MODELING PERCENTAGE OF POOR PEOPLE IN INDONESIA USING KERNEL AND FOURIER SERIES MIXED ESTIMATOR IN NONPARAMETRIC REGRESSION

I Nyoman Budiantara, Vita Ratnasari, Madu Ratna, Wahyu Wibowo, Ngizatul Afifah, Dyah Putri Rahmawati and Made Ayu Dwi Octavanny[1]
Department of Statistics, Faculty of Mathematics, Computation and Data Sciences.
Institut Teknologi Sepuluh Nopember (ITS), Surabaya, 60111, Indonesia

ABSTRACT

Poverty is a very serious problem and often faced by the countries in the world, especially developing countries. The percentage of poor people in Indonesia reached 11.47 percent in 2013. The seven provinces with the highest poverty in Indonesia are Papua, West Papua, East Nusa Tenggara, Maluku, Gorontalo, Bengkulu and Aceh. This problem is modeled using mixed nonparametric regression of Kernel and Fourier Series. The response variable of this model is percentage of poor people (y), the predictor variables that follow Kernel regression curve are Mean of Years Schooling or MYS (v_1) and Literacy Rate or LR (v_2), whereas the predictor variable that follow the Fourier Series regression curve are Unemployment Rate or UR (t_1). This modeling produces $R^2 = 62.78\%$.

KEYWORDS: Fourier Series, Kernel, Mixed Nonparametric Regression, Percentage of Poor People.

MSC: 37M10

1. INTRODUCTION

Nonparametric regression is one of the regression analysis that has high flexibility. In this method, data is expected to find its own form of regression curve estimation without being influenced by the subjectivity of the researcher [1]. There are many types of estimators in nonparametric regression models such as Kernel, Spline, Local Polynomial, Wavelet and Fourier Series. Studies on Kernel estimator have been carried out by researchers such as Okumura and Naito [2], Yao [3], Kayri and Zirhlioglu [4], Cheng, Paige, Sun and Yan [5], and Fernandes, Budiantara, Otok and Suhartono, [6]. Research on Spline estimators have been applied by other researchers such as Eubank [1], Merdekawati and Budiantara [7], Darmawi and Otok [8] and Fernandes, Budiantara, Otok, and Suhartono [9]. Research on Local Polynomial estimators is developed by Su and Ullah [10], He and Huang [11], Filho and Yao [12] and Qingguo [13]. Research on the Wavelet estimator is developed by Antoniadis, Bigot and Sapatinas [14], Amato and De Canditiis [15], and Taylor [16]. Fourier Series estimators has been applied by researchers such as Bilodeau [17], Galtchouk and Pergamenshchikov [18], Ratnasari, Budiantara, Zain, Ratna and Mariati [19], Pane, Budiantara, Zain and Otok [20] and Asrini and Budiantara [21]. Beside the above studies, there are studies about mixed nonparametric regression too, that are combination of Spline and Kernel by Budiantara, Ratnasari, Ratna and Zain [22] and mixed nonparametric regression of Kernel and Fourier Series by Afifah, Budiantara and Latra [23].
Nonparametric regression model can be applied in various fields of science, one of them is social field. In this field, nonparametric regression models can be applied to analyze poverty issues. In 2013, the percentage of poor people in Indonesia reached 11.47 percent. There are seven provinces with the highest percentage of

[1] nyomanbudiantara65@gmail.com, vita_ratna@statistika.its.ac.id, madu_r@statistika.its.ac.id, wahyu_w@statistika.its.ac.id, ngizatul.afifah@gmail.com, dyahputri1234@gmail.com,octavanny@gmail.com

poor people in Indonesia, that are Papua, West Papua, East Nusa Tenggara, Maluku, Gorontalo, Bengkulu and Aceh [24]. This issue is interesting to discuss further when we observe the scatter plots of variables that affect the percentage of poverty in the seven provinces. It found that some of these variables form Kernel regression patterns and some others form the pattern of Fourier Series. Thus the problem of the percentage of poor people in Indonesia can be investigated using mixed nonparametric regression of Kernel and Fourier Series which have been developed previously by Afifah, et al [23].

2. MATERIALS AND METHODS

2.1.1 Data Source

This study uses secondary data in 2013 from publications by the Statistics of Indonesia (BPS) with observation units as many as 102 regency/city in 7 province that are Papua, West Papua, East Nusa Tenggara, Maluku, Gorontalo, Bengkulu, and Aceh. The respon variable in this study is percentage of poor people in 102 regency/city in 2013 (*Y*). The predictor variables are Mean of Years Schooling or MYS (X_1), Unemployment Rate or UR (X_2) and Literacy Rate or LR (X_3).

2.2. Analysis Stage

The steps to solve the problems and achieve the aim of this study are as follows:

1) Conducting descriptive data analysis to know the general description of poverty data in Indonesia.
2) Create a scatter plot of data between response variables with each predictor variables.
3) Determining predictor variables which are nonparametric component approximated by Kernel function and approximated by Fourier Series function.
4) Modeling poverty in Indonesia using mixed nonparametric of Kernel and Fourier Series.
5) Determine optimal bandwidth, oscillation parameters, and smoothing parameter using Generalized Cross Validation (GCV) method.
6) Obtain the estimation of the mixed nonparametric regression curve.
7) Calculating the value of MSE and $R^{2.}$
8) Drawing conclusions.

