Judul Artikel: Spline And Kernel Mixed Estimators In Multivariable Nonparametric Regression For Dengue Hemorrhagic Fever Models.

Nama Journal: Communications in Mathematical Biology and Neuroscience. 2023 (11). 2023. https://doi.org/10.28919/cmbn/7790. E-ISSN:2052-2541

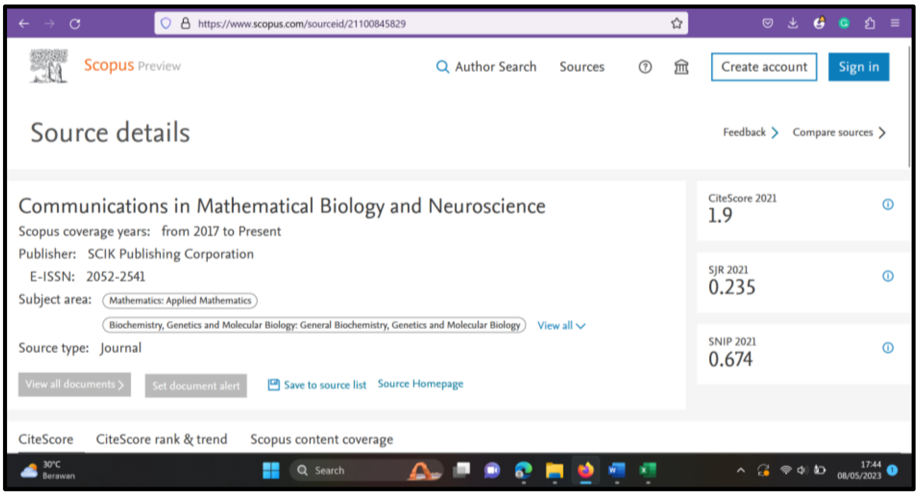
Author: Sifriyani, Andrea Tri Rian Dani, Meirinda Fauziyah, Memi Nor Hayati, Sri Wahyuningsih, Surya Prangga.

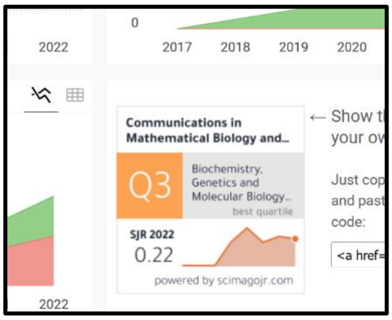
Status: Penulis pertama dan Corresponden Author

Akses Web Jurnal : https://scik.org/index.php/cmbn/article/view/7790

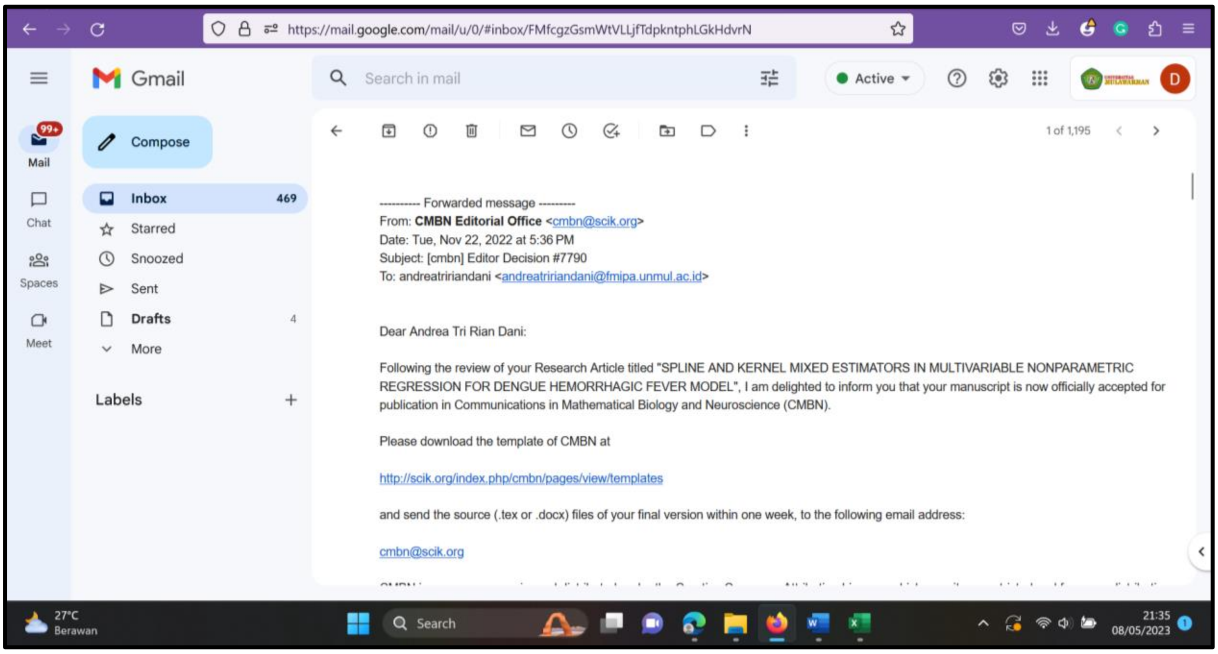
Akses Web PDF Jurnal : https://scik.org/index.php/cmbn/article/download/7790/3679

Terindeks Q3 Scopus : https://www.scopus.com/sourceid/21100845829





Bukti Correspondensi



---------- Forwarded message ---------
From: **CMBN Editorial Office** <cmbn@scik.org>
Date: Tue, Nov 22, 2022 at 5:36 PM
Subject: [cmbn] Editor Decision #7790
To: andreatririandani <andreatririandani@fmipa.unmul.ac.id>

Dear Andrea Tri Rian Dani:

Following the review of your Research Article titled "SPLINE AND KERNEL MIXED ESTIMATORS IN MULTIVARIABLE NONPARAMETRIC REGRESSION FOR DENGUE HEMORRHAGIC FEVER MODEL", I am delighted to inform you that your manuscript is now officially accepted for publication in Communications in Mathematical Biology and Neuroscience (CMBN).

