Statistical Modeling: A New Regression Curve Approximation using Mixed Estimators Truncated Spline and Epanechnikov Kernel

Sifriyani, Andrea Tri Rian Dani, Meirinda Fauziyah, and I Nyoman Budiantara

***Abstract*—In the era of computation, researchers have paid significant attention to the nonparametric regression method. Nonparametric regression has the benefit of a high degree of modeling flexibility. Developing a mixed estimator truncated spline-Epanechnikov kernel is the most recent innovation in this study. The truncated spline estimator excels at handling data whose behavior varies at predetermined intervals. In contrast, the Epanechnikov kernel estimator has a more flexible structure and excels at modeling data that does not adhere to a particular pattern. Maximum Likelihood Estimation is utilized to estimate parameters. The concluding section of this study will discuss the estimator properties of the mixed estimators truncated spline and Epanechnikov kernel models. The proposed can be utilized for additional analysis in the field of nonparametric regression.**

***Index Terms*—Nonparametric Regression, Truncated Spline, Epanechnikov Kernel, Maximum Likelihood Estimation, Mixed Estimators**

I. Introduction

NONPARAMETRIC regression is a statistical method to determine the relationship pattern between predictor variables and response variables for which the form of the relationship pattern is unknown and there is no past information [1]–[3]. This differs from parametric regression, where the relationship pattern between the predictor and response variables is known [4].

The nonparametric regression approach has received much attention from researchers because it has high flexibility in modeling. In the computational era, the relationship pattern is less rigid. Nowadays, many random, non-patterned relationship patterns are found, so the parametric regression approach is unsuitable [5].

In their study, Budiantara et al. [6] outlined how the nonparametric regression method makes it possible to adjustments to the estimated regression curve. As time goes by, approaches to nonparametric regression continue to evolve. Several estimators have been developed, including Spline [2], [7]–[9], Kernel [10]–[12], Fourier series [13]–[15], Local Polynomial [16], [17], Wavelet [18], [19], and so on. The estimators studied and developed are limited to using only one form of the estimator in the modeling process [20].

However, no one guarantees that each predictor variable has the same relationship pattern and is appropriate if modeled with only one form of an estimator. The development of mixed estimator forms is interesting and needs to be studied. The essential thinking of developing a mixed estimator form is to approach the regression curve based on each of its characteristics. Budiantara et al. [20] developed a mixed spline and kernel estimator, but this study was limited to using only one predictor for each spline and kernel component. Ratnasari et al. [5] conducted a simulation study to develop mixed truncated spline and Gaussian kernel estimators. Still, this research was limited to using only one predictor for each component. Dani et al. [21] developed a mixed truncated spline and Gaussian kernel estimator that is applied to poverty data on the island of Borneo, Indonesia. The results showed that the mixed estimator form has better accuracy than the single estimator.

The innovation that will be carried out in this study is to develop a mixed estimator form with different combinations. Based on previous research, studying the estimator form of a mixed truncated spline and Epanechnikov kernels is interesting. The two components in the mixed estimator that will be studied have their advantages and characteristics, which are expected to be frequently found in modeling cases. The truncated spline estimator can handle data whose behavior changes at certain intervals. Eubank [22] explained that estimating the regression curve with a truncated spline using the knot point can handle data pattern problems that show sharp fluctuations in data pattern changes [23]. The number of knot points and the location of the optimal knot points have a significant impact on the truncated spline estimator. [24]. The Epanechnikov kernel estimator is also better at modeling data that don't follow a certain pattern and has a more adaptable shape. The Epanechnikov kernel estimator depends on determining the optimal bandwidth [25]. The

Manuscript received January 17, 2023; revised 14 August 2023. This work was supported in part by the funding of the Ministry of Education, Culture, Research, and Technology (Kemendikbudristek) and Mulawarman University in 2022.

Sifriyani is a Lecturer at the Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia. (e-mail: sifriyani@fmipa.unmul.ac.id).

Andrea Tri Rian Dani is a Lecturer at the Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia. (e-mail: andreatririandani@fmipa.unmul.ac.id).

Meirinda Fauziyah is a Lecturer at the Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia. (e-mail: meirindafauziyah@fmipa.unmul.ac.id).

I Nyoman Budiantara is a Professor in Statistics Department, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. (e-mail: nyomanbudiantara65@gmail.com).

smoothness of the regression curve estimation results with the Epanechnikov kernel estimator is set to the bandwidth value.

Based on the description above, this article will develop a mixed estimator with a combination of different estimators. The focus of this article is to conduct a theoretical study to obtain a form of estimation of the nonparametric regression curve, an additive model between the truncated spline estimator and the Epanechnikov kernel obtained by optimizing the Maximum Likelihood Estimation (MLE).

II. PRELIMINARIES

A. Regression Analysis

Regression is a statistical technique for determining the pattern of relationships [26]. The general form of the regression model is:

$$y_i = m(x_i) + \varepsilon_i \tag{1}$$

With:

y_i : response variable

x_i : predictor variable

$m(x_i)$: regression curve to be approached

ε_i : error random $\varepsilon_i \sim N(0,\sigma^2)$.

$m(x_i)$ will be approached with a nonparametric regression technique, the Mixed Estimator model.

B. Nonparametric Regression Mixed Estimators

The mixed estimator form is a multi-predictor nonparametric regression model ($x > 1$) whose regression curve is additive. The regression curve will be approximated by two or more types of estimators [20]. For example, given paired data (x_i, v_i, y_i) where the relationship between predictor (x_i, v_i) and response (y_i) variables follow a nonparametric regression model.

$$y_i = \mu(x_i, v_i) + \varepsilon_i \tag{2}$$

Where $i = 1, 2, ..., n$.

The regression curve from $\mu(x_i, v_i)$ is assumed to be unknown, smooth in the sense of continuous, and differential. Error random ε_i follows a normal distribution $\varepsilon_i \sim N(0,\sigma^2)$. The regression curve $\mu(x_i, v_i)$ is also assumed to be additive such that $\mu(x_i, v_i)$ can be written:

$$\mu(x_i, v_i) = f(x_i) + h(v_i) \tag{3}$$

with $f(x_i)$ and $h(v_i)$ are smooth functions.