2.3. Fourier Series Function

Fourier series function are widely used in various fields of science. In the field of statistics the Fourier Series function is used to modeling data behavior that follows periodic pattern or seasonal. Let $h \in C(0,\pi)$ is any function where

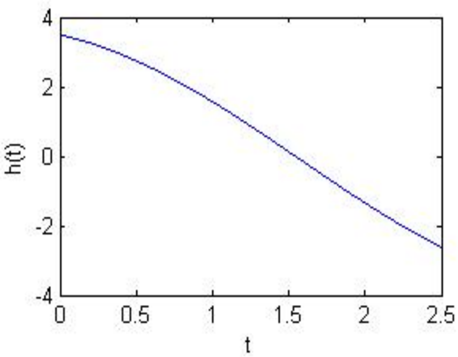
$$C(0,\pi) = \left\{ h \,;\, h \text{ continous function in } 0 < t < \pi \right\}.$$

Thus according to Bilodeau [17], Ratnasari, Budiantara, Zain, Ratna and Mariati [19], Pane, Budiantara, Zain and Otok [20], and Asrini and Budiantara [21], the function of h can be presented in the form :

$$h(t) = bt + \frac{1}{2}a_0 + \sum_{k=1}^{M} a_k \cos kt$$

with b, a_0, $a_1, a_2, ..., a_M$ are real number, and $M \in Z^+ = \left\{ 1,2,3,... \right\}$ is oscillation parameter. The value of b shows the trend of the Fourier Series function. If the value of $b = 0$, then it will be obtained the non trend Fourier Series function. Otherwise if the value of $b > 0$, it will be obtained up-trend of Fourier Series function. Whereas , if the value of $b < 0$, it will be obtained down-trend of Fourier Series function. Figure 1, Figure 2, and Figure 3 respectively present the Fourier Series function with oscillation parameter are M=1, M=5 and M=10, for $b < 0$ (down-trend).

It can be seen visually from Figure 1, Figure 2 and Figure 3, the form of Fourier Series function exceptionally depend on the value of oscillation parameter M. If the value of the oscillation parameter M is larger than the form of Fourier Series function, it will have higher oscillation. Conversely, if the value of the oscillation parameter M is smaller than the form of Fourier Series function, it has no oscillation



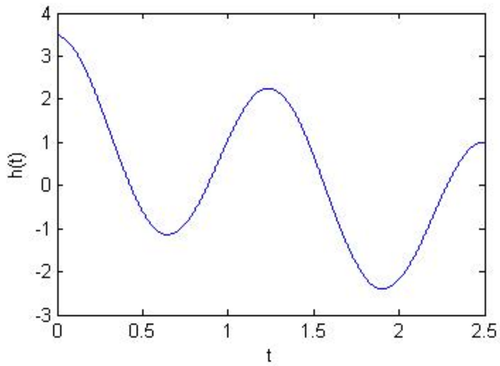


Figure 1. Fourier Series Function with *M*=1 **Figure 2**. Fourier Series Function with *M*=5

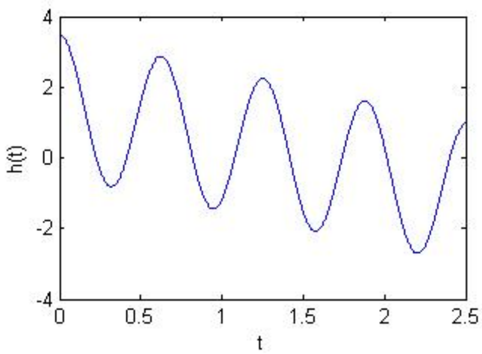


Figure 3. Fourier Series Function with *M*=10.

2.4 . Nonparametric Regression of Mixed Kernel and Deret Fourier

Afifah, et al. [23] have written about the mixed estimation of Kernel and Fourier Series in nonparametric regression. Given paired data $\left(v_{1i},...,v_{pi},t_{1i},...,t_{qi},y_i\right),\ \ i=1,2,...,n,$ following the mixed regression model i.e.

$$y_i=\mu\left(v_{1i},...,v_{pi},t_{1i},...,t_{qi}\right)+\varepsilon_i \qquad =\mu\left(\mathbf{v}_i,\mathbf{t}_i\right)+\varepsilon_i \tag{1}$$

where $\mathbf{v}_i=(v_{1i},...,v_{pi})'$ and $\mathbf{t}_i=(t_{1i},t_{2i}\ldots,t_{qi})$. The form of regression curve $\mu=(\mathbf{v}_i,\mathbf{t}_i)$ in the Equation (1) is assumed to be unknown and the curve is smooth in the continuous and differentiable. Random error ε_i is a normal distribution with mean 0 and variance σ^2. The regression curve $\mu\left(\mathbf{v}_i,\mathbf{t}_i\right)$ is assumed to be additive, so it can be written as:

$$\mu\left(\mathbf{v}_i,\mathbf{t}_i\right)=\sum_{j=1}^{p}g_j\left(v_{ji}\right)+\sum_{s=1}^{q}h_s\left(t_{si}\right) \tag{2}$$

The component:

$$\sum_{j=1}^{p}g_j\left(v_{ji}\right)$$

is Nadaraya Watson Kernel regression curve as follows:

$$\hat{g}_{\phi}\left(v\right)=n^{-1}\sum_{i=1}^{n}\frac{K_{\phi}\left(v-v_i\right)}{n^{-1}\sum_{i=1}^{n}K_{\phi}\left(v-v_i\right)}y_i \quad =n^{-1}\sum_{i=1}^{n}W_{\phi i}\left(v\right)y_i \tag{3}$$

which $\hat{g}_{\phi}\left(v\right)$ is Kernel regression estimation function and ϕ is the bandwidth. Function $W_{\phi i}\left(v\right)$ is a weighted function:

$$W_{\phi i}\left(v\right)=\sum_{i=1}^{n}\frac{K_{\phi}\left(v-v_i\right)}{n^{-1}\sum_{i=1}^{n}K_{\phi}\left(v-v_i\right)}$$

where $K_{\phi}\left(v-v_i\right)$ is kernel function with

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

In this research, K is Gaussian kernel function:

$$K\left(z\right)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right),\quad -\infty<z<\infty$$

