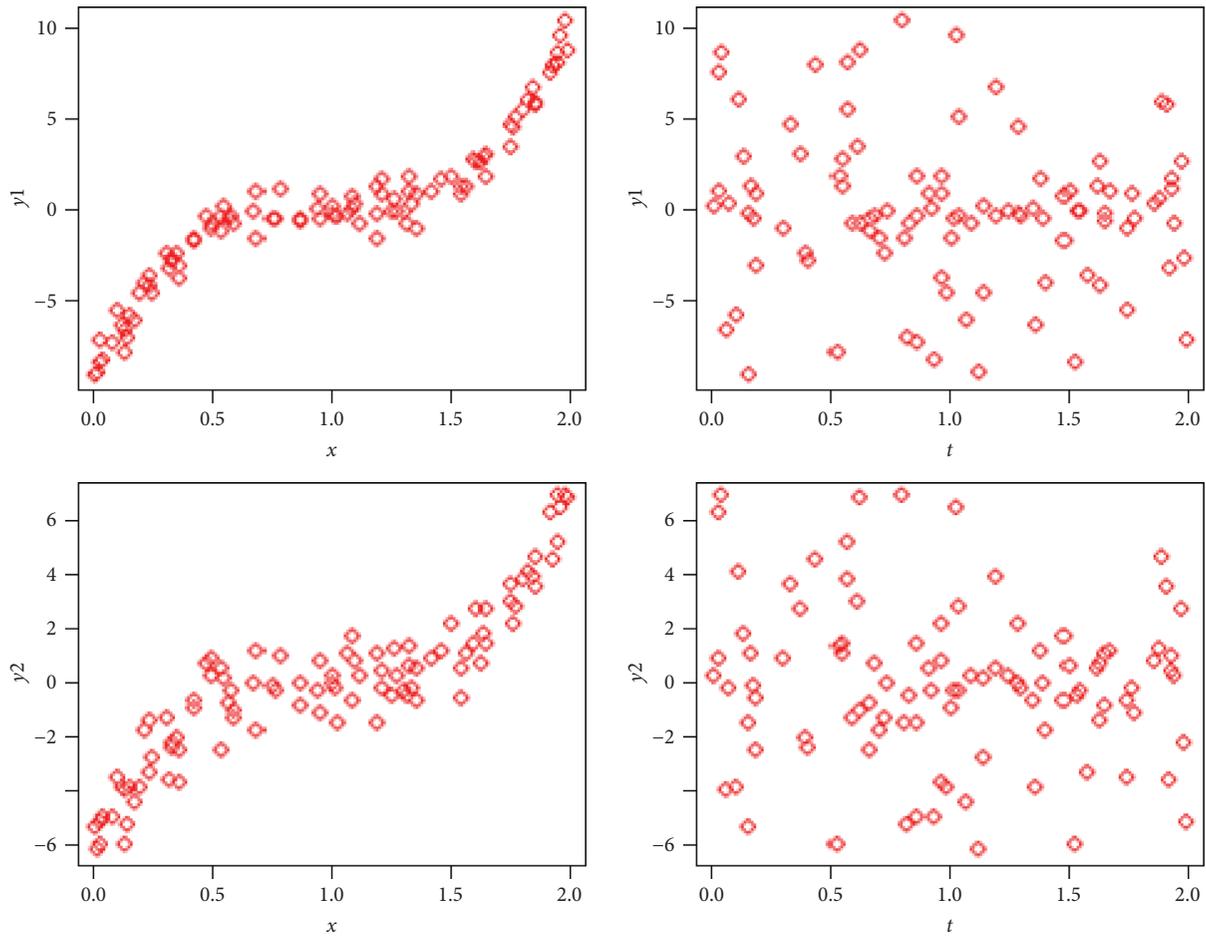


FIGURE 1: Plots of the simulation data.

TABLE 1: Combination of smoothing (λ_h) and bandwidth (α_h) parameters with the GCV value.

λ_1	λ_2	α_1	α_2	GCV
0.002413	0.000899	23.96	24.16	13.123
0.0002413	0.0000899	23.96	24.16	3.397
0.002413	0.000899	2.396	2.416	12.143
0.0002413	0.0000899	2.396	2.416	3.239
0.00002413	0.00000899	2.396	2.416	5.739
0.0002413	0.0000899	0.2396	0.2416	11494.49
0.00002413	0.00000899	0.2396	0.2416	3130.825

TABLE 2: Comparison between three estimators.