C. Additive Model

In some modeling cases, especially regression modeling, researchers often assume that they follow an additive model. The additive model is a model of the response variable that depends on the sum of the functions of the predictor variables.

$$y_i = \sum_{i=1}^{n}\sum_{j=1}^{p} f_j(x_{ji}) + \varepsilon_i \tag{4}$$

III. MAIN RESULTS

A. The Shape of Estimator of Mixed Truncated Spline and Epanechnikov Kernel Model

Given paired data $\left(x_{1i}, x_{2i}, \ldots, x_{pi}, v_{1i}, v_{2i}, \ldots, v_{qi}, y_i\right)$ and the relationship between predictor variables $x_{1i}, x_{2i}, \ldots, x_{pi}, v_{1i}, v_{2i}, \ldots, v_{qi}$ and response variables y_i follows a nonparametric regression model.

$$y_i = \mu\left(x_{1i}, x_{2i}, \ldots, x_{pi}, v_{1i}, v_{2i}, \ldots, v_{qi}\right) + \varepsilon_i \tag{5}$$

The shape of the regression curve $\mu\left(x_{1i}, x_{2i}, \ldots, x_{pi}, v_{1i}, v_{2i}, \ldots, v_{qi}\right)$ is assumed to be unknown and only assumed to be smooth, which means it is continuous and differentiable. Error random ε_i follows a normal distribution $\varepsilon_i \sim N(0,\sigma^2)$.

The regression curve is assumed to be additive, so it can be written:

$$\begin{aligned}\mu\left(x_{1i}, \ldots, x_{pi}, v_{1i}, \ldots, v_{qi}\right) = f_1\left(x_{1i}\right) + \ldots + f_p\left(x_{pi}\right) + \\ h_1\left(v_{1i}\right) + \ldots + h_q\left(v_{qi}\right)\end{aligned} \tag{6}$$

Equation (6) can be written as follows:

$$\mu(x,t) = \sum_{s=1}^{p} f_s\left(x_{si}\right) + \sum_{k=1}^{q} h_k\left(v_{ki}\right), \tag{7}$$

It is known that $f_s\left(x_{si}\right)$ and $h_k\left(v_{ki}\right)$ is a smooth function. Based on Equation (7), the main problem in the nonparametric regression curve mixed estimator is to get the regression curve estimation form $\mu(x,t)$.

$$\mu(x,t) = \sum_{s=1}^{p} f_s\left(x, v\right) + \sum_{k=1}^{q} \hat{h}_k\left(v\right) \tag{8}$$

The regression curve $f_s\left(x_{si}\right)$ will be approximated by the truncated spline of degree m_s and knot points $\eta_s = \left(\eta_{s1}, \eta_{s2}, \ldots, \eta_{sr}\right)^T$. The regression curve $h_k\left(v_{ki}\right)$ will then be reached by the Epanechnikov kernel.

Suppose given the basis for the sample space, Spline $1, x, , \ldots, x^{m_s}, (x-\eta_{s1})^{m_s} I(x \ge \eta_{s1}), \ldots, (x-\eta_{sr})^{m_s} I(x \ge \eta_{sr})$ with *I* is an Indicator function. So, the regression curve from $f_s\left(x_{si}\right)$ can be written:

$$\begin{aligned}f_s(x_{si}) = \xi_{s1} + \xi_{s1}x_{si} + \xi_{s2}x_{si}^2 + \ldots + \xi_{sm_s}x_{si}^{m_s} + \\ \varphi_{s1}(x_{si}-\eta_{s1})_+^{m_s} + \ldots + \varphi_{sr}(x_{si}-\eta_{sr})_+^{m_s},\end{aligned} \tag{9}$$

With $\xi_{s0}, \xi_{s1}, \ldots, \xi_{sm_s}, \varphi_{s1}, \ldots, \varphi_{sr}$ are the unknown regression model parameters.

Furthermore, the estimated regression curve of $h_k\left(v_{ki}\right)$ can be written as follows:

$$\hat{h}_{\alpha_k}(v) = n^{-1}\sum_{i=1}^{n}\left(\frac{T_{\alpha_k}(v-v_i)}{n^{-1}\sum_{i=1}^{n}T_{\alpha_k}(v-v_i)}\right)y_i = n^{-1}\sum_{i=1}^{n}R_{\alpha_k i}(v)y_i, \quad (10)$$

With:

$$R_{\alpha_k i}(v) = \frac{T_{\alpha_k}(v-v_i)}{n^{-1}\sum_{i=1}^{n}T_{\alpha_k}(v-v_i)}; T_{\alpha_k}(v-v_i) = \frac{1}{\alpha_k}T\left(\frac{(v-v_i)}{\alpha_k}\right)$$

T is a kernel function. The kernel function used in this study is the Epanechnikov kernel, which has the formula:

$$T(z) = \frac{3}{4}(1-z^2);\ I_{[-1,1]}(z) \quad (11)$$

with $z = \dfrac{v-v_i}{\alpha}$, α is bandwidth, v is a value determined from the predictor variable, and v_i is the i-th value of the predictor variable. The regression curve estimator in Equation (10) depends on two things, namely, the bandwidth parameter and the kernel function. The form of the mixed estimators truncated spline and Epanechnikov kernel model will be searched in Equation (8) using the Maximum Likelihood Estimation (MLE) method based on Equation (9).

Lemma 1

The regression curve $f_s(x_{si})$ given in Equation (8), then:

$$\mathbf{f} = \mathbf{X}_1\boldsymbol{\xi} + \mathbf{X}_2\boldsymbol{\varphi} \quad (12)$$

Where vector from $\mathbf{f}$, $\boldsymbol{\xi}$, and $\boldsymbol{\varphi}$ are given by:

$$\mathbf{f} = \left(\mathbf{f}_1\ \mathbf{f}_2\ \dots\ \mathbf{f}_p\right)^T; \mathbf{f}_i = \left(f_i(x_{i1})\ f_i(x_{i2})\ \dots\ f_i(x_{in})\right)^T$$

$$\boldsymbol{\xi} = \left(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \dots\ \boldsymbol{\xi}_p\right)^T, \boldsymbol{\xi}_j = \left(\xi_{j0}\ \xi_{j1}\ \dots\ \xi_{jm_j}\right)^T$$

$$\boldsymbol{\varphi} = \left(\boldsymbol{\varphi}_1\ \boldsymbol{\varphi}_2\ \dots\ \boldsymbol{\varphi}_p\right)^T, \boldsymbol{\varphi}_l = \left(\varphi_{l0}\ \varphi_{l1}\ \dots\ \varphi_{lm_l}\right)^T$$