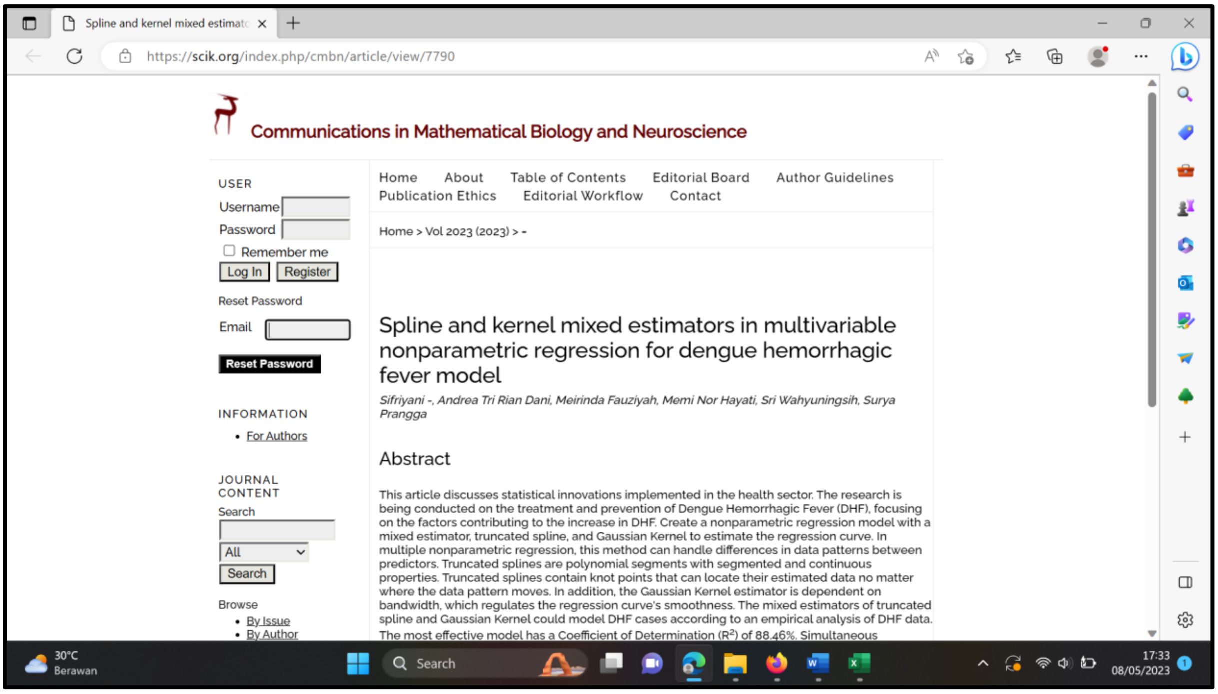
Please download the template of CMBN at

http://scik.org/index.php/cmbn/pages/view/templates

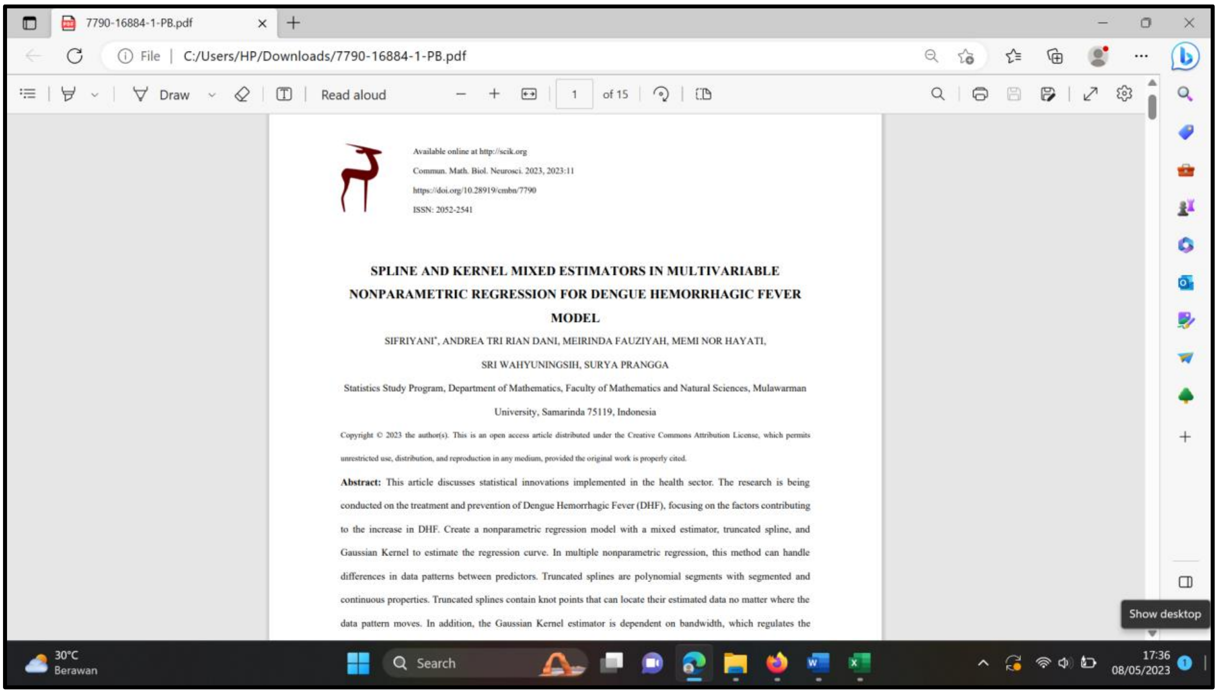
and send the source (.tex or .docx) files of your final version within one week, to the following email address:

cmbn@scik.org

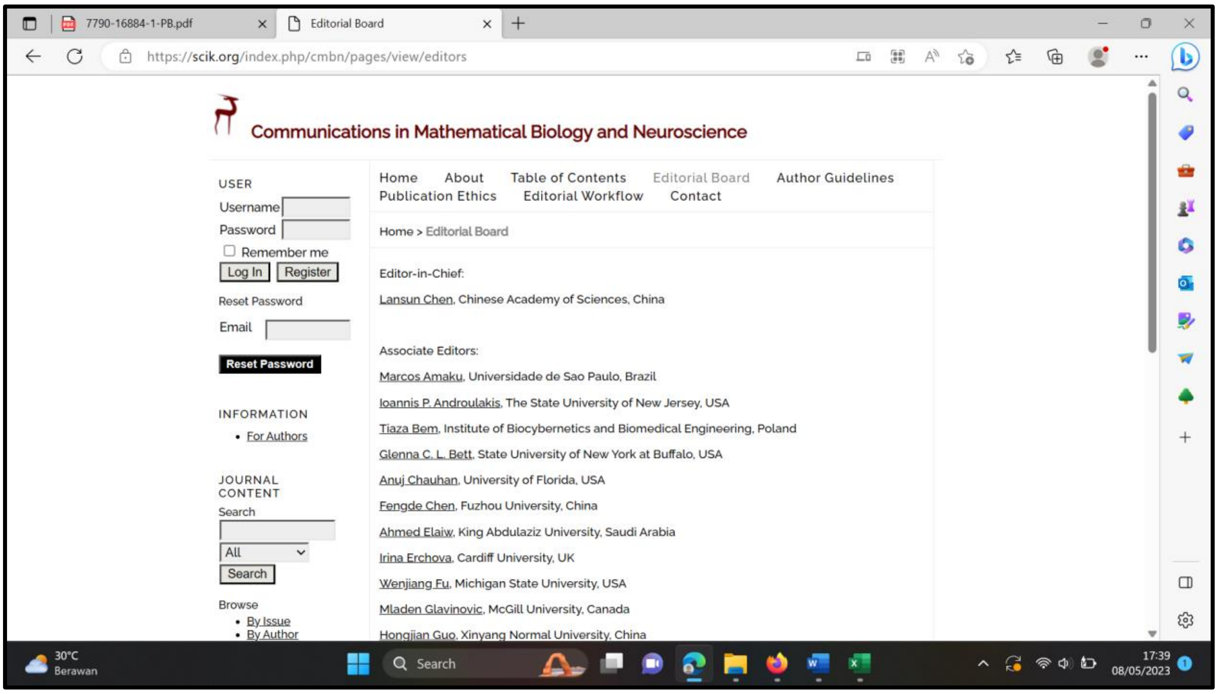
Akses Web Jurnal : https://scik.org/index.php/cmbn/article/view/7790



Akses Web PDF Jurnal : https://scik.org/index.php/cmbn/article/download/7790/3679



Tim Editor Jurnal : Editorial Board (scik.org)



7790-16884-1-PB.pdf | Editorial Board

https://scik.org/index.php/cmbn/pages/view/editors

Communications in Mathematical Biology and Neuroscience

Home | About | Table of Contents | Editorial Board | Author Guidelines | Publication Ethics | Editorial Workflow | Contact

Home > Editorial Board

USER
Username
Password
Remember me
Log In | Register
Reset Password
Email
Reset Password

INFORMATION
- For Authors

JOURNAL CONTENT
Search
All
Search
Browse
- By Issue
- By Author

Editor-in-Chief:

Lansun Chen, Chinese Academy of Sciences, China

Associate Editors:

Marcos Amaku, Universidade de Sao Paulo, Brazil

Ioannis P. Androulakis, The State University of New Jersey, USA

Tiaza Bem, Institute of Biocybernetics and Biomedical Engineering, Poland

Glenna C. L. Bett, State University of New York at Buffalo, USA

Anuj Chauhan, University of Florida, USA

Fengde Chen, Fuzhou University, China

Ahmed Elaiw, King Abdulaziz University, Saudi Arabia

Irina Erchova, Cardiff University, UK

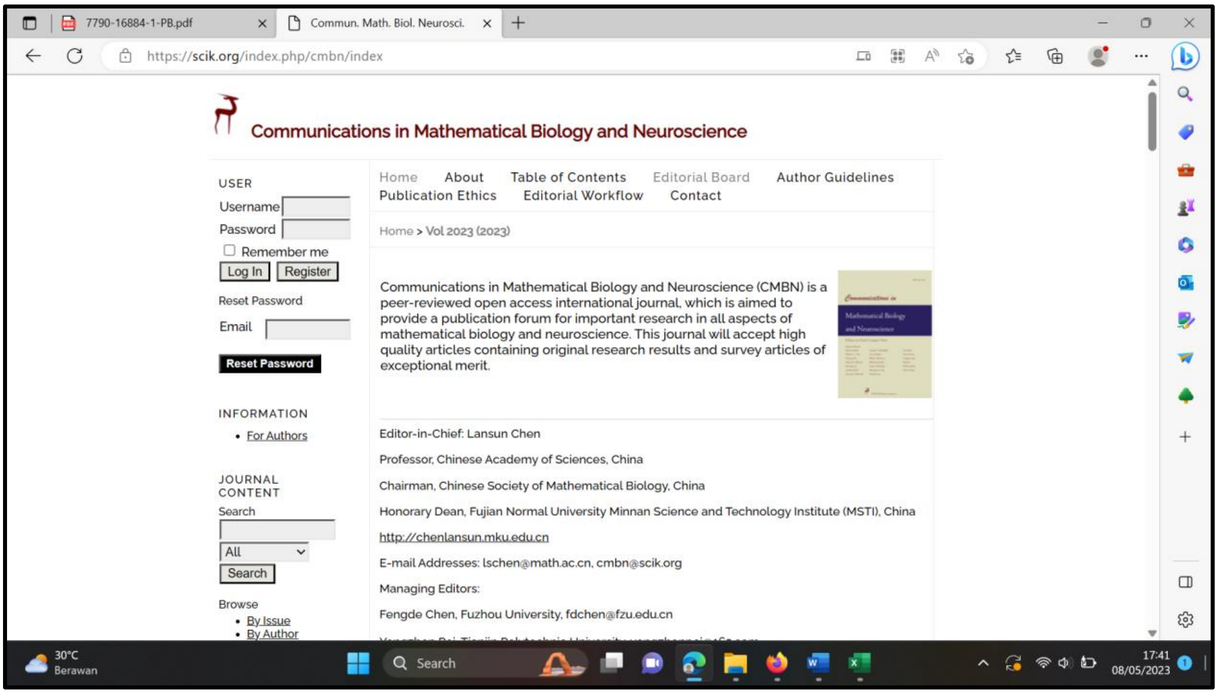
Wenjiang Fu, Michigan State University, USA

Mladen Glavinovic, McGill University, Canada

Hongjian Guo, Xinyang Normal University, China

30°C Berawan | Search | 17:39 08/05/2023

[Commun. Math. Biol. Neurosci. (scik.org)](https://scik.org/index.php/cmbn/index)



7790-16884-1-PB.pdf | Commun. Math. Biol. Neurosci.

https://scik.org/index.php/cmbn/index

Communications in Mathematical Biology and Neuroscience

Home | About | Table of Contents | Editorial Board | Author Guidelines | Publication Ethics | Editorial Workflow | Contact

Home > Vol 2023 (2023)

USER
Username
Password
Remember me
Log In | Register
Reset Password
Email
Reset Password

INFORMATION
- For Authors

JOURNAL CONTENT
Search
All
Search
Browse
- By Issue
- By Author

Communications in Mathematical Biology and Neuroscience (CMBN) is a peer-reviewed open access international journal, which is aimed to provide a publication forum for important research in all aspects of mathematical biology and neuroscience. This journal will accept high quality articles containing original research results and survey articles of exceptional merit.