Model	Minimum GCV
Mixed spline smoothing and kernel	3.2392
Spline smoothing	4.6587
Kernel	4.4085

population growth rate of Papua Province in 2018 and 2019. One of the criteria for the predictive ability of a model that can be obtained from this prediction is the Mean Absolute

Percentage Error (MAPE) value. The MAPE value obtained from the predictions in 2018 is 3.8042% or the level of accuracy is 96.1958%, and the MAPE value from the predictions in 2019

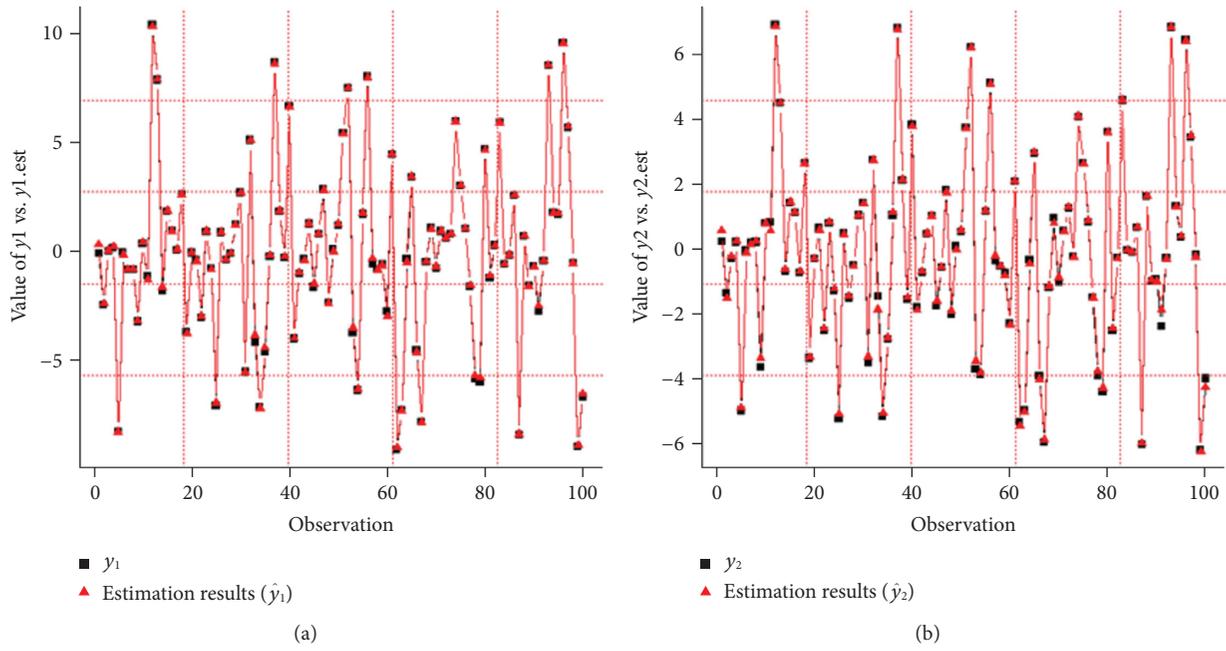


FIGURE 2: The plots between the estimation results and the original simulation data.