And matrix $\mathbf{X}_1$, $\mathbf{X}_2(\varphi)$ are given by:

$$\mathbf{X}_1 = \begin{bmatrix} 1 & x_{11} & \cdots & x_{11}^{m_1} & 1 & \cdots & x_{21}^{m_2} & 1 & \cdots & x_{p1}^{m_p} \\ 1 & x_{12} & \cdots & x_{12}^{m_1} & 1 & \cdots & x_{22}^{m_2} & 1 & \cdots & x_{p2}^{m_p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{1n}^{m_1} & 1 & \cdots & x_{2n}^{m_2} & 1 & \cdots & x_{pn}^{m_p} \end{bmatrix}$$

$$\mathbf{X}_2(\varphi) = \begin{bmatrix} (x_{11}-\eta_{11})_+^{m_1} & \cdots & (x_{21}-\eta_{21})_+^{m_2} & \cdots & (x_{p1}-\eta_{pr_p})_+^{m_p} \\ (x_{12}-\eta_{11})_+^{m_1} & \cdots & (x_{22}-\eta_{21})_+^{m_2} & \cdots & (x_{p2}-\eta_{pr_p})_+^{m_p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ (x_{1n}-\eta_{11})_+^{m_1} & \cdots & (x_{2n}-\eta_{21})_+^{m_2} & \cdots & (x_{pn}-\eta_{pr_p})_+^{m_p} \end{bmatrix}$$

Proof

Based on Equation (9)

$$f_s(x_{si}) = \xi_{s0} + \xi_{s1}x_{si} + \dots + \xi_{sm_s}x_{si}^{m_s} + = \varphi_{s1}(x_{si}-\eta_{s1})_+^{m_s} + \dots + \varphi_{sr_s}(x_{si}-\eta_{sr_s})_+^{m_s}$$

for s = 1, obtained

$$f_1(x_{1i}) = \xi_{10} + \xi_{11}x_{1i} + \dots + \xi_{1m_1}x_{1i}^{m_1} + = \varphi_{11}(x_{1i}-\eta_{11})_+^{m_1} + \dots + \varphi_{1r_1}(x_{1i}-\eta_{1r_1})_+^{m_1}$$

If it is explained for $i = 1, 2, \dots, n$:

$$f_1(x_{11}) = \xi_{10} + \xi_{11}x_{11} + \dots + \xi_{1m_1}x_{11}^{m_1} + = \varphi_{11}(x_{11}-\eta_{11})_+^{m_1} + \dots + \varphi_{1r}(x_{11}-\eta_{1r_1})_+^{m_1}$$

$$\vdots$$

$$f_1(x_{1n}) = \xi_{10} + \xi_{11}x_{1n} + \dots + \xi_{1m_1}x_{1n}^{m_1} + = \varphi_{11}(x_{1n}-\eta_{11})_+^{m_1} + \dots + \varphi_{1r}(x_{1n}-\eta_{1r_1})_+^{m_1}$$

The regression model represented by Equation (9) can be expressed in matrix form as:

$$\begin{pmatrix} f_1(x_{11}) \\ f_1(x_{12}) \\ \vdots \\ f_1(x_{1n}) \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{11}^{m_1} \\ 1 & x_{12} & \cdots & x_{12}^{m_1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & \cdots & x_{1n}^{m_1} \end{pmatrix}\begin{pmatrix} \xi_{10} \\ \xi_{11} \\ \vdots \\ \xi_{1m_1} \end{pmatrix} + = \begin{pmatrix} (x_{11}-\eta_{11})_+^{m_1} & \cdots & (x_{11}-\eta_{1r_1})_+^{m_1} \\ (x_{12}-\eta_{11})_+^{m_1} & \cdots & (x_{12}-\eta_{1r_1})_+^{m_1} \\ \vdots & \vdots & \vdots \\ (x_{1n}-\eta_{11})_+^{m_1} & \cdots & (x_{1n}-\eta_{1r_1})_+^{m_1} \end{pmatrix}\begin{pmatrix} \varphi_{11} \\ \varphi_{12} \\ \vdots \\ \varphi_{1n} \end{pmatrix}$$

Then, $\mathbf{f}_1 = \mathbf{X}_{11}\boldsymbol{\xi}_1 + \mathbf{X}_{12}(\eta_1)\boldsymbol{\varphi}_1$.

For s = 2, 3, …, p, do the same thing, so we get:

$$s = 2, \mathbf{f}_2 = \mathbf{X}_{21}\boldsymbol{\xi}_2 + \mathbf{X}_{22}(\eta_2)\boldsymbol{\varphi}_2,$$

$$s = 3, \mathbf{f}_3 = \mathbf{X}_{31}\boldsymbol{\xi}_3 + \mathbf{X}_{32}(\eta_3)\boldsymbol{\varphi}_3$$

until

$$s = p, \mathbf{f}_p = \mathbf{X}_{p1}\boldsymbol{\xi}_p + \mathbf{X}_{p2}(\eta_p)\boldsymbol{\varphi}_p.$$

We get:

$$\begin{aligned} \mathbf{f} &= \sum_{s=1}^{p}\mathbf{f}_s \\ &= \mathbf{f}_1 + \mathbf{f}_2 + \dots + \mathbf{f}_p \\ &= \mathbf{X}_{11}\boldsymbol{\xi}_1 + \mathbf{X}_{12}(\eta_1)\boldsymbol{\varphi}_1 + \dots + \\ &\quad \mathbf{X}_{p1}\boldsymbol{\xi}_p + \mathbf{X}_{p2}(\eta_p)\boldsymbol{\varphi}_p \\ &= \mathbf{X}_{11}\boldsymbol{\xi}_1 + \dots + \mathbf{X}_{p1}\boldsymbol{\xi}_p + \mathbf{X}_{12}(\eta_1)\boldsymbol{\varphi}_1 + \\ &\quad \dots + \mathbf{X}_{p2}(\eta_p)\boldsymbol{\varphi}_p \end{aligned}$$

$$= \begin{pmatrix} \mathbf{X}_{11} & \mathbf{X}_{21} & \cdots & \mathbf{X}_{p1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \vdots \\ \boldsymbol{\xi}_p \end{pmatrix} +$$

$$= \begin{pmatrix} \mathbf{X}_{12}(\eta_1) & \mathbf{X}_{22}(\eta_2) & \cdots & \mathbf{X}_{p2}(\eta_p) \end{pmatrix} \begin{pmatrix} \boldsymbol{\varphi}_1 \\ \boldsymbol{\varphi}_2 \\ \vdots \\ \boldsymbol{\varphi}_p \end{pmatrix}$$

$$= \mathbf{X}_1\boldsymbol{\xi} + \mathbf{X}_2(\eta)\boldsymbol{\varphi}$$

With vectors $\boldsymbol{\xi}$, $\boldsymbol{\varphi}$, and matrix $\mathbf{X}_1, \mathbf{X}_2(\eta)$ are given by Equation (12).

