Modelling Return on Assets (ROA) using nonparametric regression spline truncated for longitudinal data

Mustain Ramli¹, Vita Ratnasari² I Nyoman Budiantara³

¹,²,³ Department of Statistics, Faculty of Science and Data Analytics, Sepuluh Nopember Institute of Technology, Jl. Arif Rahman Hakim, Surabaya 60111, Indonesia

(*)Corresponding author : mustainramli1994@gmail.com¹, vitaratna70@gmail.com², nyomanbudiantara65@gmail.com³

Abstract. Nonparametric regression is one of the statistical methods used to see the relationship between response variables and predictor variables with unknown curves. The method often used in estimating the curve of nonparametric regression is the spline method. Spline is one of the nonparametric regression approaches that have very specific and very good statistical and visual interpretations. Spline truncated nonparametric regression is not only used to modelling data with cross-section type but can be used to modelling longitudinal data. Longitudinal data is data whose observations are repeated on each subject within a certain time. In this research, the modelling of Return On Assets (ROA) will be based on the factors that influence it with using 30 subjects of go public banks listed on the Indonesia stock exchange during 2012-2018 using nonparametric regression spline truncated for longitudinal data, with the weighting matrix using the variance-covariance matrix of errors. Based on the analysis data that has been done by using the weighting matrix of the error variance-covariance matrix and one-knot point with 14 Increments, the best model is obtained, namely using the weighting matrix based of the error variance-covariance matrix one-knot point with the smallest GCV value of 15.12 with MSE value of 0.739 and \( R^2 \) value of 85.56%.

1. Introduction
Regession analysis is a statistical analysis used to find out the similarity in the pattern of relationships between predictor variables and response variables. There are two models of estimation approaches in regression analysis, namely parametric regression and nonparametric regression. The parametric regression approach is used in the shape of the regression curve is known [1]. When the shape of the curve is unknown, nonparametric regression is a model approach that assumes the curve pattern is unknown and is contained within a particular function [2]. The approach using nonparametric regression will produce a better model because the data are expected to find their own model estimation form without being influenced by the subjectivity of the researcher and also this approach is very flexible [1]. Several models of nonparametric regression approach are of concern and most often used by many researchers including kernel nonparametric regression [3], spline nonparametric regression [4] - [2], nonparametric regression Fourier series [5] and Wavelet [6]... Among the nonparametric regression models obtained above, spline is a model that has a very specific and very good statistical and visual interpretation [7]. Spline is one of the polynomial type, namely polynomial which has segmented properties. This segmented nature provides more than ordinary polynomials,
making it possible to adjust more effectively to the characteristics of a data. Wahba [4] shows that spline has statistical properties that are useful for relationship analysis in regression. Spline also has a very good ability for data whose behavior is altered at certain sub-intervals [8].

Researchers who use spline truncated nonparametric regression in their research include Saputro [9] modelling nonparametric regression spline truncated on the data of Human Development Index (HDI) in Indonesia. Ratnasari [10] estimation of nonparametric regression curve using mixed estimator of multivariable truncated spline and multivariable kernel. Sudiarsa [11] combined estimator Fourier series and spline truncated in multivariable nonparametric regression.

In addition to spline truncated nonparametric regression modelling using cross-section data, there are also nonparametric spline truncated studies using longitudinal data. Longitudinal data is data obtained from repeated observations for each subject in different periods. According to Diggle [12], longitudinal data is a combination of cross-section data with time-series data where longitudinal data between subjects are mutually independent of each other and between observations are interdependent so there is a correlation. Research with a spline truncated nonparametric approach for longitudinal data is mostly done in the health sector, but can also be applied in other fields including the social and economic fields. The advantage of using longitudinal data is that it can determine changes that occur in individuals, does not require a lot of subjects because of repeated observations and also more efficient estimation because every observation is carried out [13].

Some researchers who use longitudinal data, among others, Fernandes [14] with the aim of its research is to obtain a form of function from a nonparametric multipredictor regression on longitudinal data and obtain a spline estimator in estimating nonparametric bi-responses in longitudinal data. Damaliana [15] comparing between mGCV and aGCV methods to choose the optimal knot points in semiparametric regression with spline truncated using longitudinal data.

Modelling using spline truncated nonparametric regression for longitudinal data has been done but with the assumption that the weighting matrix is given so that in this study the development is done by using the weighting matrix obtained based on the estimation of the error variance-covariance matrix. The model of nonparametric regression spline truncated for longitudinal data will be applied on the modelling Return On Assets (ROA) for go public banks that listed on the Indonesian stock exchange within period 2012-2018 based on factors that are suspected to be influential.