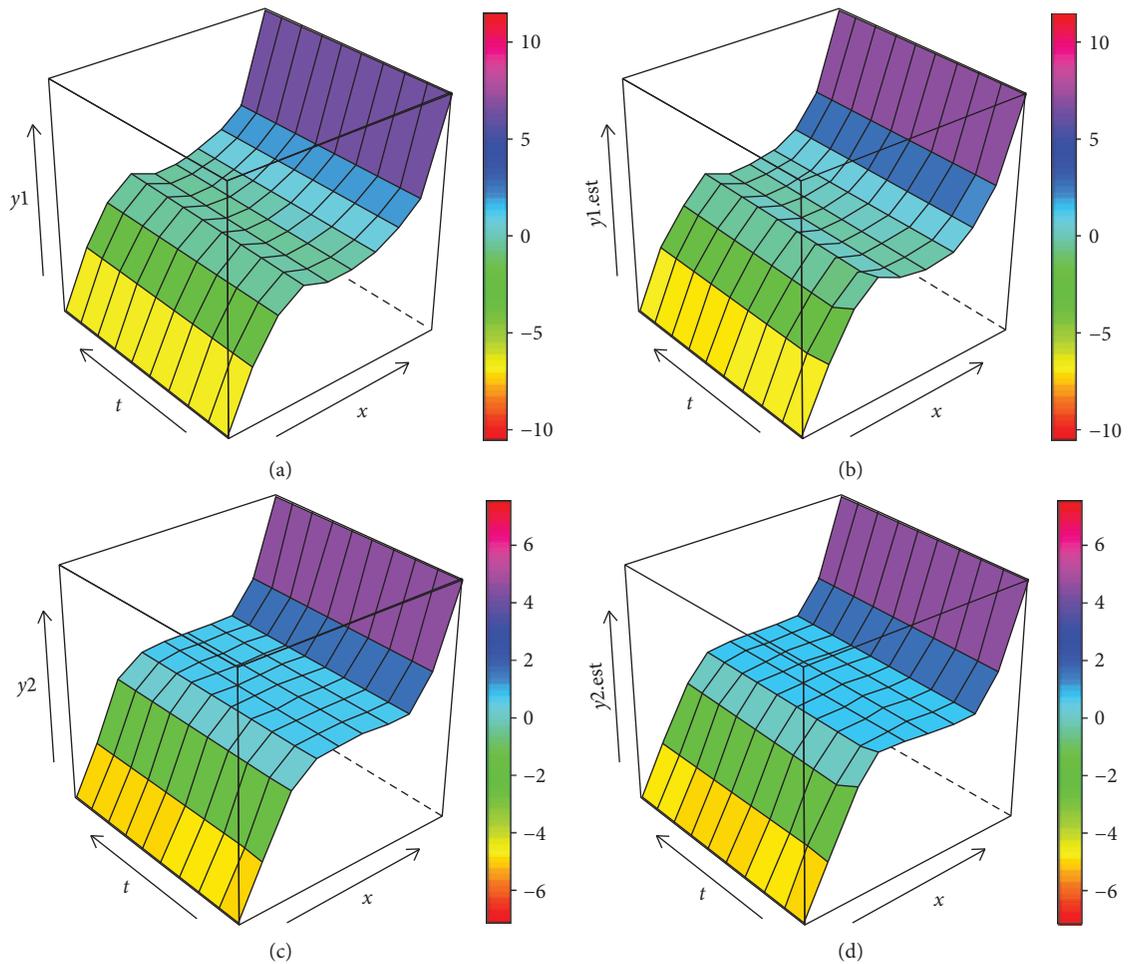


FIGURE 3: The 3D surface plots of the function for simulation and its estimator. (a) Surface plots of the simulation function for y_1 . (b) Surface plots of the estimator function for y_1 . (c) Surface plots of the simulation function for y_2 . (d) Surface plots of the estimator function for y_2 .