Lemma 2

When the estimator for Epanechnikov kernel regression is given by Equation (10), and the regression model is given by Equation (8), then the sum squares of the errors is:

$$\|\boldsymbol{\varepsilon}\|^2 = \left\| \left[\mathbf{I} - \mathbf{M}(\alpha)\right]\mathbf{y} - \mathbf{X}(\eta)\boldsymbol{\Phi} \right\|^2, \qquad (13)$$

Where $\|\boldsymbol{\varepsilon}\|^2$ is a length vector from $\boldsymbol{\varepsilon}$.

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}; \ \mathbf{X}(\eta) = \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2(\eta) \end{pmatrix}, \boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\varphi} \end{pmatrix}, \text{ and}$$

$$\mathbf{M}(\alpha) = \begin{pmatrix} \sum_{k=1}^{q} R_{\alpha_k 1}(v_1) & \sum_{k=1}^{q} R_{\alpha_k 2}(v_1) & \cdots & \sum_{k=1}^{q} R_{\alpha_k n}(v_1) \\ \sum_{k=1}^{q} R_{\alpha_k 1}(v_2) & \sum_{k=1}^{q} R_{\alpha_k 2}(v_2) & \cdots & \sum_{k=1}^{q} R_{\alpha_k n}(v_2) \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{k=1}^{q} R_{\alpha_k 1}(v_n) & \sum_{k=1}^{q} R_{\alpha_k 2}(v_n) & \cdots & \sum_{k=1}^{q} R_{\alpha_k n}(v_n) \end{pmatrix}$$

Proof

Based on Equation (10)

$\hat{h}_{\alpha_k}(v_i) = n^{-1}\sum_{i=1}^{n} R_{\alpha_k}(v_i) y_i$ for each $k = 1, 2, \ldots, q$.

For $k = 1$, get $\hat{h}_{\alpha_1}(v_i) = n^{-1}\sum_{i=1}^{n} R_{\alpha_{1i}}(v_i) y_i$. Since it applies to $i = 1, 2, \ldots n,$ then:

$$\hat{h}_{\alpha_1}(v_1) = R_{\alpha_1 1}(v_1)y_1 + R_{\alpha_1 2}(v_1)y_2 + \ldots + R_{\alpha_1 n}(v_1)y_n$$
$$\hat{h}_{\alpha_1}(v_2) = R_{\alpha_1 1}(v_2)y_1 + R_{\alpha_1 2}(v_2)y_2 + \ldots + R_{\alpha_1 n}(v_2)y_n$$
$$\vdots$$
$$\hat{h}_{\alpha_1}(v_n) = R_{\alpha_1 1}(v_n)y_1 + R_{\alpha_1 2}(v_n)y_2 + \ldots + R_{\alpha_1 n}(v_n)y_n$$

We obtain in matrix form:

$$\begin{pmatrix} \hat{h}_{\alpha_1}(v_1) \\ \hat{h}_{\alpha_1}(v_2) \\ \vdots \\ \hat{h}_{\alpha_1}(v_n) \end{pmatrix} = \begin{pmatrix} R_{\alpha_1 1}(v_1) & R_{\alpha_1 2}(v_1) & \cdots & R_{\alpha_1 n}(v_1) \\ R_{\alpha_1 1}(v_2) & R_{\alpha_1 2}(v_2) & \cdots & R_{\alpha_1 n}(v_2) \\ \vdots & \vdots & \vdots & \vdots \\ R_{\alpha_1 1}(v_n) & R_{\alpha_1 2}(v_n) & \cdots & R_{\alpha_1 n}(v_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$\hat{h}_1 = \mathbf{M}(\alpha_1)\mathbf{y} \qquad (14)$$

For k = 2, 3, …, q, do the same thing, so we get: $k = 2,\ h_2 = \mathbf{M}(\alpha_2)\mathbf{y}$ until $k = q, \hat{h}_q = \mathbf{M}(\alpha_q)\mathbf{y}$. So we get:

$$\mathbf{M}(\alpha) = \begin{pmatrix} \sum_{k=1}^{q} R_{\alpha_k 1}(v_1) & \sum_{k=1}^{q} R_{\alpha_k 2}(v_1) & \cdots & \sum_{k=1}^{q} R_{\alpha_k n}(v_1) \\ \sum_{k=1}^{q} R_{\alpha_k 1}(v_2) & \sum_{k=1}^{q} R_{\alpha_k 2}(v_2) & \cdots & \sum_{k=1}^{q} R_{\alpha_k n}(v_2) \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{k=1}^{q} R_{\alpha_k 1}(v_n) & \sum_{k=1}^{q} R_{\alpha_k 2}(v_n) & \cdots & \sum_{k=1}^{q} R_{\alpha_k n}(v_n) \end{pmatrix}$$

The regression model in Equation (8), Equation (14), and based on Lemma 1, gives:

$$\begin{aligned} \mathbf{y} &= \sum_{s=1}^{p} \mathbf{f}_s + \sum_{k=1}^{q} \mathbf{h}_k + \boldsymbol{\varepsilon} \\ &= \mathbf{X}_1\boldsymbol{\xi} + \mathbf{X}_2(\eta)\boldsymbol{\varphi} + \mathbf{M}(\alpha)\mathbf{y} + \boldsymbol{\varepsilon} \\ &= \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2(\eta) \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\varphi} \end{pmatrix} + \mathbf{M}(\alpha)\mathbf{y} + \boldsymbol{\varepsilon} \\ &= \mathbf{X}(\eta)\boldsymbol{\Phi} + \mathbf{M}(\alpha)\mathbf{y} + \varepsilon \end{aligned} \qquad (15)$$

Equation (15) gives the sum squares of the errors:

$$\begin{aligned} \|\boldsymbol{\varepsilon}\|^2 &= \|\mathbf{y} - \mathbf{X}(\eta)\boldsymbol{\Phi} - \mathbf{M}(\alpha)\mathbf{y}\|^2 \\ &= \left\| \left[\mathbf{I} - \mathbf{M}(\alpha)\right]\mathbf{y} - \mathbf{X}(\eta)\boldsymbol{\Phi} \right\|^2 \end{aligned} \qquad (16)$$

Theorem 1

Based on Equation (16) it is obtained that the sum of the squared errors of the mixed estimators truncated spline and Epanechnikov kernel. The error has a Normal multivariate with a mean of zero and $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T) = \sigma^2\mathbf{I}$. $L(\boldsymbol{\Phi}, \sigma^2 | \eta, \alpha)$ is a Likelihood function, so by using the Maximum Likelihood Estimation (MLE) estimator for the parameter vector $\boldsymbol{\Phi}$ obtained by optimization.