While the component:

$$\sum_{s=1}^{q}h_s\left(t_{si}\right)$$

is Fourier series component. It approximated by the function $T_s\left(t_{si}\right)$ i.e.

$$T_s\left(t_{si}\right)=b_s t_{si}+\frac{1}{2}a_{0s}+\sum_{k=1}^{M}a_{ks}\cos kt_{si}. \tag{4}$$

The mixed estimator of kernel and fourier series in nonparametric regression (1) are given by Theorem 1 as follows:

Theorem 1. If in the Equation (1) the kernel regression component estimator is $\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)$ and Fourier series estimator is $\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)$ then the mixed estimator $\boldsymbol{\mu}\left(\mathbf{v}_i,\mathbf{t}_i\right)$ is given by $\hat{\boldsymbol{\mu}}_{\boldsymbol{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)=\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)+\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)$ where $\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)=\mathbf{V}\left(\boldsymbol{\Phi}\right)\mathbf{y}$, $\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)=\mathbf{X}\hat{\mathbf{a}}\left(\boldsymbol{\lambda}\right)$, with $\hat{\mathbf{a}}\left(\lambda\right)=C\left(\boldsymbol{\Phi},\lambda,M\right)\mathbf{y}$, $C\left(\boldsymbol{\Phi},\lambda,M\right)=\left(\mathbf{X}'\mathbf{X}+n\mathbf{D}\left(\boldsymbol{\lambda}\right)\right)^{-1}\mathbf{X}'\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right)$. So Equation (1) can be written as follows:

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)=\mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\mathbf{y}.$$

Proof.
If component of kernel function in the Equation (2) is estimated by the estimator component in the Equation (3), then

$$\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)=\mathbf{V}\left(\boldsymbol{\Phi}\right)\mathbf{y} \tag{5}$$

where

$$\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)=\begin{bmatrix}\hat{g}_{j\phi_j}\left(v_{j1}\right)\\ \hat{g}_{j\phi_j}\left(v_{j2}\right)\\ \mathrm{M}\\ \hat{g}_{j\phi_j}\left(v_{jn}\right)\end{bmatrix},\ \mathbf{y}=\begin{bmatrix}y_1\\ y_2\\ \mathrm{M}\\ y_n\end{bmatrix},\ \text{and }\mathbf{V}\left(\boldsymbol{\Phi}\right)=\begin{bmatrix}n^{-1}\sum_{j=1}^{p}W_{\phi_j 1}\left(v_{j1}\right) & \mathrm{K} & n^{-1}\sum_{j=1}^{p}W_{\phi_j n}\left(v_{j1}\right)\\ n^{-1}\sum_{j=1}^{p}W_{\phi_j 1}\left(v_{j2}\right) & \mathrm{K} & n^{-1}\sum_{j=1}^{p}W_{\phi_j n}\left(v_{j2}\right)\\ \mathrm{M} & \mathrm{O} & \mathrm{M}\\ n^{-1}\sum_{j=1}^{p}W_{\phi_j 1}\left(v_{jn}\right) & \mathrm{K} & n^{-1}\sum_{j=1}^{p}W_{\phi_j n}\left(v_{jn}\right)\end{bmatrix},$$

vector $\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)$ is $n\times 1$, vector $\mathbf{y}$ is $n\times 1$, and matrix $\mathbf{V}(\boldsymbol{\Phi})$ is $n\times n$.

If Fourier series component in the Equation (2) is approximated by Equation (4) then

$$\sum_{s=1}^{q}\mathbf{T}_s\left(t_s\right)=\mathbf{Xa}) \tag{6}$$

Where

$$\mathbf{X}=\begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_p\end{bmatrix},\ \mathbf{a}=\begin{bmatrix}\mathbf{a}_1' & \mathbf{a}_2' & \cdots & \mathbf{a}_p'\end{bmatrix}',\ \mathbf{T}_s\left(t_s\right)=\begin{bmatrix}T_s\left(t_{s1}\right) & T_s\left(t_{s2}\right) & \cdots & T_s\left(t_{sn}\right)\end{bmatrix}',$$

$$\mathbf{X}_s=\begin{bmatrix}t_{s1} & 1 & \cos t_{s1} & \cos 2t_{s1} & \cdots & \cos Mt_{s1}\\ t_{s2} & 1 & \cos t_{s2} & \cos 2t_{s2} & \cdots & \cos Mt_{s2}\\ t_{s3} & 1 & \cos t_{s3} & \cos 2t_{s3} & \cdots & \cos Mt_{s3}\\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots\\ t_{sn} & 1 & \cos t_{sn} & \cos 2t_{sn} & \cdots & \cos Mt_{sn}\end{bmatrix},\ \text{and } \mathbf{a}_s=\begin{pmatrix}b_s\\ \frac{1}{2}a_{0s}\\ a_{1s}\\ \vdots\\ a_{Ms}\end{pmatrix}$$

The matrix $\mathbf{X}$ is $n\times\left(p\left(M+2\right)\right)$, the vector $\mathbf{a}$ is $p\left(M+2\right)\times 1$, the matrix $\mathbf{X}_s$ is $n\times\left(M+2\right)$, the vector $\mathbf{a}_s$ is $\left(M+2\right)\times 1$, and M is the oscillation parameter.

If the regression model is given by Equation (1), the kernel function is given by Equation (5) and the Fourier series function is given by Equation (6), then goodness of fit function is

$$\begin{aligned}R\left(\mathbf{a}\right)&=n^{-1}\left\|\mathbf{y}-\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)-\sum_{s=1}^{q}\mathbf{T}_s\left(t_s\right)\right\|^2\\ &=n^{-1}\left\|\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right)\mathbf{y}-\mathbf{Xa}\right\|^2\end{aligned} \tag{7}$$

where I is $n\times n$ identity matrix.