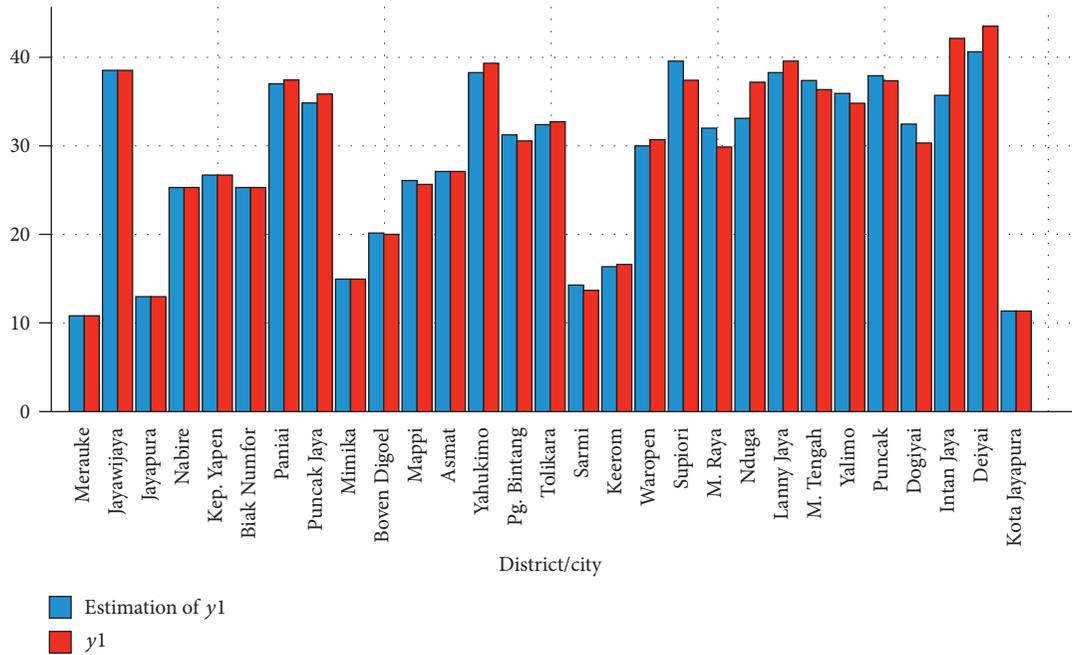


FIGURE 4: Comparison of the actual PPP and its estimation in Papua Province.

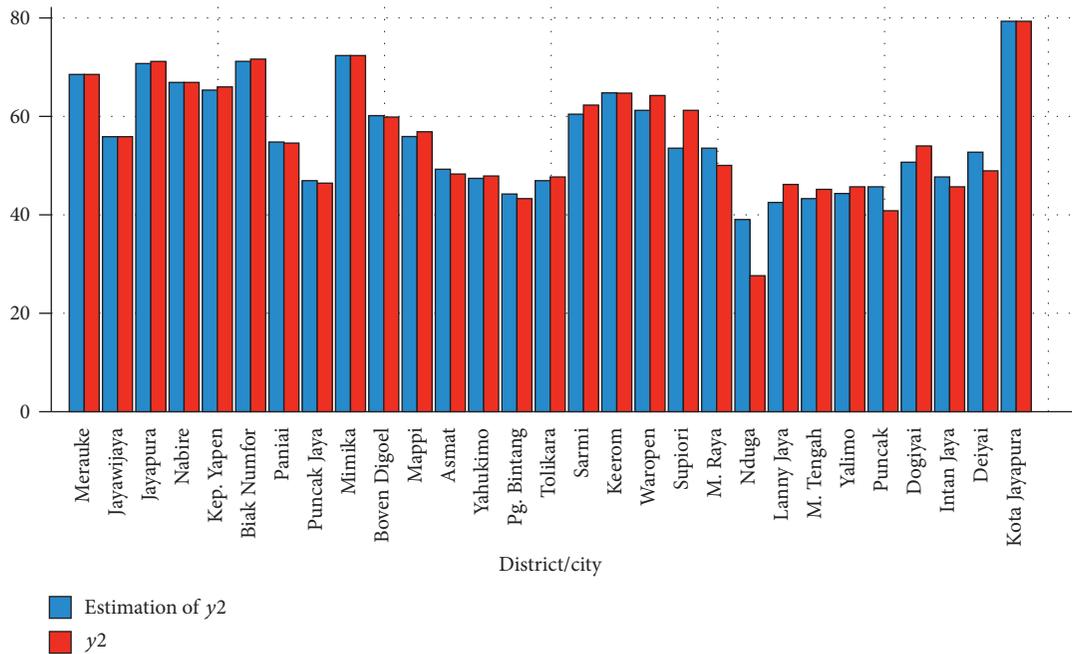


FIGURE 5: Comparison of the actual HDI and its estimation in Papua Province.

is 5.1658% or the level of accuracy is 94.8342%. These MAPE values are less than 10%, which indicates that the biresponse nonparametric regression model with mixed spline smoothing and kernel estimators has a good predictive ability to predict PPP and HDI in Papua Province.