$$Max\left\{L(\boldsymbol{\Phi}, \sigma^2 | \eta, \alpha\right\} = Min\left\{ \left\| \begin{matrix} \left[\mathbf{I} - \mathbf{M}(\alpha)\right]\mathbf{y} - \\ \mathbf{X}(\eta)\boldsymbol{\Phi} \end{matrix} \right\|^2 \right\} \qquad (17)$$

Proof

A mixed estimators truncated spline and Epanechnikov kernel model is presented as in Equation (8), so the Likelihood function $L(\boldsymbol{\Phi}, \sigma^2 | \eta, \alpha)$ will be given by:

$$L(\mathbf{\Phi},\sigma^2|\eta,\alpha)=\prod_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi\sigma^2}}exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right)\right)$$
$$=\left(2\pi\sigma^2\right)^{-n/2}exp\left(-\frac{1}{2\sigma^2}\|\boldsymbol{\varepsilon}\|^2\right)$$

Based on Lemma 2, the Likelihood function is obtained:

$$L(.)=\left(2\pi\sigma^2\right)^{-n/2}exp\left(-\frac{1}{2\sigma^2}\left\|\begin{matrix}[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}-\\ \mathbf{X}(\eta)\mathbf{\Phi}\end{matrix}\right\|^2\right) \quad (18)$$

The estimator for parameter $\mathbf{\Phi}$ is obtained using the MLE. The log-likelihood function is written:

$$l(.)=\log L(\mathbf{\Phi},\sigma^2|\eta,\alpha)$$
$$=-\frac{n}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left\|\begin{matrix}[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}-\\ \mathbf{X}(\eta)\mathbf{\Phi}\end{matrix}\right\|^2 \quad (19)$$

Equation (19) will be maximum if the component $\|[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}-\mathbf{X}(\eta)\mathbf{\Phi}\|^2$ have a minimum value.

$$max\{L(.)\}=min\left\{\|[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}-\mathbf{X}(\eta)\mathbf{\Phi}\|^2\right\} \quad (20)$$

The form of the mixed estimator truncated spline and Epanechnikov kernel model, which is given in full in Theorem 2, can be found using Lemma 1, Lemma 2, and Theorem 1.

Theorem 2

Given a mixed estimators truncated spline and Epanechnikov kernel model is presented as in Equation (8), the sum squares of errors presented in Lemma 2, and the MLE estimator for the parameter $\mathbf{\Phi}$ is obtained from the results in the optimization of Theorem 1.

$$\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)=\hat{\mathbf{f}}_{\eta,\alpha}(x,v)+\hat{h}_{\alpha}(v) \quad (21)$$

With

$$\hat{\mathbf{f}}_{\gamma,\alpha}(x,v)=\mathbf{X}(\eta)\hat{\mathbf{\Phi}}(\eta,\alpha)$$
$$\hat{h}_{\alpha}(v)=\mathbf{M}(\alpha)\mathbf{y}$$
$$\hat{\mathbf{\Phi}}(\eta,\alpha)=\left[\left(\mathbf{X}(\eta)\right)^T\mathbf{X}(\eta)\right]^{-1}\left(\mathbf{X}(\eta)\right)^T\left(\mathbf{I}\text{-}\mathbf{M}(\alpha)\right)\mathbf{y}$$

Noted, if the matrix form from $\mathbf{X}(\eta)$ and $\mathbf{M}(\alpha)$ given by Lemma 1 and 2.

Proof

The MLE estimator from the parameter vector $\hat{\mathbf{\Phi}}$:

$$Min\left\{\|[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}-\mathbf{X}(\eta)\mathbf{\Phi}\|^2\right\}=Min\left\{D\left((\mathbf{\Phi},\sigma^2|\eta,\alpha)\right)\right\}$$

With

$$D(.)=\|[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}-\mathbf{X}(\eta)\mathbf{\Phi}\|^2$$
$$=\|[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}\|^2-2\mathbf{\Phi}^T\left(\left(\mathbf{X}(\eta)\right)^T\right)[\mathbf{I}-\mathbf{M}(\alpha)]\mathbf{y}+$$
$$\mathbf{\Phi}^T\left(\mathbf{X}(\eta)\right)^T\mathbf{X}(\eta)\mathbf{\Phi}$$

Furthermore, by a partial derivative of the parameter $\mathbf{\Phi}$ and then equal to 0:

$$\frac{\partial}{\partial\mathbf{\Phi}}\left[D\left((\mathbf{\Phi},\sigma^2|\eta,\alpha)\right)\right]=\mathbf{0}$$

The estimation results of the parameters $\hat{\mathbf{\Phi}}$ using MLE are given by:

$$\hat{\mathbf{\Phi}}(\eta,\alpha)=\mathbf{Q}(\eta,\alpha)\mathbf{y} \quad (22)$$

With

$$\mathbf{Q}(\eta,\alpha)=\left[\left(\mathbf{X}(\eta)\right)^T\left(\mathbf{X}(\eta)\right)\right]^{-1}\left(\mathbf{X}(\eta)\right)^T\left(\mathbf{I}-\mathbf{M}(\alpha)\right).$$

Based on Equation (22), the truncated spline estimator given by $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)=\mathbf{X}(\eta)\hat{\mathbf{\Phi}}(\eta,\alpha)$, and for the estimation of the regression curve approximated by the Epanechnikov kernel estimator $\hat{h}_{\alpha}(v)$ is given by $\hat{h}_{\alpha}(v)=\mathbf{M}(\alpha)\mathbf{y}$.

Lemma 3

If the estimators $\mathbf{\Phi}(\eta,\alpha),\hat{\mathbf{h}}_{\alpha}(v),\hat{\mathbf{f}}_{\eta,\alpha}(x,v)$ and $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)$ are given in Theorem 2, then:

$\hat{\mathbf{f}}_{\eta,\alpha}(x,v)=\mathbf{A}(\eta,\alpha)\mathbf{y}$ and $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)=\mathbf{P}(\eta,\alpha)\mathbf{y}$.