Editor-in-Chief: Lansun Chen

Professor, Chinese Academy of Sciences, China

Chairman, Chinese Society of Mathematical Biology, China

Honorary Dean, Fujian Normal University Minnan Science and Technology Institute (MSTI), China

http://chenlansun.mku.edu.cn

E-mail Addresses: lschen@math.ac.cn, cmbn@scik.org

Managing Editors:

Fengde Chen, Fuzhou University, fdchen@fzu.edu.cn

30°C Berawan | Search | 17:41 08/05/2023

Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2023, 2023:X
https://doi.org/10.28919/cmbn/7790
ISSN: 2052-2541

SPLINE AND KERNEL MIXED ESTIMATORS IN MULTIVARIABLE NONPARAMETRIC REGRESSION FOR DENGUE HEMORRHAGIC FEVER MODEL

SIFRIYANI*, ANDREA TRI RIAN DANI, MEIRINDA FAUZIYAH, MEMI NOR HAYATI, SRI WAHYUNINGSIH, SURYA PRANGGA

Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda 75119, Indonesia

Copyright © 2023 the author(s). 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.

Abstract: This article discusses statistical innovations implemented in the health sector. The research is being conducted on the treatment and prevention of Dengue Hemorrhagic Fever (DHF), focusing on the factors contributing to the increase in DHF. Create a nonparametric regression model with a mixed estimator, truncated spline, and Gaussian Kernel to estimate the regression curve. In multiple nonparametric regression, this method can handle differences in data patterns between predictors. Truncated splines are polynomial segments with segmented and continuous properties. Truncated splines contain knot points that can locate their estimated data no matter where the data pattern moves. In addition, the Gaussian Kernel estimator is dependent on bandwidth, which regulates the regression curve's smoothness. The mixed estimators of truncated spline and Gaussian Kernel could model DHF cases according to an empirical analysis of DHF data. The most effective model has a Coefficient of Determination (R^2) of 88.46%. Simultaneous hypothesis testing indicates that the model contains at least one significant predictor variable.

*Corresponding author
E-mail address: sifriyani@fmipa.unmul.ac.id
Received October 19, 2022

Keywords: dengue hemorrhagic fever (DHF) modeling; Gaussian kernel; maximum likelihood (ML); mixed estimators; nonparametric regression; truncated spline.

2020 AMS Subject Classification: 92C60.

1. INTRODUCTION

The relationship between the response and predictor variable, whose purpose is unknown, can be identified statistically using nonparametric regression [1], [2]. Nonparametric regression is not rigid in defining the regression function [3], does not require certain assumptions like linear regression, where the error must be normally distributed, and does not force the regression curve to be linear. This is in contrast to parametric regression, which causes the regression curve to follow a specific model, such as a linear model. The observational data determine the advantage of nonparametric regression, the regression curve without being forced to adjust a specific function [4], [5]. Nonparametric regression assumes that the data derive their form of estimation from the regression curve without regard for the researcher's subjectivity [6]. As a result, the nonparametric regression model approach is both adaptable and objective [2], [3].

The estimator approach used in nonparametric regression includes truncated spline and Kernel. Truncated spline is polynomial pieces that have segmented and continuous properties [7], [8]. One of the advantages of the truncated spline is that this model tends to find its own estimate of the data wherever the data pattern moves [9], [10]. This advantage occurs because, in the truncated spline, knot points indicate changes in data behavior patterns [4], [5]. While the Kernel estimator has the advantage that it is flexible [11], the mathematical form is easy and can achieve a relatively fast level of convergence [12]. The Kernel approach depends on bandwidth, which controls the smoothness of the estimation curve [5], [12]–[14]. Estimating the regression curve with the Kernel estimator approach adjusts from the value of the smoothing parameter λ.

According to Budiantara et al. [15], the nonparametric and semiparametric regression models developed by the researchers so far, if explored more deeply, basically there are very heavy and basic assumptions in the model. Each predictor in the multi-predictor nonparametric regression is considered to have the same pattern, so the researchers force the use of only one form of the model estimator for all predictor variables. Therefore, using only one form of the estimator in various forms of different data relationship patterns will certainly result in the resulting estimator not being compatible with the data pattern. As a result, the estimation of the regression model is not good and produces a significant error. Therefore, to overcome this problem, several researchers have developed a nonparametric mixed regression curve estimator in which an appropriate curve estimator approximates each data pattern in the nonparametric regression model. There are several studies that have developed and reviewed mixed estimator models, including [1], [16]–[19].

Dengue Hemorrhagic Fever is one of the problems in Indonesia's health sector. DHF is caused by the bite of the Aedes Aegypti mosquito [20], which usually attacks tropical and subtropical areas of the world [21], [22], one of which is Indonesia. Based on data from the World Health Organization (WHO), Indonesia has the 2nd rank with the most significant DHF cases among 30 endemic areas (Ministry of Health, 2018). In 2020, there were 108303 patients with DHF cases, and 747 died. Meanwhile, the number of DHF cases in 2021 was 73518, and 705 died. The number decreased by 32.12% compared to the previous year, but this case still needs special attention.

Based on the description that has been explained, so the purpose of this study is to conduct a study of the nonparametric regression Mixed Estimator of Truncated Spline and Gaussian Kernel (MTs-GK) model in the additive multi-predictor nonparametric model and the implementation of the model in the case study of Dengue Hemorrhagic Fever (DHF) with a special issue of the factors that influence the increase in DHF.

2. PRELIMINARIES

A. Mixed Estimators of Truncated Spline and Gaussian Kernel

A mixed estimator is a multi-predictor nonparametric regression model that uses two or more types of estimators to approximate the regression curve [15], [16]. Budiantara et al. [15] were the first to develop a mixed truncated spline and Kernel nonparametric regression model. A Mixed estimator is a model approach in nonparametric regression where more than one estimator is used [18], [23], [24]. The form of the regression curve for each relationship between the predictor and the response variable will be approximated by two or more estimators according to the characteristics of the relationship [25].