4. Conclusions

This paper presents the biresponse nonparametric regression model with mixed spline smoothing and kernel

estimators. This new mixed estimator is obtained through two-stage estimation, i.e., the first stage using the PWLS to obtain the spline smoothing component, followed by the second stage that employs the WLS to estimate the kernel component. This mixed estimator is formed to handle the different data patterns between each predictor in the biresponse case, so this estimator can provide better estimation results. Selection of the best model for the proposed estimator is carried out by selecting a model that produces a minimum GCV value. The simulation results show the

biresponse mixed spline smoothing and kernel estimator provides better results compared to the biresponse spline smoothing or biresponse kernel estimator. Furthermore, this proposed estimator can be appropriately applied to model the percentage of the poor population (PPP) and human development index (HDI) in Papua Province and gives satisfactory results. The limitation of this study is we only use one predictor variable for each component of the estimators, both spline smoothing and kernel estimators. For future work, this biresponse mixed estimator can be developed with more than one predictor for each estimator component. Apart from this limitation, this study is useful for our insight into mixed estimators in biresponse nonparametric regression.

Data Availability

The data in this article are available in the BPS of the Papua Province repository (<https://papua.bps.go.id>).

Conflicts of Interest

The authors declare that there are no conflicts of interest in this paper.

Acknowledgments

The authors thank the Ministry of Research, Technology, and Higher Education (Ristekdikti), Republic of Indonesia, for supporting this research through the PMDSU scholarship.

References

- [1] W. Härdle, *Applied Nonparametric Regression*, Vol. 19, Cambridge University Press, Cambridge UK, 1990.
- [2] I. N. Budiantara, V. Ratnasari, M. Ratna, W. Wibowo et al., "Modeling percentage of poor people in Indonesia using kernel and fourier series mixed estimator in nonparametric regression," *Investigacion Operacional*, vol. 40, no. 4, pp. 538–550, 2019.
- [3] R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, CRC Press, Boca Raton, FL, USA, 1999.
- [4] H. Becher, G. Kauermann, P. Khomski, and B. Kouyaté, "Using penalized splines to model age- and season-of-birth-dependent effects of childhood mortality risk factors in rural Burkina Faso," *Biometrical Journal*, vol. 51, no. 1, pp. 110–122, 2009.
- [5] Y. Wang, W. Guo, and M. B. Brown, "Spline smoothing for bivariate data with applications to association between hormones," *Statistica Sinica*, vol. 10, no. 2, pp. 377–397, 2000.
- [6] P. Hall and L. S. Huang, "Nonparametric kernel regression subject to monotonicity constraints," *Annals of Statistics*, vol. 29, no. 3, pp. 624–647, 2001.
- [7] H. Okumura and K. Naito, "Non-parametric kernel regression for multinomial data," *Journal of Multivariate Analysis*, vol. 97, no. 9, pp. 2009–2022, 2006.
- [8] P. Du, C. F. Parmeter, and J. S. Racine, "Nonparametric kernel regression with multiple predictors and multiple shape constraints," *Statistica Sinica*, vol. 23, no. 3, pp. 1347–1371, 2013.
- [9] N. Chamidah and T. Saifudin, "Estimation of children growth curve based on kernel smoothing in multi-response non-parametric regression," *Applied Mathematical Sciences*, vol. 7, no. 37–40, pp. 1839–1847, 2013.
- [10] E. Erçelik and M. Nadar, "Nonparametric density estimation based on beta prime kernel," *Communications in Statistics-Theory and Methods*, vol. 49, no. 2, pp. 325–342, 2020.
- [11] M. Bilodeau, "Fourier smoother and additive models," *Canadian Journal of Statistics*, vol. 20, no. 3, pp. 257–269, 1992.
- [12] U. Amato, A. Antoniadis, and I. De Feis, "Fourier series approximation of separable models," *Journal of Computational and Applied Mathematics*, vol. 146, no. 2, pp. 459–479, 2002.
- [13] R. Hidayat, I. N. Budiantara, B. W. Otok, and V. Ratnasari, "The regression curve estimation by using mixed smoothing spline and kernel (MsS-K) model," *Communications in Statistics-Theory and Methods*, pp. 1–12, 2020.
- [14] N. P. A. M. Mariati, I. N. Budiantara, and V. Ratnasari, "Combination estimation of smoothing spline and fourier series in nonparametric regression," *Journal of Mathematics*, vol. 2020, 10 pages, 2020, <https://doi.org/10.1155/2020/4712531>, Article ID 4712531.
- [15] M. A. D. Octavanny, I. N. Budiantara, H. Kuswanto, and D. P. Rahmawati, "Nonparametric regression model for longitudinal data with mixed truncated spline and fourier series," In *Abstract and Applied Analysis*, vol. 2020, Article ID 4710745, 11 pages, 2020.
- [16] P. J. Green and B. S. Yandell, *Semi-parametric Generalized Linear Models*, pp. 44–55, Springer, Berlin, Germany, 1985.
- [17] M. Roozbeh and M. Arashi, "New ridge regression estimator in semiparametric regression models," *Communications in Statistics-Simulation and Computation*, vol. 45, no. 10, pp. 3683–3715, 2016.
- [18] M. Roozbeh and M. Najarian, "Efficiency of the QR class estimator in semiparametric regression models to combat multicollinearity," *Journal of Statistical Computation and Simulation*, vol. 88, no. 9, pp. 1804–1825, 2018.
- [19] F. Akdeniz and M. Roozbeh, "Efficiency of the generalized-difference-based weighted mixed almost unbiased two-parameter estimator in partially linear model," *Communications in Statistics-Theory and Methods*, vol. 46, no. 24, pp. 12259–12280, 2017.
- [20] P. Craven and G. Wahba, "Smoothing noisy data with spline functions," *Numerische Mathematik*, vol. 31, no. 4, pp. 377–403, 1978.
- [21] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.
- [22] G. Wahba, *Spline Models for Observational Data*, SIAM, Philadelphia, PA, USA, 1990.
- [23] M. Amini and M. Roozbeh, "Optimal partial ridge estimation in restricted semiparametric regression models," *Journal of Multivariate Analysis*, vol. 136, pp. 26–40, 2015.
- [24] M. Roozbeh, "Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion," *Computational Statistics & Data Analysis*, vol. 117, pp. 45–61, 2018.
- [25] M. Roozbeh, G. Hesamian, and M. G. Akbari, "Ridge estimation in semi-parametric regression models under the stochastic restriction and correlated elliptically contoured errors," *Journal of Computational and Applied Mathematics*, vol. 378, p. 112940, 2020.
- [26] A. A. R. Fernandes, I. N. Budiantara, B. W. Otok, and L. Suhartono, "Spline estimator for bi-responses