Where:

$$\mathbf{A}(\eta,\alpha)=\mathbf{X}(\eta)\left[\left(\mathbf{X}(\eta)\right)^T\mathbf{X}(\eta)\right]^{-1}\left(\mathbf{X}(\eta)\right)^T\left(\mathbf{I}\text{-}\mathbf{M}(\alpha)\right)$$
$$\mathbf{P}(\eta,\alpha)=\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha).$$

Proof

In Theorem 2, given $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)=\mathbf{X}(\eta)\hat{\mathbf{\Phi}}(\eta,\alpha)$, so:

$$\hat{\mathbf{f}}_{\eta,\alpha}(x,v)=\mathbf{X}(\eta)\begin{bmatrix}\left[\left(\mathbf{X}(\eta)\right)^T\left(\mathbf{X}(\eta)\right)\right]^{-1}\left(\mathbf{X}(\eta)\right)^T\\ \left(\mathbf{I}-\mathbf{M}(\alpha)\right)\mathbf{y}\end{bmatrix}$$

Can be summarized as $\mathbf{A}(\eta,\alpha)\mathbf{y}$, with:

$$\mathbf{A}(\eta,\alpha)=\mathbf{X}(\eta)\left[\left(\mathbf{X}(\eta)\right)^T\mathbf{X}(\eta)\right]^{-1}\left(\mathbf{X}(\eta)\right)^T\left(\mathbf{I}\text{-}\mathbf{M}(\alpha)\right).$$

In theorem 2, given $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)=\hat{\mathbf{f}}_{\eta,\alpha}(x,v)+\hat{h}_{\alpha}(v)$, so:

$$\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)=\hat{\mathbf{f}}_{\ ,\alpha}(x,v)+\hat{h}_{\alpha}(v)$$
$$=\left(\mathbf{A}(\eta,\alpha)\mathbf{y}\right)+\left(\mathbf{M}(\alpha)\mathbf{y}\right)$$
$$=\left[\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha)\right]\mathbf{y}$$

Then $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)=\mathbf{P}(\eta,\alpha)\mathbf{y}$.

B. The Properties of the Mixed Estimator Truncated Spline and Epanechnikov Kernel Model

The next step is to investigate the properties of the estimator $\hat{\mathbf{\Phi}}(\eta,\alpha)$, Epanechnikov kernel estimator $\hat{h}_{\alpha}(v)$, truncated spline estimator $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)$, and mixed estimator $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)$.

Theorem 3

Given a mixed estimators truncated spline and Epanechnikov kernel model is presented as in Equation (8), and it is known if $\hat{\boldsymbol{\Phi}}(\eta,\alpha)$, $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)$, $\hat{h}_{\alpha}(v)$, and $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)$ is an estimator given in Theorem 2, it can be seen that:

a) $\hat{\boldsymbol{\Phi}}(\eta,\alpha)$, $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)$, $\hat{h}_{\alpha}(v)$, and $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)$ are estimators that are biased for $\boldsymbol{\Phi}$, $\mathbf{f}_{\eta,\alpha}(x,v)$, $\mathbf{h}_{\alpha}(v)$, and $\boldsymbol{\mu}_{\eta,\alpha}(x,v)$.

b) $\hat{\boldsymbol{\Phi}}(\eta,\alpha)$, $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)$, $\hat{h}_{\alpha}(v)$, and $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)$ are linear estimators in observations $\mathbf{y}$.

Proof

a) In Theorem 2 and Lemma 3, given:

$\hat{\boldsymbol{\Phi}}(\eta,\alpha)=\left[(\mathbf{X}(\eta))^T(\mathbf{X}(\eta))\right]^{-1}(\mathbf{X}(\eta))^T(\mathbf{I}-\mathbf{M}(\alpha))\mathbf{y}$,

$\hat{\mathbf{f}}_{\eta,\alpha}(x,v)=\mathbf{A}(\eta,\alpha)\mathbf{y}$,

$\hat{h}_{\alpha}(v)=\mathbf{M}(\alpha)\mathbf{y}$, and

$\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)=\left[\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha)\right]\mathbf{y}$.

Then,

$$
\begin{aligned}
\mathrm{E}\left[\hat{\boldsymbol{\Phi}}(\eta,\alpha)\right] &= E\left(\left[(\mathbf{X}(\eta))^T(\mathbf{X}(\eta))\right]^{-1}(\mathbf{X}(\eta))^T(\mathbf{I}-\mathbf{M}(\alpha))\mathbf{y}\right)\\
&= \left[(\mathbf{X}(\eta))^T(\mathbf{X}(\eta))\right]^{-1}(\mathbf{X}(\eta))^T(\mathbf{I}-\mathbf{M}(\alpha))\\
&\quad \left[\mathbf{X}(\eta)\boldsymbol{\Phi}+\sum_{k=1}^{q}\mathbf{g}_k(t)\right]\\
&= \boldsymbol{\Phi}+\left[(\mathbf{X}(\eta))^T(\mathbf{X}(\eta))\right]^{-1}(\mathbf{X}(\eta))^T\sum_{k=1}^{q}\mathbf{g}_k(t)-\\
&\quad \left[(\mathbf{X}(\eta))^T(\mathbf{X}(\eta))\right]^{-1}
\end{aligned}
$$

$$\mathrm{E}\left[\hat{\boldsymbol{\Phi}}(\eta,\alpha)\right]\neq\boldsymbol{\Phi}$$

$$
\begin{aligned}
\mathrm{E}\left[\hat{\mathbf{f}}_{\eta,\alpha}(x,v)\right] &= E(\mathbf{A}(\eta,\alpha)\mathbf{y})\\
&= \mathbf{A}(\eta,\alpha)\left[\sum_{s=1}^{p}\mathbf{f}_s(x)+\sum_{k=1}^{q}h_k(v)\right]\\
&= \mathbf{A}(\eta,\alpha)\sum_{s=1}^{p}\mathbf{f}_s(x)+\mathbf{A}(\eta,\alpha)\sum_{k=1}^{q}h_k(v)
\end{aligned}
$$

$$\mathrm{E}\left[\hat{\mathbf{f}}_{\eta,\alpha}(x,v)\right]\neq\mathbf{f}_{\eta,\alpha}(x,v)$$

$$
\begin{aligned}
\mathrm{E}\left[\hat{h}_{\alpha}(v)\right] &= E(\mathbf{A}(\eta,\alpha)\mathbf{y})\\
&= \mathbf{A}(\eta,\alpha)\left[\sum_{s=1}^{p}\mathbf{f}_s(x)+\sum_{k=1}^{q}h_k(v)\right]\\
&= \mathbf{A}(\eta,\alpha)\sum_{s=1}^{p}\mathbf{f}_s(x)+\mathbf{A}(\eta,\alpha)\sum_{k=1}^{q}h_k(v)
\end{aligned}
$$

$$\mathrm{E}\left[\hat{h}_{\alpha}(v)\right]\neq h_{\alpha}(v)$$

$$
\begin{aligned}
\mathrm{E}\left[\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)\right] &= E(\left[\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha)\right]\mathbf{y})\\
&= \left[\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha)\right]\left[\sum_{s=1}^{p}\mathbf{f}_s(x)+\sum_{k=1}^{q}h_k(v)\right]\\
&= \left[\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha)\right]\sum_{s=1}^{p}\mathbf{f}_s(x)+\\
&\quad \left[\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha)\right]\sum_{k=1}^{q}h_k(v)
\end{aligned}
$$

$$\mathrm{E}\left[\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)\right]\neq\boldsymbol{\mu}_{\eta,\alpha}(x,v)$$

b) In Theorem 2 and Lemma 3, given:

$\hat{\boldsymbol{\Phi}}(\eta,\alpha)=\left[(\mathbf{X}(\eta))^T(\mathbf{X}(\eta))\right]^{-1}(\mathbf{X}(\eta))^T(\mathbf{I}-\mathbf{M}(\alpha))\mathbf{y}$,

$\hat{\mathbf{f}}_{\eta,\alpha}(x,v)=\mathbf{A}(\eta,\alpha)\mathbf{y}$, $\hat{h}_{\alpha}(v)=\mathbf{M}(\alpha)\mathbf{y}$, and $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)=\mathbf{P}(\eta,\alpha)\mathbf{y}$.

It is known that the $\hat{\boldsymbol{\Phi}}(\eta,\alpha)$, $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)$, $\hat{h}_{\alpha}(v)$, and $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)$ are linear estimators in observations $\mathbf{y}$.

IV. Conclusion

Utilizing the mixed estimator truncated spline and Epanechnikov kernel model, a successful theoretical study for a new approximation of a regression curve has been conducted.

a. The mixed estimator truncated spline and Epanechnikov kernel:

$$y_i=\mu\left(x_{1i},\ldots,x_{pi},v_{1i},\ldots,v_{qi}\right)+\varepsilon_i,\ i=1,2,\ldots,n$$

With $\mu(x,t)=\sum_{s=1}^{p}f_s(x,v)+\sum_{k=1}^{q}\hat{h}_k(v)$.

The regression curve $f_s(x_{si})$ will be approximated by the truncated spline of degree m_s and knot points $\eta_s=(\eta_{s1},\eta_{s2},\ldots,\eta_{sr})^T$. The regression curve $h_k(v_{ki})$, then will be approximated by the Epanechnikov kernel. If it is denoted in matrix form, then:

$$
\begin{aligned}
\mathbf{y} &= \sum_{s=1}^{p}\mathbf{f}_s+\sum_{k=1}^{q}\mathbf{h}_k+\boldsymbol{\varepsilon}\\
&= \mathbf{X}(\eta)\boldsymbol{\Phi}+\mathbf{M}(\alpha)\mathbf{y}+\boldsymbol{\varepsilon}
\end{aligned}
$$

using the maximum likelihood estimation (MLE) method, the following parameter estimates are obtained:

$$\hat{\boldsymbol{\Phi}}(\eta,\alpha)=\mathbf{Q}(\eta,\alpha)\mathbf{y}$$

with

$$\mathbf{Q}(\eta,\alpha)=\left[(\mathbf{X}(\eta))^T(\mathbf{X}(\eta))\right]^{-1}(\mathbf{X}(\eta))^T(\mathbf{I}-\mathbf{M}(\alpha)).$$

The estimation results of the regression curve from the mixed nonparametric regression estimator form of the truncated spline and Epanechnikov kernel are:

$$
\begin{aligned}
\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v) &= \hat{\mathbf{f}}_{\eta,\alpha}(x,v)+\hat{h}_{\alpha}(v)\\
&= \left[\mathbf{A}(\eta,\alpha)+\mathbf{M}(\alpha)\right]\mathbf{y}
\end{aligned}
$$

with

$$\mathbf{A}(\eta,\alpha)=\mathbf{X}(\eta)\left[\left(\mathbf{X}(\eta)\right)^{T}\mathbf{X}(\eta)\right]^{-1}\left(\mathbf{X}(\eta)\right)^{T}\left(\mathbf{I}-\mathbf{M}(\alpha)\right)$$

.

b. By tracing the properties for each estimator, it is known that the estimators $\hat{\mathbf{\Phi}}(\eta,\alpha)$, $\hat{\mathbf{f}}_{\eta,\alpha}(x,v)$, $\hat{h}_{\alpha}(v)$, dan $\hat{\boldsymbol{\mu}}_{\eta,\alpha}(x,v)$ are biased estimators but are still linear estimators in observations $\mathbf{y}$.

The mixed estimator truncated spline and Epanechnikov kernel model can be applied to multivariable regression cases. Suppose, several predictor variables with a response variable that follows the characteristics of the truncated spline estimator and other predictor variables that follow the characteristics of the Epanechnikov kernel estimator, then this approach can be applied.

REFERENCES

[1] G. Wahba and Y. Wang, “Spline Function,” *Encycl. Stat. Sci.*, pp. 1–27, 2014, doi: 10.4135/9781446247501.n3679.

[2] G. Wahba, “Improper Priors, Spline Smoothing and the Problem of Guarding Against Model Errors in Regression,” *J. R. Stat. Soc. Ser. B*, vol. 40, no. 3, pp. 364–372, 1978, doi: 10.1111/j.2517-6161.1978.tb01050.x.

[3] P. Speckman, “Convergence in Nonparametric Regression Models,” *Ann. Stat.*, vol. 13, no. 3, pp. 970–983, 1985.

[4] R. Hidayat, I. N. Budiantara, B. W. Otok, and V. Ratnasari, “An extended model of penalized spline with the addition of Kernel Functions in nonparametric regression model,” *Appl. Math. Inf. Sci.*, vol. 13, no. 3, pp. 453–460, 2019, doi: 10.18576/amis/130318.

[5] V. Ratnasari, I. N. Budiantara, and A. T. R. Dani, “Nonparametric Regression Mixed Estimators of Truncated Spline and Gaussian Kernel based on Cross-Validation (CV), Generalized Cross-Validation (GCV), and Unbiased Risk (UBR) Methods,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 11, no. 6, 2021.

[6] I. N. Budiantara *et al.*, “Modeling Percentage of Poor People In Indonesia Using Kernel and Fourier Series Mixed Estimator In Nonparametric Regression,” *Rev. Investig. OPERACIONAL*, vol. 40, no. 4, pp. 538–550, 2019.