If the Fourier series function is given by Equation (6), then penalty function and smoothing parameter for the Penalized Least Square (PLS) method are

$$\sum_{s=1}^{q}\lambda_s J\left(T_s\left(t_{si}\right)\right)=\mathbf{a}'\mathbf{D}\left(\boldsymbol{\lambda}\right)\mathbf{a} \tag{8}$$

where λ_s is smoothing parameter, $\mathbf{D}\left(\boldsymbol{\lambda}\right)=diag\begin{pmatrix}\mathbf{D}_1\left(\boldsymbol{\lambda}\right) & \mathbf{D}_2\left(\boldsymbol{\lambda}\right) & \cdots & \mathbf{D}_p\left(\boldsymbol{\lambda}\right)\end{pmatrix}$, and

$\mathbf{D}_s\left(\boldsymbol{\lambda}\right)=diag\begin{pmatrix}0 & 0 & \lambda_s 1^4 & \lambda_s 2^4 & \cdots & \lambda_s M^4\end{pmatrix}$. The matrix $\mathbf{D}\left(\boldsymbol{\lambda}\right)$ is $\left(p\times\left(M+2\right)\right)\times\left(p\times\left(M+2\right)\right)$ and the matrix $\mathbf{D}_s\left(\boldsymbol{\lambda}\right)$ is $\left(M+2\right)\times\left(M+2\right)$.

If the goodness of fit function is given by Equation (7) while penalty function and smoothing parameter are given by Equation (8), then estimation for parameter $\mathbf{a}$ is [25]:

$$\begin{aligned}&\underset{\mathbf{a}}{Min}\left\{R\left(\mathbf{a}\right)+\sum_{s=1}^{q}\lambda_s J\left(T_s\left(t_{si}\right)\right)\right\}\\ &=\underset{\mathbf{a}}{Min}\left\{n^{-1}\left\|\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right)\mathbf{y}-\mathbf{Xa}\right\|^2+\mathbf{a}'\mathbf{D}\left(\boldsymbol{\lambda}\right)\mathbf{a}\right\}\\ &=\underset{\mathbf{a}}{Min}\{n^{-1}\mathbf{y}'\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right)'\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right)\mathbf{y}\\ &\quad-2n^{-1}\mathbf{a}'\mathbf{X}'\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right)\mathbf{y}+\mathbf{a}\left(n^{-1}\mathbf{X}'\mathbf{X}+\mathbf{D}\left(\boldsymbol{\lambda}\right)\right)\mathbf{a}\}\end{aligned}$$

Therefore estimator for $\mathbf{a}$ is

$$\hat{\mathbf{a}}\left(\lambda\right)=C\left(\boldsymbol{\Phi},\lambda,M\right)\mathbf{y}. \tag{9}$$

Based on Equation (9), we obtained estimator for fourier series, that is:

$$\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)=\mathbf{X}\hat{\mathbf{a}}\left(\boldsymbol{\lambda}\right). \tag{10}$$

Based on Equation (10), we obtained

$$\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_{si}\right)=\mathbf{S}\left(\mathbf{\Phi},\boldsymbol{\lambda},M\right)\mathbf{y} \tag{11}$$

where $\mathbf{S}\left(\mathbf{\Phi},\boldsymbol{\lambda},M\right)=\mathbf{X}\left(\mathbf{X}'\mathbf{X}+n\mathbf{D}(\boldsymbol{\lambda})\right)^{-1}\mathbf{X}'\left(\mathbf{I}-\mathbf{V}\left(\mathbf{\Phi}\right)\right).$

Using Equation (9) and Equation (5) the mixed estimator for nonparametric regression is

$$\hat{\boldsymbol{\mu}}_{\mathbf{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)=\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)+\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right) \tag{12}$$

Equation (12) can be written:

$$\hat{\boldsymbol{\mu}}_{\mathbf{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)=\mathbf{Z}\left(\mathbf{\Phi},\boldsymbol{\lambda},M\right)\mathbf{y} \tag{13}$$

with $\mathbf{Z}\left(\mathbf{\Phi},\boldsymbol{\lambda},M\right)=\left(\mathbf{V}\left(\mathbf{\Phi}\right)+\mathbf{S}\left(\mathbf{\Phi},\boldsymbol{\lambda},M\right)\right).$

Furthermore, in this section we deal with the properties of estimator by fourier series component, kernel component and mixed of fourier series and kernel in nonparametric regression. The properties of those estimator are given by Lemma 1.

Lemma 1. If the estimator of fourier series component is $\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)$, estimator of kernel component is: $\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)$ and mixed estimator of kernel and Fourier series is $\hat{\boldsymbol{\mu}}_{\mathbf{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)$ as given by Theorem 1, then each of the estimator are biased estimator, but still are classified as linear estimator in y observation.

Proof

Expectation of Fourier Series component estimator can be written as:

$$E\left(\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)\right)=E\left(\mathbf{X}\hat{\mathbf{a}}\left(\boldsymbol{\lambda}\right)\right)=E\left(\mathbf{X}C\left(\mathbf{\Phi},\lambda,M\right)\mathbf{y}\right)=\mathbf{X}C\left(\mathbf{\Phi},\lambda,M\right)E(\mathbf{y})$$

$$\mathbf{X}C\left(\mathbf{\Phi},\lambda,M\right)\left(\sum_{j=1}^{p}g_j\left(v_j\right)+\sum_{s=1}^{q}h_s\left(t_s\right)\right)=\mathbf{X}C\left(\mathbf{\Phi},\lambda,M\right)\sum_{j=1}^{p}g_j\left(v_j\right)+\mathbf{X}C\left(\mathbf{\Phi},\lambda,M\right)\sum_{s=1}^{q}h_s\left(t_s\right)$$

$$\neq\sum_{s=1}^{q}h_s\left(t_s\right).$$

Moreover, the estimator of Fourier series component could be presented as follows:

$$\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)=\mathbf{X}C\left(\mathbf{\Phi},\lambda,M\right)\mathbf{y}.$$

thus, the form shows that the estimator of the Fourier series component is linearity in y observation.
Furthermore, the expectation value of estimator kernel component can be written as:

$$E\left(\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)\right)=E\left(\mathbf{V}\left(\mathbf{\Phi}\right)\mathbf{y}\right)=\mathbf{V}\left(\mathbf{\Phi}\right)E(\mathbf{y})=\mathbf{V}\left(\mathbf{\Phi}\right)\sum_{j=1}^{p}g_j\left(v_j\right)+\mathbf{V}\left(\mathbf{\Phi}\right)\sum_{s=1}^{q}h_s\left(t_s\right)\neq\sum_{j=1}^{p}g_j\left(v_j\right)$$