2. Theoretical Review
In this section, we will review some of the theories used.

2.1. Nonparametric Regression Spline Truncated
Nonparametric regression spline truncated is a regression method that has a regression curve designed by modifying polynomial functions. One nonparametric regression model that has a very specific and very good statistical and visual interpretation is a spline. Spline is polynomial pieces that have segmented properties (piecewise polynomial) at the point of the knot [1]. Spline method is very good in modelling data whose patterns change at certain sub-intervals [8]. Given data in pairs $\left( x_{i1}, x_{i2}, \ldots, x_{ip}, y_i \right)$, $i = 1, 2, \ldots, n$ with the relationship between predictor variables $y(x)$ and response variables $y$ assumed to follow a nonparametric regression model then the spline function $f(x)$ with degrees and knot points $K_1, K_2, \ldots, K_r$ can be written as follows:

$$f(x) = \sum_{j=1}^{p} \left( \sum_{h=0}^{d_j} \beta_{jh} x_i^h + \sum_{k=1}^{r} \beta_{rk} (x_i - K_k)^{q_k} \right)$$

with truncated function:
(x_i - K_{ik})^q = \begin{cases} 
(x_i - K_{ik})^q, & x_i \geq K_{ik} \\
0, & x_i < K_{ik} 
\end{cases}

In general, the spline truncated model can be presented in the form of equation (2) below:

\[ y_i = \sum_{l=1}^{p} \left( \sum_{h=0}^{q} \alpha_{hi} x_{li}^h + \sum_{k=1}^{r} \beta_{hi} (x_{li} - K_{ik})^q \right) + \varepsilon_i \]

Next, the model in equation (2) above can be written in the following form:

\[ y_i = \alpha_0 + \alpha_{11} x_{i1} + \alpha_{12} x_{i2} + \ldots + \alpha_{q-1p} x_{pi}^q + \alpha_{qp} x_{pi} + \beta_{11} (x_{i1} - K_{11})^q + \ldots + \beta_{pr} (x_{ip} - K_{pr})^q + \varepsilon_i \]

Equation (3) above can be written in matrix notation as follows:

\[ \mathbf{y} = \mathbf{X}(K) \mathbf{B} + \varepsilon \]

where,

\[ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X}(K) = \begin{pmatrix} 1 & x_{i1} & x_{i2} & \cdots & x_{i1}^q & x_{i1}^{q-1} & (x_{i1} - K_{11})^q & \cdots & (x_{ip} - K_{pr})^q \\ 1 & x_{21} & x_{22} & \cdots & x_{21}^q & x_{21}^{q-1} & (x_{12} - K_{11})^q & \cdots & (x_{2p} - K_{pr})^q \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{n1}^q & x_{n1}^{q-1} & (x_{in} - K_{11})^q & \cdots & (x_{np} - K_{pr})^q \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \alpha_0 \\ \alpha_{11} \\ \alpha_{12} \\ \cdots \\ \alpha_{q-1p} \\ \alpha_{qp} \\ \cdots \\ \beta_{11} \\ \beta_{pr} \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{pmatrix} \]

If \( q = 1 \), then equation (3) above can be written as a linear spline function with \( r \) knot points as follows:

\[ y_i = \alpha_0 + \alpha_{1i} x_{i1} + \alpha_{2i} x_{i2} + \ldots + \alpha_{pi} x_{pi} + \beta_{1} (x_{i1} - K_{11})_+ + \ldots + \beta_{pr} (x_{ip} - K_{pr})_+ + \varepsilon_i, \quad i = 1, 2, \ldots, n \]

2.2. Nonparametric Regression Spline Truncated for Longitudinal Data

The model of nonparametric regression spline truncated on equation (5) above is a form of model for cross-section data, but the model can be developed in longitudinal data [13]. In nonparametric spline truncated regression for longitudinal data there are as many mutually independent objects, where each object is observed at a certain time. In general, spline truncated nonparametric regression models for longitudinal data with one predictor variable can be written in the following form:

\[ y_{ij} = f(x_{ij}) + \varepsilon_{ij}, \quad i = 1, 2, \ldots, n; \; j = 1, 2, \ldots, t \]