[7] S. Sifriyani, M. Rasjid, D. Rosadi, S. Anwar, R. D. Wahyuni, and S. Jalaluddin, “Spatial-Temporal Epidemiology of COVID-19 Using a Geographically and Temporally Weighted Regression Model,” *Symmetry (Basel).*, vol. 14, no. 4, Apr. 2022, doi: 10.3390/sym14040742.

[8] D. D. Cox and F. Sullivan, “Penalized Likelihood-type Estimators for Generalized Nonparametric Regression,” *J. Multivar. Anal.*, vol. 56, no. 2, pp. 185–206, 1996, doi: 10.1006/jmva.1996.0010.

[9] A. A. R. Fernandes, I. N. Budiantara, and B. W. Otok, “Reproducing Kernel Hilbert space and penalized weighted least square in nonparametric regression,” *Appl. Math. Sci.*, vol. 8, no. 145–148, pp. 7289–7300, 2014, doi: 10.12988/ams.2014.49759.

[10] R. Putra and M. G. Fadhlurrahman, “Determination of The Best Knot and Bandwidth in Geographically Weighted Truncated Spline Nonparametric Regression Using Generalized Cross Validation,” *MethodsX*, p. 101994, 2023, doi: 10.1016/j.mex.2022.101994.

[11] T. H. Ali, “Modification of the adaptive Nadaraya-Watson kernel method for nonparametric regression (simulation study),” *Commun. Stat. Simul. Comput.*, vol. 51, no. 2, pp. 391–403, 2022, doi: 10.1080/03610918.2019.1652319.

[12] W. H. Wong, “On the Consistency of Cross-Validation in Kernel Nonparametric Regression,” *Ann. Stat.*, vol. 11, no. 4, pp. 1136–1141, 1983.

[13] M. Bilodeau, “Fourier smoother and additive models,” *Can. J. Stat.*, vol. 20, no. 3, pp. 257–269, 1992, doi: 10.2307/3315313.

[14] M. A. D. Octavanny, I. N. Budiantara, H. Kuswanto, and D. P. Rahmawati, “A New Mixed Estimator in Nonparametric Regression for Longitudinal Data,” *J. Math.*, vol. 2021, 2021, doi: 10.1155/2021/3909401.

[15] R. Pane, I. N. Budiantara, I. Zain, and B. W. Otok, “Parametric and nonparametric estimators in fourier series semiparametric regression and their characteristics,” *Appl. Math. Sci.*, vol. 8, no. 101–104, pp. 5053–5064, 2014, doi: 10.12988/ams.2014.46472.

[16] C. Adam and I. Gijbels, *Local polynomial expectile regression*, vol. 74, no. 2. Springer Japan, 2022. doi: 10.1007/s10463-021-00799-y.

[17] S. E. Ahmed, D. Aydin, and E. Yilmaz, “Semiparametric Time-Series Model Using Local Polynomial: An Application on the Effects of Financial Risk Factors on Crop Yield,” *J. Risk Financ. Manag.*, vol. 15, no. 3, 2022, doi: 10.3390/jrfm15030141.

[18] J. Opsomer, Y. Wang, and Y. Yang, “Nonparametric Regression with Correlated Errors,” *Stat. Sci.*, vol. 16, no. 2, pp. 134–153, 2001, doi: 10.1214/ss/1009213287.

[19] U. Amato, A. Antoniadis, I. De Feis, and I. Gijbels, “Wavelet-based robust estimation and variable selection in nonparametric additive models,” *Stat. Comput.*, vol. 32, no. 1, Feb. 2022, doi: 10.1007/s11222-021-10065-z.

[20] I. N. Budiantara, V. Ratnasari, M. Ratna, and I. Zain, “The Combination of Spline and Kernel Estimator for Nonparametric Regression and its Properties,” *Appl. Math. Sci.*, vol. 9, no. 122, pp. 6083–6094, 2015, doi: 10.12988/ams.2015.58517.

[21] A. T. R. Dani, V. Ratnasari, and I. N. Budiantara, “Optimal Knots Point and Bandwidth Selection in Modeling Mixed Estimator Nonparametric Regression,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1115, no. 1, p. 012020, Mar. 2021, doi: 10.1088/1757-899x/1115/1/012020.

[22] R. L. Eubank, “Nonparametric Regression and Spline Smoothing.” Marcel Dekker, New York, 1999.

[23] Sifriyani, S. H. Kartiko, I. N. Budiantara, and Gunardi, “Development of nonparametric geographically weighted regression using truncated spline approach,” *Songklanakarin J. Sci. Technol.*, vol. 40, no. 4, pp. 909–920, 2018, doi: 10.14456/sjst-psu.2018.98.

[24] X. Lin and D. Zhang, “Inference in generalized additive mixed models by using smoothing splines,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 61, no. 2, pp. 381–400, 1999, doi: 10.1111/1467-9868.00183.

[25] W. Härdle and A. W. Bowman, “Bootstrapping in nonparametric regression: Local adaptive smoothing and confidence bands,” *J. Am. Stat. Assoc.*, vol. 83, no. 401, pp. 102–110, 1988, doi: 10.1080/01621459.1988.10478572.

[26] N. R. Drapper and H. Smith, *Applied Regression Analysis*, 3rd ed. Canada: John Wiley and Sons Inc., 1998.

Sifriyani born in Raha on November 23, 1982. She is a Lecturer at the Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia. She received Ph.D Degree at Gadjah Mada University, Yogyakarta, Indonesia. Her research interest is Statistical Modeling with the Nonparametric Regression approach and Its Applications, Spatial Analysis, and Spatio-Temporal. The results of her research are published in several National Journals, International Proceedings, and International Journals Indexed by Scopus.

Andrea Tri Rian Dani born in Samarinda on January 16, 1998. He is a Lecturer at the Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia. He received M. Stat. Degree at Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. His research interest is Statistical Modeling with the Nonparametric-Semiparametric Regression approach and Its Applications. The results of his research have been published in several National Journals, International Proceedings, and International Journals Indexed by Scopus.

Meirinda Fauziyah born in Samarinda on May 01, 1995. She is a Lecturer at the Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia. She received M. Stat. Degree at Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Her research interest is Spatial Statistics Modelling and Extreme Value Theory approach and Its Applications. The results of her research have been published in several National Journals and International Proceedings by Scopus.

I Nyoman Budiantara born in Keliki on June 03, 1965. He is a Professor in the Statistics Department, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. He received Ph.D Degree at Gadjah Mada University, Yogyakarta, Indonesia. His research interest is Nonparametric-Semiparametric Regression and Its Applications. The results of his research are mostly published in National Journals, International Proceedings, and International Journals Indexed by Scopus.