For example, given paired data (x_i, v_i, y_i) and the relationship between predictor variables (x_i, v_i) with response variable (y_i) following a nonparametric regression model:

$$y_i = \mu(x_i, v_i) + \varepsilon_i \qquad (1)$$

$\mu(x_i, v_i)$ is the regression curve, with assumed to be unknown, smooth, and follows an additive model so that $\mu(x_i, v_i)$ we can write in the form in Equation (2).

$$\mu(x_i, v_i) = m(x_i) + h(v_i) \qquad (2)$$

Based on Equation (2), the regression curve $m(x_i)$ will be estimated with a truncated spline estimator, while the regression curve $h(v_i)$ with a Kernel estimator.

The truncated spline estimator is a segmented polynomial model [26], [27]. For example, given paired data (x_i, y_i) where the relationship from the predictor (x_i) and response variable (y_i) follow a truncated spline nonparametric regression model [7], [28].

$$y_i = \beta_0 + \sum_{j=1}^{m}\sum_{p=1}^{q} \beta_{jp} x_{pi}^{j} + \sum_{k=1}^{r}\sum_{p=1}^{q} \beta_{(m+k)p} \left(x_{pi} - K_{kp}\right)_+^m + \varepsilon_i \qquad (3)$$

The truncated function is:

$$\left(x_{pi} - K_{kp}\right)_+^m = \begin{cases} \left(x_{pi} - K_{kp}\right)^m & x_p \geq K_{kp} \\ 0 & x_p < K_{kp} \end{cases}$$

In matrix form, Equation (3) is as follows:

$$\mathbf{y} = \mathbf{X}(K)\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (4)$$

As a result, the regression curve estimation using the truncated spline estimator can be written in Equation (5).

$$\hat{\mathbf{m}}(x) = \mathbf{X}(K)\hat{\boldsymbol{\beta}} \qquad (5)$$

Kernel estimator has a good ability to model data that does not have a certain pattern [11], [29],

[30]. For example, given paired data, (v_i, y_i) where is the relationship between the predictor (v_i) and with response variable (y_i) following the Kernel nonparametric regression model.

$$y_i = h(v_i) + \varepsilon_i \tag{6}$$

The regression curve $h(v_i)$ will be estimated with the Kernel estimator in Equation 7.

$$\begin{aligned}\hat{h}_\lambda(v_i) &= \frac{1}{n}\sum_{i=1}^{n}\frac{K_\lambda(v-v_i)}{\frac{1}{n}\sum_{i=1}^{n}K_\lambda(v-v_i)}y_i \\ &= \frac{1}{n}\sum_{i=1}^{n}R_{\lambda i}(v)y_i\end{aligned} \tag{7}$$

With

$$R_{\lambda i}(v) = \frac{K_\lambda(v-v_i)}{\frac{1}{n}\sum_{i=1}^{n}K_\lambda(v-v_i)}$$

$$K_\lambda(v-v_i) = \frac{1}{\lambda}K\left(\frac{v-v_i}{\lambda}\right)$$

K is an abbreviation for Kernel Function. The Gaussian Kernel function is used in this research:

$$K(v) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{1}{2}\left(-v^2\right)\right) \tag{8}$$

Based on equations (7) and (8), we can write them in matrix form as:

$$\hat{\mathbf{h}}(v) = \mathbf{D}(\lambda)\mathbf{y} \tag{9}$$

Furthermore, based on Equation (2) and the form of each estimator in Equations (5) and (9), we can write:

$$\mathbf{y} = \mathbf{X}(K)\boldsymbol{\beta} + \mathbf{D}(\lambda)\mathbf{y} + \boldsymbol{\varepsilon} \tag{10}$$

Vector $\boldsymbol{\varepsilon}$ has a size $(n \times 1)$, so that based on Equation (10), then:

$$\begin{aligned}\boldsymbol{\varepsilon} &= \mathbf{y} - \left(\mathbf{X}(K)\boldsymbol{\beta} - \mathbf{D}(\lambda)\mathbf{y}\right) \\ &= \left(\mathbf{I} - \mathbf{D}(\lambda)\right)\mathbf{y} - \mathbf{X}(K)\boldsymbol{\beta}\end{aligned} \tag{11}$$

The parameter estimation of $\boldsymbol{\beta}$ used the Least Squares (LS). The estimated results for $\hat{\boldsymbol{\beta}}$ is:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}(K)^T \mathbf{X}(K)\right)^{-1} \mathbf{X}(K)^T \left(\mathbf{I} - \mathbf{D}(\lambda)\right)\mathbf{y} \tag{12}$$

Equation (12) can be summarized as:

$$\hat{\boldsymbol{\beta}} = \mathbf{A}(K,\lambda)\mathbf{y} \tag{13}$$

Where $\mathbf{A}(K,\lambda) = \left(\mathbf{X}(K)^T \mathbf{X}(K)\right)^{-1} \mathbf{X}(K)^T \left(\mathbf{I} - \mathbf{D}(\lambda)\right)$.

In Equation (5), the regression curve estimation is written using the truncated spline estimator is $\hat{\mathbf{m}}(x) = \mathbf{X}(K)\hat{\boldsymbol{\beta}}$, then based on Equation (12), we can write:

$$\begin{aligned} \hat{\mathbf{m}}(x) &= \mathbf{X}(K)\hat{\boldsymbol{\beta}} \\ &= \mathbf{X}(K)\left[\left(\mathbf{X}(K)^T \mathbf{X}(K)\right)^{-1} \mathbf{X}(K)^T \left(\mathbf{I} - \mathbf{D}(\lambda)\right)\mathbf{y}\right] \end{aligned} \tag{14}$$

A brief summary of Equation (14) is:

$$\hat{\mathbf{m}}(x) = \mathbf{S}(K,\lambda)\mathbf{y} \tag{15}$$

According to Equation (15) and the shape of the estimator for each component in Equation (9) and (15), the mixed estimator of truncated spline and Gaussian Kernel will be obtained as follows:

$$\begin{aligned} \hat{\boldsymbol{\mu}}(x,v) &= \hat{\mathbf{m}}(x) + \hat{\mathbf{h}}(v) \\ &= \left(\mathbf{S}(K,\lambda) + \mathbf{D}(\lambda)\right)\mathbf{y} \\ &= \mathbf{B}(K,\lambda)\mathbf{y} \end{aligned} \tag{16}$$