In general, the spline truncated model on longitudinal data for one predictor variable can be presented in the form of equation (7) below:

\[ y_{ij} = \sum_{h=0}^{q} \alpha_{hi} x_{ij}^h + \sum_{k=1}^{r} \beta_{hi} (x_{ij} - K_{ki})^q + \varepsilon_{ij} \]

with truncated function:
\[(x_j - K_n)_j^2 = \begin{cases} (x_j - K_{ki})^q, & x_j \geq K_{(k+q)i} \\ 0, & x_j < K_{(k+q)i} \end{cases}\]

Equation (7) is a form of nonparametric regression spline truncated for longitudinal data with one predictor variable. If the predictor variable is used as much as \(p\), then the resulting model of nonparametric regression spline truncated multivariable for longitudinal data as follows:

\[y_{ij} = \sum_{l=0}^{p} \left( \sum_{h=0}^{q} \alpha_{hl} x_{ij}^h + \sum_{k=1}^{n} \beta_{kl} (x_{ij} - K_{ki})_l^q \right) + \epsilon_{ij}\]  

Equation (8) if written into matrix notation is as follows:

\[\mathbf{y} = \mathbf{X}^{(K)} \mathbf{B} + \mathbf{\epsilon}\]  

Based on equation model (9) estimator for parameters \(\mathbf{B}\) obtained by completing Weighted Least Square (WLS) optimization as follows:

\[
\min_{\mathbf{B} \in \mathbb{R}^{1(p+1)}} \left\{ \| \mathbf{y} - \mathbf{X}^{(K)} \mathbf{B} \| W^{-1} \left( \mathbf{y} - \mathbf{X}^{(K)} \mathbf{B} \right) \right\}
\]

by minimizing the square of error in equation (10) above, then an estimate is obtained \(\hat{\mathbf{B}}\) as follows:

\[
\hat{\mathbf{B}} = \left( \mathbf{X}^{(K)}^T \mathbf{W}^{-1} \mathbf{X}^{(K)} \right)^{-1} \mathbf{X}^{(K)}^T \mathbf{W}^{-1} \mathbf{y}
\]

where, \(\mathbf{y}\) is a response vector with the size \(nt \times 1\), \(\mathbf{X}^{(K)}\) is a predictor matrix of polynomial components and truncated components with the size \(nt \times n(1 + p(r + 1))\) with the parameters vector \(\mathbf{B}\) with the size \(n(1 + p(r + 1)) \times 1\) and \(\mathbf{\epsilon}\) is an error vector with the size \(nt \times 1\), while \(\mathbf{W}\) a weighting matrix (variance-covariance matrix) with the size \(nt \times nt\) and contains a diagonal \((W_1, W_2, \ldots, W_n)\).

2.3. Optimal Knot Point Selection

To get the best spline regression model, the optimal point is found that best fits the data. One method that is widely used in choosing the optimal knot point is Generalized Cross Validation (GCV). To obtain the optimal knot point can be seen from the minimum GCV value. The GCV method is generally defined as follows [13]:

\[
\text{GCV} = \frac{\text{MSE}}{\left( (nt)^{-1} (tr[\mathbf{I} - \mathbf{A}^{(K)}]) \right)^2} = \frac{(nt)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{t} (y_{ij} - \hat{y}_{ij})^2}{\left( (nt)^{-1} (tr[\mathbf{I} - \mathbf{A}^{(K)}]) \right)^2}
\]

where, \(\mathbf{A} = \mathbf{X}^{(K)} \mathbf{X}^{(K)}^T \mathbf{W}^{-1} \mathbf{X}^{(K)} \mathbf{X}^{(K)}^T \mathbf{W}^{-1}\)

3. Result and Discussion

This study uses secondary data in the form of annual financial reports from go public banks in Indonesia which are listed on the Indonesia Stock Exchange within period 2012-2018. The number of observations in this study are 30 go public banks in Indonesia which trades its shares on the Indonesia stock exchange during the study period. The list of go public banks companies in Indonesia is obtained
through the web site www.sahamok.com. The response variable used in this study is Return On Assets (ROA) which is influenced by several predictor variables which are defined as follows:

| Variable | Descriptions |
|----------|--------------|
| $y_{ij}$ | Return On Assets (ROA) ratio on the bank to $-i$, observation to $-j$ |
| $x_{ij1}$ | Capital Adequacy Ratio (CAR) ratio on the bank to $-i$, observation to $-j$ |
| $x_{ij2}$ | Non Performing Loan (NPL) ratio on the bank to $-i$, observation to $-j$ |
| $x_{ij3}$ | Net Interest Margin (NIM) ratio on the bank to $-i$, observation to $-j$ |
| $x_{ij4}$ | Operating Costs and Operating Income (BOPO) ratio on the bank to $-i$, observation to $-j$ |
| $x_{ij5}$ | Loan To Deposit (LDR) ratio on the bank to $-i$, observation to $-j$ |