The estimator of the kernel component could also be presented as:

$$\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)=\mathbf{V}\left(\mathbf{\Phi}\right)\mathbf{y}. \tag{14}$$

Based on form estimator of kernel component in Equation (14) it shown that the estimator of kernel component is a linear estimator in y observation.
Finally, the mixed estimator of Fourier series and kernel can be presented in the form of

$$\hat{\boldsymbol{\mu}}_{\mathbf{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v},\mathbf{t}\right)=\mathbf{Z}\left(\mathbf{\Phi},\boldsymbol{\lambda},M\right)\mathbf{y}$$

it shown that the mixed estimator of fourier series and kernel is a linear estimator in y observation.
Otherwise, expectation of the mixed estimator of fourier series and kernel can be written as follows:

$$
\begin{aligned}
E\left(\hat{\boldsymbol{\mu}}_{\boldsymbol{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v},\mathbf{t}\right)\right) &= E\left(\mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\mathbf{y}\right)\\
&= \mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)E(\mathbf{y})\\
&= \mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\left(\sum_{j=1}^{p} g_j\left(v_j\right)+\sum_{s=1}^{q} h_s\left(t_s\right)\right)\\
&= \mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\left(\sum_{j=1}^{p} g_j\left(v_j\right)+\sum_{s=1}^{q} h_s\left(t_s\right)\right)\\
&\neq \left(\sum_{j=1}^{p} g_j\left(v_j\right)+\sum_{s=1}^{q} h_s\left(t_s\right)\right) = \mu\left(\mathbf{v},\mathbf{t}\right).
\end{aligned}
$$

thus, the estimator of the Fourier series component, the estimator of the kernel component and the mixed estimator of the Fourier series and the kernel are biased estimator, and the linear estimator in y the observation.

Mixed estimators $\hat{\boldsymbol{\mu}}_{\boldsymbol{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)$ dependent on bandwidth $\boldsymbol{\Phi}=\left(\phi_1,\phi_2,K\ ,\phi_p\right)'$, oscillation parameters M, and smoothing parameters $\boldsymbol{\lambda}=\left(\lambda_1,\lambda_2,K\ ,\lambda_q\right)'$. If the bandwidth parameter and the smoothing parameter are very small $\phi_j \to 0,\ j=1,2,..,p,\ \ \lambda_s \to 0,\ s=1,2,..,q$ and the oscillation parameter M are very large $(M \to \infty)$ then the mixed estimator of the Fourier Series and kernel will be very roughness. Conversely if the bandwidth parameter and the smoothing parameter are very large $\phi_j \to \infty,\ \ j=1,2,..,p,\ \ \lambda_s \to \infty,\ \ s=1,2,..,q$ and the oscillation parameter M is very small, then the mixed estimator of Fourier series and kernel will be very smooth.

The best of mixed estimator of the Fourier series and the kernel is the estimator that contains the optimal bandwidth, the smoothing parameters and the oscillation parameter. To obtain optimal bandwidth, smoothing parameter and oscillation parameter, we can use various methods, such as Cross Validation (CV) method, Unbiased Risk (UBR) method, Generalized Cross Validation (GCV) method or Generalized Maximum Likelihood (GML) method.

In this study the choosing of the optimal bandwidth, smoothing parameter and oscillation parameter, done by using generalization of GCV method, this methods are given by Wahba [26], Budiantara [27], Ratnasari, et. al [28] and Budiantara, et al [22].

Mixed estimators $\hat{\boldsymbol{\mu}}_{\boldsymbol{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)$ dependent on bandwidth $\boldsymbol{\Phi}=\left(\phi_1,\phi_2,K\ ,\phi_p\right)'$, oscillation parameters ***M***, and smoothing parameters $\boldsymbol{\lambda}=\left(\lambda_1,\lambda_2,K\ ,\lambda_q\right)'$. GCV method can provide the optimal bandwidth, smoothing parameter and oscillation parameter, that is:

$$
\mathrm{GCV}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)=\frac{\mathrm{MSE}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)}{\left(n^{-1}\mathrm{trace}\left(\mathbf{I}-\mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\right)\right)^2} \tag{15}
$$

where $\mathrm{MSE}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)=n^{-1}\mathbf{y}'\left(\mathbf{I}-\mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\right)'\left(\mathbf{I}-\mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\right)\mathbf{y}$. The optimal bandwidth, smoothing parameter and oscillation parameter are obtained from the smallest GCV values. The other word, the optimal bandwidth, smoothing parameter and oscillation parameter can be obtained by optimization:

$\mathrm{GCV}\left(\boldsymbol{\Phi}_{\mathrm{opt}},\boldsymbol{\lambda}_{\mathrm{opt}},M_{\mathrm{opt}}\right)=\underset{\boldsymbol{\Phi},\boldsymbol{\lambda},M}{\mathrm{Min}}\left(\dfrac{\mathrm{MSE}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)}{\left(n^{-1}\mathrm{trace}\left(\mathbf{I}-\mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\right)\right)^2}\right)$. According to Wahba [26], the GCV method has an advantage. The advantage of the GCV method is GCV method has asymptotically optimal properties that it is not be possessed by other method

3. RESULT AND DISCUSSION

In 2013, there were 7 provinces with highest percentage of poorer people compared to other provinces in Indonesia, which are Papua (31.53 percent), West Papua (27.14 percent), East Nusa Tenggara (20.24 percent), Maluku (19.27 percent), Gorontalo (18.01 percent), Bengkulu (17.75 percent) and Aceh (17.72 percent). Relationship pattern between the response variable with each predictor variable can be seen from the scatter plot chart. The scatter plot results for each response variable and predictor variable as shown in Figure 4, Figure 5 and Figure 6.