- nonparametric regression model for longitudinal data,” *Applied Mathematical Sciences*, vol. 8, no. 113–116, pp. 5653–5665, 2014.
- [27] S. Madan, “Human development and poverty—a perspective across Indian states,” *Statistika*, vol. 49, no. 4, pp. 81–94, 2012.
- [28] A. Akhmad and S. Amir, “The effect of economic growth and income inequality on poverty in Indonesia,” *IOSR Journal of Economics and Finance*, vol. 9, no. 4, pp. 20–26, 2018.
- [29] S. G. Grubaugh, “Economic growth and growth in human development,” *Applied Econometrics and International Development*, vol. 15, no. 2, pp. 5–16, 2015.
- [30] S. Mallick and N. Ghani, “A review of the relationship between poverty, population growth and environment,” *Pakistan Development Review*, vol. 44, no. 4, pp. 597–613, 2005.
- [31] K. P. Sinaga and O. D. Tulus, “Analisis Pengaruh Produk Domestik Regional Bruto, Pendidikan dan Pengangguran terhadap Kemiskinan di Kab/Kota Propinsi Sumatera Utara Tahun 2010-2011,” *Saintia Matematika*, vol. 1, no. 3, pp. 261–271, 2013.
- [32] T. R. Malthus, “Review: an essay on the principle of population,” *Forum for Modern Language Studies*, vol. 41, no. 1, pp. 114–115, 2005.
- [33] M. K. S. Budhi, “Analisis faktor-faktor yang berpengaruh terhadap pengentasan kemiskinan di bali: analisis FEM data panel,” *Jurnal Ekonomi Kuantitatif Terapan*, vol. 6, no. 1, pp. 1–70, 2013.