3.1. Estimation of Error Variance-Covariance Matrix

Many researchers do modelling using nonparametric regression spline truncated for longitudinal data, but with the assumption that the weighting matrix is given, so in this research will be developed with the weighting matrix used is a variance-covariance matrix of errors. Given $\varepsilon_{ij}$ are random errors from the results of the prediction on the subject to $-i$, observation to $-j$ assumed to be multivariate normal distribution, with a mean $E(\varepsilon) = \mathbf{0}$ and variance-covariance matrix $Var(\varepsilon) = \mathbf{W}$, then using the Maximum Likelihood Estimator (MLE) method the estimate of the variance-covariance matrix error is obtained $\hat{\mathbf{W}_i}$ as follows:

$$\hat{\mathbf{W}_i} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \mathbf{X}_i (K) \mathbf{\hat{B})} (y_i - \mathbf{X}_i (K) \mathbf{\hat{B})}^T$$

(13)

3.2. Application of Nonparametric Regression Spline Truncated for Longitudinal Data on Return On Assets (ROA)

In modelling using regression, the main thing that needs to be done is to look at the relationship patterns of the data, so that the researcher can determine the right method to use in modelling. The following is a scatter plot pattern of the relationship between Return On Assets (ROA) with factors that are thought to be influential.

![Figure 1. Scatter plot ROA with CAR](image1)

![Figure 2. Scatter plot ROA with NPL](image2)
One of the pattern is very difficult to identify as a knot point. In this study, the selection of the best model is based on the use of Generalized Cross Validation (GCV), which is the selection of optimal knot points. The first step that needs to be considered in determining the selection of the best model is the analysis that has been done, the GCV values obtained as follows:

\[ \text{GCV} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]

where \( y_i \) is the observed value and \( \hat{y}_i \) is the predicted value.

### 3.3. The Optimal Knot Point Selection

The first step that needs to be considered in determining the selection of the best model is using Generalized Cross Validation (GCV), which is the selection of optimal knot points. In this study, the selection of knot points is based on the error variance-covariance matrix with one-knot point and 14 increments. Based on the data analysis that has been done, the GCV values are obtained as follows:

| Increment | Knot Values | GCV     |
|-----------|-------------|---------|
| \( K_{111} \) | \( K_{112} \) | \( K_{1304} \) | \( K_{1305} \) |       |
| 1         | 14.80       | 1.90    | ...    | 83.75  | 77.47  | 1739.59 |
| 2         | 15.86       | 2.03    | ...    | 84.74  | 77.95  | 31.46   |
| 3         | 16.91       | 2.15    | ...    | 85.73  | 78.44  | 25.25   |
| 4         | 17.97       | 2.28    | ...    | 86.72  | 78.92  | 22.42   |
The interpretation of the model above is as follows: Rakyat Indonesia (BRI) is obtained as follows: operational costs and operating income (BOPO), and loan to deposit (LDR), then the model of variables capital adequacy ratio (CAR), non-performing loans (NPL), net interest margin (NIM), exchange within period 2012-2018, namely bank of rakyat indonesia (BRI) with the predictor variable of capital adequacy ratio (CAR), non-performing loans (NPL), interest margin (NIM), and loan to deposit (LDR), then the model of nonparametric regression spline truncated for longitudinal data onReturn on Assets (ROA) at go public bank on the Indonesia stock exchange, with the highest average return on assets (ROA) at go public bank on the Indonesia stock exchange, with the highest average return on assets (ROA) at go public bank on the Indonesia stock exchange, with the highest average return on assets (ROA) at go public bank on the Indonesia stock exchange, with the highest average return on assets (ROA) at go public bank on the Indonesia stock exchange, with the highest average return on assets (ROA) at go public bank on the Indonesia stock exchange. Based on Table 2 above, the minimum GCV values obtained at 13 increment with one-knot point and weighting matrix is by using the error variance-covariance matrix with GCV value of 15.12 with MSE value of 0.739 and $R^2$ value of 85.56% which means that the best model in modelling Return on Assets (ROA) based on factors thought to be influential is a model of nonparametric regression spline truncated for longitudinal data with one-knot point and the weighting matrix used is the error variance-covariance matrix with the following equation:

$$y_{ij} = \alpha_i + \sum_{l=1}^{5} (\alpha_{il} x_{ijl} + \beta_{il} (x_{ijl} - K_{il})_+ + \beta_{2il} (x_{ijl} - K_{2il})_+ ) + \epsilon_{ij}$$

where, $i = 1, 2, \ldots, 30$ and $j = 1, 2, \ldots, 7$

3.4. Interpretation Models of Nonparametric Regression Spline Truncated for Longitudinal Data

Modelling using nonparametric regression spline truncated has a very good interpretation when the characteristics of the data are changing at certain sub-intervals marked by knot points. For an illustration, the following will explain the interpretation of the model with the selected bank is the bank with the highest average return on assets (ROA) at go public bank on the Indonesia stock exchange within period 2012-2018, namely Bank of Rakyat Indonesia (BRI) with the predictor variable of capital adequacy ratio (CAR), non-performing loans (NPL), net interest margin (NIM), operational costs and operating income (BOPO), and loan to deposit (LDR), then the model of nonparametric regression spline truncated for longitudinal data on return on assets (ROA) at Bank of Rakyat Indonesia (BRI) is obtained as follows:

$$\hat{y}_{7j} = 0.028 + 0.050 x_{7j1} + 0.169 x_{7j2} - 0.036 x_{7j3} - 0.091 x_{7j4} + 0.122 x_{7j5} - 0.041(x_{7j1} - 22.10)_+ + 0.027(x_{7j2} - 2.06)_+ - 0.088(x_{7j3} - 8.39)_+ - 0.211(x_{7j4} - 67.82)_+ - 0.276(x_{7j5} - 88.18)_+$$

The interpretation of the model above is as follows:

1) For the predictor variable of capital adequacy ratio (CAR)

$$\hat{y}_{7j} = 0.028 + 0.050 x_{7j1} \quad ; \quad x_{7j1} < 22.10$$

$$\hat{y}_{7j} = 0.934 + 0.009 x_{7j1} \quad ; \quad x_{7j1} \geq 22.10$$

2) For the predictor variable of non-performing loan (NPL)
\[ \hat{y}_{ij} = 0.028 + 0.169x_{ij2} \quad ; x_{ij2} < 2.06 \]
\[ \hat{y}_{ij} = -0.276 + 0.196x_{ij2} \quad ; x_{ij2} \geq 2.06 \]

3) For the predictor variable of Net Interest Margin (NIM)
\[ \hat{y}_{ij} = 0.028 - 0.036x_{ij3} \quad ; x_{ij3} < 8.39 \]
\[ \hat{y}_{ij} = 0.766 + 0.124x_{ij3} \quad ; x_{ij3} \geq 8.39 \]

4) For the predictor variable of Operational Costs and Operating Income (BOPO)
\[ \hat{y}_{ij} = 0.028 - 0.091x_{ij4} \quad ; x_{ij4} < 67.82 \]
\[ \hat{y}_{ij} = 14.338 - 0.302x_{ij4} \quad ; x_{ij4} \geq 67.82 \]

5) For the predictor variable of Loan To Deposit (LDR)
\[ \hat{y}_{ij} = 0.028 + 0.122x_{ij5} \quad ; x_{ij5} < 88.18 \]
\[ \hat{y}_{ij} = 24.366 - 0.154x_{ij5} \quad ; x_{ij5} \geq 88.18 \]

Based on the Bank of Rakyat Indonesia (BRI) model which is used as an illustration in this interpretation, it can be concluded that when the Capital Adequacy Ratio (CAR) of Bank of Rakyat Indonesia (BRI) is less than 22.10 and if the level of the Capital Adequacy Ratio (CAR) rises by one percent then Return On Assets (ROA) increases by 0.05 assuming the other predictor variables are fixed. Next, if the Capital Adequacy Ratio (CAR) of Bank of Rakyat Indonesia (BRI) is greater than or equal to 22.10 and if the Capital Adequacy Ratio (CAR) rises by one percent then Return On Assets (ROA) decreases by 0.169 assuming the other predictor variables are fixed. For variable Non Performing Loan (NPL) of Bank of Rakyat Indonesia (BRI) is less than 2.06 and if the level of the Non Performing Loan (NPL) rises by one percent then Return On Assets (ROA) increases by 0.169 assuming the other predictor variables are fixed. Next, if the Non Performing Loan (NPL) of Bank of Rakyat Indonesia (BRI) is greater than or equal to 2.06 and if the Non Performing Loan (NPL) rises by one percent then Return On Assets (ROA) decreases by 0.196 assuming the other predictor variables are fixed. For the other predictor variables and the other bank's subject can be interpreted with the same interpretation.