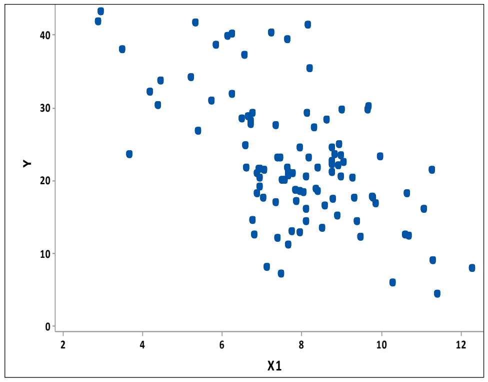


Figure 4. Scatter Plot Percentage of Poor People (Y) *vs* Mean of Years Schooling (X_1)

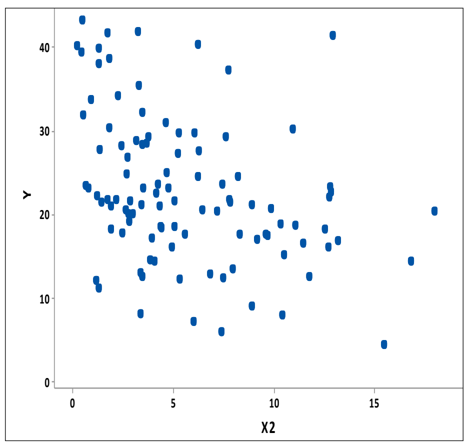


Figure 5. Scatter Plot Percentage of Poor People (Y) *vs* Unemployment Rate (X_2)

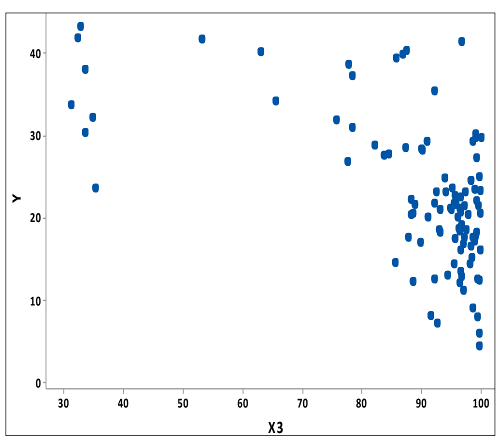


Figure 6. Scatter Plot Percentage of Poor People (Y) *vs* Literacy Rate (X_3).

Figure 4 and Figure 6 show that the pattern of relationships between response variables and predictor variables that have no particular pattern. Therefore, Mean of Years Schooling (MYR) (X_1) and Literacy Rate

(LR) (X_3) variables can be approached by nonparametric kernel estimator. Figure 5 shows that the pattern of relationship between response variable and predictor variable tends to be repetive at 0.44; 3.20; 6.22; 7.70; 12.89; 12.79 and 17.97 then 0.50; 1.15; 3.34; 5.97; 7.35; 10.40 and 15.46 with a downward trend. Thus theoretically, Unemployment Rate (UR) (X_2) variable can be approached by Fourier series.
In addition, we can use GCV to determine which variables are approached by the kernel or the Fourier series. GCV values can be obtained from Equation (15). Based on the calculation results obtained GCV values, for all possible mixed models of kernels and Fourier series as shown in Table 1 and Table 2.

Table 1. GCV of Models with 2 Kernels and 1 Fourier Series

No.	Variable		GCV
	Kernel	Deret Fourier	
1	X_2, X_3	X_1	46.332
2	X_1, X_3	X_2	44.166
3	X_1, X_2	X_3	48.261

Table 2. GCV of Models with 1 Kernel and 2 Fourier Series

No.	Variable		GCV
	Kernel	Deret Fourier	
1	X_1	X_2, X_3	51.017
2	X_2	X_1, X_3	50.714
3	X_3	X_1, X_2	46.788

Based on Table 1, if the kernel component consists of 2 variables and Fourier series component consists of 1 variable. The minimum GCV is 44.166 which kernel component are X_1 dan X_3 while fourier series component is X_2. Based on Table 2, if kernel component consists of 1 variable and the fourier series component consists of 2 variables. The minimum GCV is 46.788 which kernel component is X_3 while fourier series component are X_1 and X_2. Based on Table 1 and Table 2, among all possibilities obtained the smallest GCV value is 44.166. Thus, the percentage of poor population is approached with nonparametric regression of kernel mix and Fourier series, which predictor variables MYS (X_1) and LR (X_3) are approached kernel function while predictor variable UR (X_2) is approached with Fourier series function. Therefore, predictor variable MYS is symbolized by v_1 and LR is symbolized by v_2, while variable predictor UR is symbolized by t_1.
Determination of bandwidth values, smoothing parameters and oscillation parameters use GCV values too. Comparison of GCV values for each model in Table 3.

Table 3. Comparison of GCV value for Model Selection

No.	ϕ_1	ϕ_2	λ	M	GCV
1	0.047	4.146	10^{-3}	1	42.888
2	0.047	4.146	10^{-3}	2	43.439
3	0.047	4.146	1	3	43.617
4	0.047	4.146	10^{-3}	4	43.063
5	0.047	4.146	10^{-3}	5	43.419

Based on Table 3, the minimum GCV is 42.888, therefore $\phi_1 = 0.047$; $\phi_2 = 4.146$; $M = 1$ and $\lambda = 10^{-3}$. The estimates of the Fourier series component are presented in Table 4.