Matrix $\mathbf{B}(K,\lambda)$ very dependent on $\mathbf{S}(K,\lambda)$ which is a component of the truncated spline estimator, where the optimal location and number of knot points must be determined, and matrix $\mathbf{D}(\lambda)$, which is a component of the Kernel estimator, needs to find the correct bandwidth value. In this study, the method used to select the optimal knot point and bandwidth is Unbiased Risk (UBR) [4], [17], [18] with the formula in Equation (17).

$$UBR(K_{opt},\lambda_{opt}) = \frac{1}{n}\left(\begin{array}{l} \left\|\left(\mathbf{I} - \mathbf{B}(K,\lambda)\right)\mathbf{y}\right\|^2 + \dfrac{\hat{\sigma}^2}{n} trace\left[\mathbf{I} - \mathbf{B}(K,\lambda)\right]^2 + \\ \dfrac{\hat{\sigma}^2}{n} trace\left[\mathbf{B}(K,\lambda)^2\right]^2 \end{array} \right) \tag{17}$$

Where:

$$\hat{\sigma}^2 = \frac{\left\| (\mathbf{I} - \mathbf{B}(K,\lambda))\mathbf{y} \right\|^2}{tr\left((\mathbf{I} - \mathbf{B}(K,\lambda))\mathbf{y} \right)}$$

B. Simultaneous Testing Hypothesis

Hypothesis testing can only be done on the truncated spline estimator component based on the mixed estimator model in Equation (11) using the Likelihood Ratio Test (LRT).

Hypothesis Formulation:

H_0 : $\beta_1 = \beta_2 = ... = \beta_{(m+r)p} = 0$

H_1 : there is at least one $\beta_j \neq 0$, $j = 1, 2, ..., (m+r)p$

In summary, the following ANOVA table gives the simultaneous hypothesis testing process for the parameters in the mixed estimator model.

Source	Degree of Freedom (df)	Sum of Squares (SS)	Mean Squares (MS)	F-Test
Regression	$(m+r)p$	$SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$	$MSR = \frac{SSR}{(m+r)p}$	$\mathrm{F}_{test} = \frac{MSR}{MSE}$
Error	$n - ((m+r)p) - 1$	$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$	$MSE = \frac{SSR}{n - ((m+r)p) - 1}$	
Total	$n - 1$	$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$		

TABLE 1. ANOVA

Under the null hypothesis H_0, test statistics from F_{test} following the F Distribution, with the degree of freedom (df) is ($(m+r)p$, $n - ((m+r)p) - 1$).

3. RESEARCH METHODOLOGY

A. Data Sources

The research used secondary data from Wahab Syahrani General Hospital (AWS Hospital) in Samarinda. The variables of this study are described in Table 2.

Variable	Notation	Description	Unit	Data Scale
DHF patient's platelet count when first blood check	y	The platelet count of DHF patients when they first do a blood check	μl	Continuous
Number of Hematocrits	x_1	The number of hematocrits found in patients with DHF	%	
Number of Hemoglobin	x_2	The number of hemoglobin cells found in patients with DHF	*g/dL*	
Number of Leukocyte	x_3	The number of leukocytes found in patients with DHF	*g/dL*	

TABLE 2. Description of Study Variables and Unit Data

B. Data Analysis Technique

To answer the research purposes, necessary to develop research steps. The research steps used in this study are:

1. Create a scatter plot to show the relationship between each predictor variable and the response.
2. Determine the predictor variables for the truncated spline and Kernel components.
3. Modeling case data of patients with dengue fever with the response variable (y) being the patient platelet count using a mixed truncated spline and Gaussian Kernel estimator model based on Equation (16).

4. Select the optimal knot point and bandwidth based on the minimum UBR value with the Formula in Equation (17). Each predictor variable in this study has the same number of knot points (1 to 3). The bandwidth values tested are in the interval of 0.05 to 5.
5. Determine the best model of the mixed estimator truncated spline and Kernel based on the minimum UBR value and then calculate the Coefficient of Determination (R^2) value.
6. Simultaneous hypothesis testing for the best model based on ANOVA in Table 1.

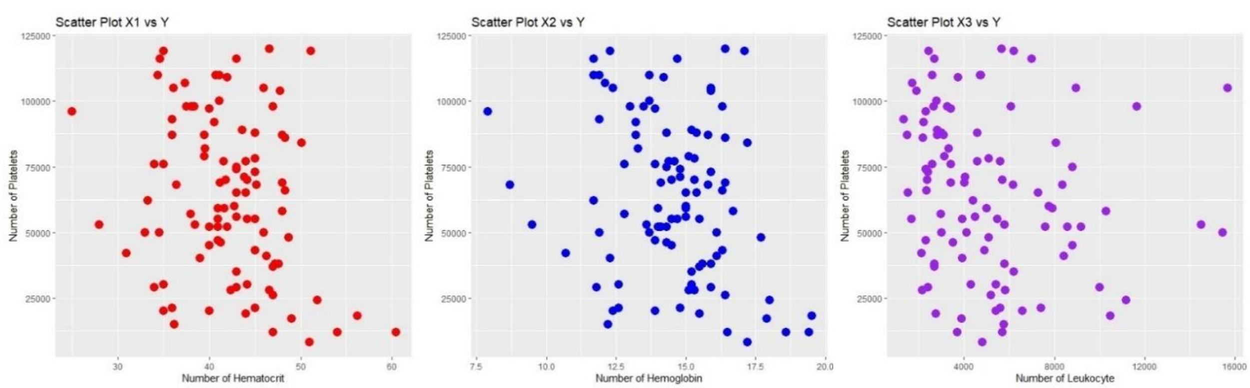
4. MAIN RESULTS

In this section, we will explain the results of the study mixed estimator truncated spline and Gaussian Kernel applied to data on the platelet count of DHF patients.

A. Scatter Plot

The first step in the modeling process using a mixed estimator is creating a scatter diagram for each variable. The scatter diagram for each predictor variable to the response variable is shown in Figure 1.

FIGURE 1. Scatter Plot



(a) (b) (c)

Based on Figure 1, it can be determined which type of estimator will be used for each predictor variable. A more detailed summary of the results of determining the estimator is presented in Table 3.