4. Conclusion
Based on the analysis described above, it can be concluded that modelling nonparametric regression spline truncated for longitudinal data on Return On Assets (ROA) based on factors that are suspected to be influential with using 30 subjects of go public banks listed on the Indonesia stock exchange during 2012-2018. Obtained the best model chosen based on the weighting matrix of the error variance-covariance matrix and the number of knot point is one with a minimum GCV value of 15.12, MSE value of 0.739 and $R^2$ 85.56%. The weighting matrix which is a variance-covariance matrix of the error can be determined by the following formula:

\[ \mathbf{W}_i = \frac{1}{n} \sum_{i=1}^{n} \left( \mathbf{y}_i - \mathbf{X}_i \mathbf{K} \mathbf{\hat{B}} \right) \left( \mathbf{y}_i - \mathbf{X}_i \mathbf{K} \mathbf{\hat{B}} \right)^T \]

References
[1] Eubank, R. L. 1999. Nonparametric Regression and Spline Smoothing Second Edition. Marcel Dekker. New York.
[2] Budiantara, I. N. 2000. “Metode U, GML, CV dan GCV Dalam Regresi Nonparametrik Spline”. Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI). No. 06, pp. 285-290.
[3] Hardle, W. 1990. Applied Nonparametric Regression. Cambridge University Press. New York.
[4] Wahba, G. 1990. Spline Model for Observational Data, SIAM XII, Philadelphia.
[5] Wisisono, I. N. R., Nurwahidah, A. I., and Andriyana, Y. 2018. “Regresi Nonparametrik dengan
Pendekatan Deret Fourier pada Data Debit Air Sungai Citarum”. Jurnal Matematika. Vol. 04, No. 02, pp. 75-82.

[6] Antomiadis, A., Bigot., and Spatinas, T. 2001. “Wavelet Estimator in Nonparametric Regression: A Comparative Simulation Study”. Journal of Statistical Software. Vol. 06, pp. 1-83.

[7] Budiantara, I. N. 2009. Spline Dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang. Institut Teknologi Sepuluh November. Surabaya.

[8] Budiantara, I. N., Lestari, B., dan Islamiyati, A. 2010. “Estimator Spline Terbobot dalam Regresi Nonparametrik dan Semiparametrik Heterokedastisitas untuk Data Longitudinal”. Hibah Penelitian Kompetensi. DP2M-Dikti: Jakarta.

[9] Saputro, D. R. S., Demu, K. R., and Widyaningsih, P. 2018. “Nonparametric Truncated Spline Regression on the Data of Human Development Index (HDI) in Indonesia”. 2th International Conference on Statistics, Mathematics, Teaching, and Research. IOP Publishing.

[10] Ratnasari, V., Budiantara, I. N., Ratna, M., and Zain, I. 2016. “Estimation of Nonparametric Regression Curve Using Mixed Estimator of Multivariable Truncated Spline and Multiviable Kernel”. Global Journal of Pure & Applied Mathematics. Vol. 12, No. 06, pp. 5047-5057.

[11] Sudiarsa, I. W., Budiantara, I. N., Suhartono, and Purnami, S. W. 2016. “Combined Estimator Fourier Series and Spline Truncated in Multivariable Nonparametric Regression”. Applied Mathematics Science. Vol. 09, No. 100, pp. 4997-5010.

[12] Diggle, P. J., Heagerty, P., Liang, K. Y., and Zeger, S. L. 2002. Analysis of Longitudinal Data 2nd Edition. Oxford Statistical Science 25. New York.

[13] Wu, H. and Zhang, J. T. 2006. Nonparametric Regression Methods for Longitudinal Data Analysis. A John Wiley and Son Inc. New Jersey.

[14] Fernandes, A. A. R., Budiantara, I. N., Otok, B. W., and Suhartono. 2015. “Spline Estimator for Bi-Responses and Multi-Predictors Nonparametric Regression Model in Case of Longitudinal Data”. Journal of Mathematics and Statistics. Vol. 11, No. 02, pp. 61-69.

[15] Damaliana, A. T., Budiantara, I. N., and Ratnasari, V. 2019. “Comparing between mGCV and aGCV Methods to Choose the Optimal Knot Points in Semiparametric Regression with Spline Truncated Using Longitudinal Data”. 9th Annual Basic Science International Conference. IOP Publishing.