Table 4. Fourier Series Component Parameter estimation

Parameter	$b(\lambda)$	$a_0(\lambda)$	$a_1(\lambda)$	λ
Estimation	-0.148	2.140	1.799	0.001

So the mixed nonparametric regression model of kernel and Fourier series is as follows:

$$\hat{y}_i = \sum_{i=1}^{102} \frac{\frac{1}{0.047} K\left(\frac{v_1 - v_{1i}}{0.047}\right)}{\sum_{i=1}^{102} \frac{1}{0.047} K\left(\frac{v_1 - v_{1i}}{0.047}\right)} y_i + \sum_{i=1}^{102} \frac{\frac{1}{4.146} K\left(\frac{v_2 - v_{2i}}{4.146}\right)}{\sum_{i=1}^{102} \frac{1}{4.146} K\left(\frac{v_j - v_{2i}}{4.146}\right)} y_i - 0.148 t_{1i} + 1.070 + 1.799 \cos t_{1i}$$

This modeling resulted in R^2 of 62.78 percent and MSE of 23.370. R^2 value is higher than using only the kernel estimator or only Fourier series or multiple linear regression. Table with comparisons of R^2 value are in Table 5.

Table 5. Comparison of R^2 values

Modeling	**R^2**
Multiple Linear Regression	37.76%
Kernel	49.36%
Fourier Series	42.48%
Mixed of Kernel and Fourier Series	62.78%

Table 5 show that the R^2 value of the model with mixed nonparametric component of kernel Fourier series is 62,78%. It means MYR, LR and UR variable is able to explain the response variable (percentage of poor people) equal to 62,78 percent.

From the model above, we obtain the scatter plot between y and $\hat{y}$ value as shown in Figure 7. Based on Figure 7, there are 15 regency/city that have considerable differences between y and $\hat{y}$ that are Yapen Waropen Regency, Waropen Regency, Supiori Regency, Teluk Wondama Regency, Teluk Bintuni Regency, Manokwari Regency, Sorong Regency, Belu Regency, East Flores Regency, Ngada Regency, Nagekeo Regency, Southwest Maluku Regency, Gorontalo City, Lebong Regency and Central Bengkulu Regency. Thus there is the possibility of existing factors other than MYR, LR and UR that affect the prediction of the percentage of poor people in 15 regency/city. This is can be the reason for the large difference between the prediction of poor people percentage to the actual value.

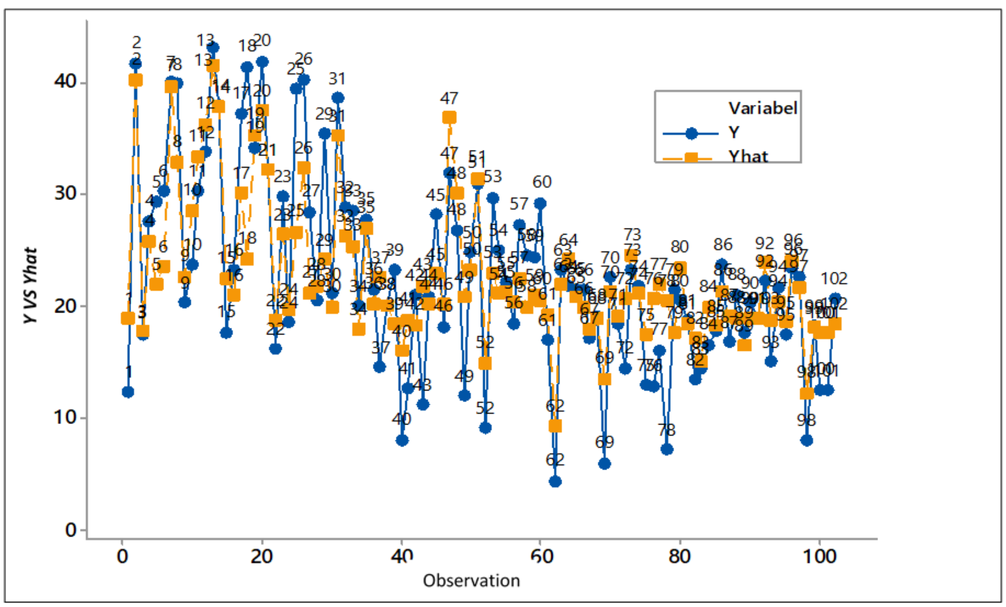


Figure 7. Scatter Plot y vs $\hat{y}$

4. CONCLUSION

If given paired data $\left(v_{1i},\ldots,v_{pi},t_{1i},\ldots,t_{qi},y_i\right)$, $i=1,2,\ldots,n$, following the mixed regression model i.e.

$$y_i = \mu\left(\mathbf{v}_i,\mathbf{t}_i\right)+\varepsilon_i$$

where $\mathbf{v}_i=(v_{1i},\ldots,v_{pi})'$ and $\mathbf{t}_i=(t_{1i},t_{2i}\ldots,t_{qi})$.