Variable	Notation	Description	Estimator
Predictor	x_1	Number of Hematocrits	Truncated Spline
	x_2	Number of Hemoglobin	
	v_1	Number of Leukocyte	Kernel

TABLE 3. Components of Truncated Spline and Gaussian Kernel Estimator

B. Modeling using Mixed Estimator Truncated Spline and Gaussian Kernel

The best model from the mixed estimator truncated spline and Gaussian Kernel was selected by comparing the smallest Unbiased Risk (UBR) value among various models based on the number of knot points and bandwidth. In this study, the number of knots used is the same for each predictor variable, namely, 1 to 3. The bandwidth values tested are in the interval of 0.05 to 5. The modeling results using a mixed estimator truncated spline and Gaussian Kernel are in Table 4.

Number of Knot Point	**Knot Point Location**		**Bandwidth**	**UBR Value**
	x_1	x_2	v_1	
1 knot	37.21	11.90	0.55	97.46
2 knots	**35.98** **37.21**	**11.50** **11.90**	**2.80**	**96.00**
3 knots	44.53 45.75 46.97	14.30 14.70 15.10	2.55	99.26

TABLE 4. Summary of Modeling Results

Based on Table 4, the minimum UBR value is 96.00 with an optimal bandwidth of 2.80, and the optimal knot point location for each predictor variable modeled with a truncated spline estimator is:

Variable x_1 (Number of Hematocrit)

$K_1 = 35.98 \qquad K_2 = 37.21$

Variable x_2 (Number of Hemoglobin)

$K_1 = 11.50 \qquad K_2 = 11.90$

Using the knot point and bandwidth optimal, a nonparametric regression model with mixed estimator truncated spline and Gaussian Kernel is written in Equation (18).

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_{11}x_{1i} + \hat{\beta}_{12}x_{2i} + \hat{\beta}_{21}\left(x_{1i} - K_{11}\right)_+ + \hat{\beta}_{31}\left(x_{1i} - K_{21}\right)_+ + \hat{\beta}_{22}\left(x_{2i} - K_{12}\right)_+ +$$
$$= \hat{\beta}_{32}\left(x_{2i} - K_{22}\right)_+ + \frac{1}{100}\sum_{i=1}^{100}\frac{K_\lambda(v-v_i)}{\frac{1}{100}\sum_{i=1}^{100}K_\lambda(v-v_i)}y_i$$

$$\hat{y}_i = 22487.37 - 541.58x_{1i} + 284.55x_{2i} + 2718.32\left(x_{1i} - K_{11}\right)_+ -$$
$$= 1603.31\left(x_{1i} - K_{21}\right)_+ - 16156.97\left(x_{2i} - K_{12}\right)_+ +$$
$$= 13711.35\left(x_{2i} - K_{22}\right)_+ + \frac{1}{100}\sum_{i=1}^{100}\frac{K_{2.80}(v-v_i)}{\frac{1}{100}\sum_{i=1}^{100}K_{2.80}(v-v_i)}y_i \qquad (18)$$

The coefficient of determination (R^2) of this model is 88.46%. This means that 88.46% of the platelet count of patients with Dengue Hemorrhagic Fever (DHF) can be explained by the variables Number of Hematocrit, Number of Hemoglobin, and Number of Leukocyte in the mixed estimator model of truncated spline and Gaussian Kernel with two-knot points and optimal bandwidth.

C. Simultaneous Testing Hypothesis

The next step will be to simultaneously test the hypothesis for the parameters in the model based on the best model from the mixed estimator truncated spline and Gaussian Kernel.

H_0 : $\beta_1 = \beta_2 = ... = \beta_{(m+r)p} = 0$

H_1 : there is at least one $\beta_j \neq 0$, $\quad j = 1, 2, ..., (m+r)p$

The ANOVA table for the results of hypothesis testing is presented in Table 5.

Source	**Degree of Freedom (df)**	**Sum of Squares (SS)**	**Mean Squares (MS)**	**F-Test**	**P-Value**
Regression	6	82542921847	13757153641	118.78	2.26e-41
Error	93	10771268700	115820094		
Total	99	91430590000			

TABLE 5. Summary of ANOVA

Based on Table 5, it can be seen that the F_{test} (118.78) is greater than the $F_{(0.05;6;93)}$ (2.19) or P-Value (2.26e-41) is smaller than the value $\alpha = 0.05$, so the decision is rejected H_0. This means simultaneously; there is at least one $\beta_j \neq 0$ or at least one significant predictor variable in the model.

5. CONCLUSION

A mixed estimator truncated spline and Gaussian Kernel model was used to successfully model the cases of Dengue Hemorrhagic Fever (DHF) patients. A nonparametric regression model of mixed estimator truncated spline and Gaussian Kernel with 2-knot points and optimal bandwidth is the best model based on the lowest UBR value.

$$\hat{y}_i = 22487.37 - 541.58x_{1i} + 284.55x_{2i} + 2718.32\left(x_{1i} - K_{11}\right)_+ - 1603.31\left(x_{1i} - K_{21}\right)_+ - \\ = 16156.97\left(x_{2i} - K_{12}\right)_+ + 13711.35\left(x_{2i} - K_{22}\right)_+ + \frac{1}{100}\sum_{i=1}^{100}\frac{K_{2.80}(v - v_i)}{\frac{1}{100}\sum_{i=1}^{100}K_{2.80}(v - v_i)}y_i$$

The best model's coefficient of determination (R^2) is 88.46%. Based on the results of simultaneous hypothesis testing, it can be concluded that simultaneously there is at least one significant predictor variable in the model.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest.

ACKNOWLEDGEMENTS

The author(s) gratefully acknowledge the funding of DIPA BLU – PNBP FMIPA / Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia, in 2022 [No: 1715/UN17.7/LT/2022].

REFERENCES

[1] N.P.A.M. Mariati, I.N. Budiantara, V. Ratnasari, Combination Estimation of Smoothing Spline and Fourier Series in Nonparametric Regression, J. Math. 2020 (2020), 4712531. https://doi.org/10.1155/2020/4712531.

[2] W. Hu, T. Pan, D. Kong, et al. Nonparametric matrix response regression with application to brain imaging data analysis, Biometrics. 77 (2020), 1227–1240. https://doi.org/10.1111/biom.13362.

[3] Rismal, I.N. Budiantara, D.D. Prastyo, Mixture model of spline truncated and kernel in multivariable nonparametric regression, AIP Conf. Proc. 1739 (2016), 020085. https://doi.org/10.1063/1.4952565.

[4] G. Wahba, Y. Wang, Spline function, in: Encyclopedia of Environmental Change, SAGE Publications, Thousand Oaks, California, 2014. https://doi.org/10.4135/9781446247501.n3679.

[5] R.L. Eubank, Nonparametric regression and spline smoothing, Marcel Dekker, New York, 1999.

[6] Sifriyani, I.N. Budiantara, S.H. Kartiko, et al. A new method of hypothesis test for truncated spline nonparametric regression influenced by spatial heterogeneity and application, Abstr. Appl. Anal. 2018 (2018), 9769150. https://doi.org/10.1155/2018/9769150.

[7] Sifriyani, S.H. Kartiko, I.N. Budiantara, et al. Development of nonparametric geographically weighted regressionusing truncated spline approach, Songklanakarin J. Sci. Technol. 40 (2018), 909-920. https://doi.org/10.14456/SJST-PSU.2018.98.

[8] S. Sifriyani, M. Rasjid, D. Rosadi, et al. Spatial-Temporal epidemiology of COVID-19 using a geographically and temporally weighted regression model, Symmetry. 14 (2022), 742. https://doi.org/10.3390/sym14040742.

[9] N. Chamidah, B. Lestari, I.N. Budiantara, et al. Consistency and asymptotic normality of estimator for parameters in multiresponse multipredictor semiparametric regression model, Symmetry. 14 (2022), 336. https://doi.org/10.3390/sym14020336.

[10] M.A. Sahraei, H. Duman, M.Y. Çodur, et al. Prediction of transportation energy demand: Multivariate adaptive regression splines, Energy. 224 (2021), 120090. https://doi.org/10.1016/j.energy.2021.120090.

[11] T.H. Ali, Modification of the adaptive Nadaraya-Watson kernel method for nonparametric regression (simulation study), Commun. Stat. - Simul. Comput. 51 (2019), 391–403. https://doi.org/10.1080/03610918.2019.1652319.

[12] A.C. Guidoum, Kernel estimator and bandwidth selection for density and its derivatives: The kedd Package, Algeria, 2020.

[13] P. Du, C.F. Parmeter, J.S. Racine, Nonparametric kernel regression with multiple predictors and multiple shape constraints, Department of Economics Working Papers, McMaster University, Canada, 2012.

[14] I.N. Budiantara, M. Ratna, I. Zain, et al. Modeling the percentage of poor people in Indonesia using spline nonparametric regression approach, Int. J. Basic Appl. Sci. 12 (2012), 119–124.

[15] I.N. Budiantara, V. Ratnasari, M. Ratna, et al. The combination of spline and kernel estimator for nonparametric regression and its properties, Appl. Math. Sci. 9 (2015), 6083–6094. https://doi.org/10.12988/ams.2015.58517.

[16] V. Ratnasari, I.N. Budiantara, M. Ratna, et al. Estimation of nonparametric regression curve using mixed estimator of multivariable truncated Spline and multivariable Kernel, Glob. J. Pure Appl. Math. 12 (2016), 5047-5057.

[17] V. Ratnasari, I.N. Budiantara, 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. Inform. Technol. 11 (2021), 2400. https://doi.org/10.18517/ijaseit.11.6.14464.

[18] A.T.R. Dani, V. Ratnasari, I.N. Budiantara, Optimal knots point and bandwidth selection in modeling mixed estimator nonparametric regression, IOP Conf. Ser.: Mater. Sci. Eng. 1115 (2021), 012020. https://doi.org/10.1088/1757-899X/1115/1/012020.

[19] R. Hidayat, I.N. Budiantara, B.W. Otok, et al. The regression curve estimation by using mixed smoothing spline and kernel (MsS-K) model, Commun. Stat. - Theory Methods. 50 (2021), 3942–3953. https://doi.org/10.1080/03610926.2019.1710201.

[20] D.J. Gubler, Dengue and dengue hemorrhagic fever, Clin. Microbiol. Rev. 11 (1998), 480–496. https://doi.org/10.1128/CMR.11.3.480.

[21] C.H. Lin, T.H. Wen, Using geographically weighted regression (GWR) to explore spatial varying relationships of immature mosquitoes and human densities with the incidence of dengue, Int. J. Environ. Res. Public Health. 8 (2011), 2798–2815. https://doi.org/10.3390/ijerph8072798.

[22] W.H. Wang, A.N. Urbina, M.R. Chang, et al. Dengue hemorrhagic fever – A systemic literature review of current perspectives on pathogenesis, prevention and control, J. Microbiol. Immunol. Infect. 53 (2020), 963–978. https://doi.org/10.1016/j.jmii.2020.03.007.

[23] R. Hidayat, I.N. Budiantara, B.W. Otok, et al. Kernel-spline estimation of additive nonparametric regression

model, IOP Conf. Ser.: Mater. Sci. Eng. 546 (2019), 052028. https://doi.org/10.1088/1757-899X/546/5/052028.
[24] D.P. Rahmawati, I.N. Budiantara, D.D. Prastyo, et al. Mixed spline smoothing and kernel estimator in biresponse nonparametric regression, Int. J. Math. Math. Sci. 2021 (2021), 6611084. https://doi.org/10.1155/2021/6611084.
[25] N.P.A.M. Mariati, I.N. Budiantara, V. Ratnasari, The application of mixed smoothing spline and Fourier series model in nonparametric regression, Symmetry. 13 (2021), 2094. https://doi.org/10.3390/sym13112094.
[26] N. Chamidah, B. Lestari, A. Massaid, et al. Estimating mean arterial pressure affected by stress scores using spline nonparametric regression model approach, Commun. Math. Biol. Neurosci. 2020 (2020), 72. https://doi.org/10.28919/cmbn/4963.
[27] G.P.D. Sohibien, L. Laome, A. Choiruddin, et al. COVID-19 pandemic's impact on return on asset and financing of islamic commercial banks: evidence from Indonesia, Sustainability. 14 (2022), 1128. https://doi.org/10.3390/su14031128..
[28] Sifriyani, Haryatmi, I.N. Budiantara, et al. Geographically weighted regression with spline approach, Far East J. Math. Sci. 101 (2017), 1183–1196. https://doi.org/10.17654/MS101061183.
[29] L.R. Cheruiyot, Local linear regression estimator on the boundary correction in nonparametric regression estimation, J. Stat. Theory Appl. 19 (2020), 460-471. https://doi.org/10.2991/jsta.d.201016.001..
[30] H. Husain, I.N. Budiantara, I. Zain, Mixed estimator of spline truncated, Fourier series, and kernel in biresponse semiparametric regression model, IOP Conf. Ser.: Earth Environ. Sci. 880 (2021) 012046. https://doi.org/10.1088/1755-1315/880/1/012046.