The regression curve $\mu\left(\mathbf{v}_i,\mathbf{t}_i\right)$ is assumed to be additive, so it can be written as:

$$\mu\left(\mathbf{v}_i,\mathbf{t}_i\right)=\sum_{j=1}^{p} g_j\left(v_{ji}\right)+\sum_{s=1}^{q} h_s\left(t_{si}\right).$$

The component

$$\sum_{j=1}^{p} g_j\left(v_{ji}\right)$$

is the kernel component, and the component :

$$\sum_{s=1}^{q} h_s\left(t_{si}\right)$$

is the Fourier series component. Then the mixed estimator $\boldsymbol{\mu}\left(\mathbf{v}_i,\mathbf{t}_i\right)$ is given by

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)=\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)+\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)=\mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)\mathbf{y}.$$

where

$$\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)=\mathbf{V}\left(\boldsymbol{\Phi}\right)\mathbf{y},\ \sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)=\mathbf{X}\hat{\mathbf{a}}\left(\boldsymbol{\lambda}\right),\ \ \hat{\mathbf{a}}\left(\lambda\right)=C\left(\boldsymbol{\Phi},\lambda,M\right)\mathbf{y},$$

$$C\left(\boldsymbol{\Phi},\lambda,M\right)=\left(\mathbf{X}'\mathbf{X}+n\mathbf{D}\left(\boldsymbol{\lambda}\right)\right)^{-1}\mathbf{X}'\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right),\ \mathbf{Z}\left(\boldsymbol{\Phi},\boldsymbol{\lambda},M\right)=\mathbf{V}\left(\boldsymbol{\Phi}\right)+X\left(\mathbf{X}'\mathbf{X}+n\mathbf{D}\left(\boldsymbol{\lambda}\right)\right)^{-1}\mathbf{X}'\left(\mathbf{I}-\mathbf{V}\left(\boldsymbol{\Phi}\right)\right)$$

The estimator of the kernel component is $\sum_{j=1}^{p}\hat{\mathbf{g}}_{j\phi_j}\left(v_j\right)$, the estimator of the Fourier series component is $\sum_{s=1}^{q}\hat{\mathbf{h}}_{\lambda_s,M}\left(t_s\right)$ and the estimator of the mixed kernel and the Fourier series is $\hat{\boldsymbol{\mu}}_{\boldsymbol{\Phi},\boldsymbol{\lambda},M}\left(\mathbf{v}_i,\mathbf{t}_i\right)$. It can be seen that each of these estimators are biased estimators, but keep are classified as linear estimators in y the observation.

Estimator of the mixed kernel and the Fourier series in the nonparametric regression are strongly influenced by the bandwidth parameter, the oscillation parameter and the smoothing parameter. The best of the mixed estimator of kernel and Fourier series are obtained from the optimal values of the bandwidth parameter, the oscillation parameter and the smoothing parameter. The optimal value of the bandwidth parameter, the oscillation parameter and the smoothing parameter can be obtained from the smallest GCV value.

Mixed Nonparametric regression model of kernel and Fourier series applied to percentage of poor people in Indonesiain 2013. In this research, the variable response is the percentage of poor people (y), while the predictor variables are Mean Years Schooling or MYS (v_1), Literacy Rate or LR (v_2) and Unemployment Rate or UR (t_1). This modeling uses $\phi_1 = 0.047$; $\phi_2 = 4.146$; $M = 1$ and $\lambda = 10^{-3}$. It produces R^2 of 62.78 percent and MSE of 23.370. Therefore, variables MYS, LR and UR can explain the percentage of poor peole as many as 62.78 percent. Mixed nonparametric model isbetter than kernel modeling, Fourier series model and multiple linear regression. From the modeling, there are 15 regency/city that have a considerable difference between y and $\hat{y}$. This is due to the possibility of existing other factors in addition to MYS, LR and UR that affect the prediction of the percentage of the poor people in 15 regency/city.

RECEIVED: JULY, 2018
REVISED: APRIL, 2019

REFERENCES

[1] AFIFAH, N., BUDIANTARA, I.N., and LATRA, I.N. (2016): Mixed Estimator of Kernel and Fourier Series in Nonparametric Regression. **Proceeding of Academics World International Conference**, 1-6.

[2] ANTONIADIS, A., BIGOT, J. and SAPATINAS, Y. (2001): Wavelet Estimators in Nonparametric Regression: A Comparative Simulation Study. **Journal of Statistical Software**, 6, 1-83.

[3] AMATO, U., and DE-CANDITIIS, D. (2001): Convergenee in Probability of Mallows and GCV Wavelet and Fourier Regularization Methods. **Statistics & Probability Letters**, 54, 325-329

[4] ASRINI, L. J., and BUDIANTARA, I. N. (2014): Fourier Series Semiparametric Regression Models (Case Study: The Production of Lawland Rice Irrigation in Central Java). **ARPN Journal of Engineering and Applied Sciences**, 9, 1501-1506.

[5] BILODEAU, M (1992): Fourier Smoother and Additive Models. **The Canadian Journal of Statistic**, 20, 257 – 269.

[6] BADAN PUSAT STATISTIK (2016) : BPS. http://bps.go.id/index.php/linkTabelStatis/981.

[7] BUDIANTARA, I N. (2000): Metode U, GML, CV dan GCV Dalam Regresi Nonparametrik Spline. **Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)**, 6, 41-45.

[8] BUDIANTARA, I. N., RATNASARI, V., RATNA, M., and ZAIN, I. (2015): The Combination of Spline and Kernel "Estimator for Nonparametric Regression and its Properties. **Journal of Applied Mathematical Sciences**, 9, 6083-6094.

[9] CHENG, M. Y., PAIGE, R. L., SUN, S. and YAN, K. (2010): Variance Reduction for Kernel Estimators in Clustered/Longitudinal Data Analysis. **Journal of Statistical Planning and Inference**, 140, 1389-1397.

[10] DARMAWI, H. and OTOK, B.W. (2014): Bootstrap Pada Regresi Linier dan Spline Truncated. **Jurnal Statistika FMIPA Universitas Islam Bandung**, 8, 47-54.

[11] EUBANK, R.L. (1999): **Nonparametric Regression and Spline Smoothing**, Volume 157, 2nd edition, Marcel Dekker, Inc**.**, New York.

[12] FERNANDES, A., A. R., BUDIANTARA, I. N., OTOK, B. W., and SUHARTONO (2014): Reproducing Kernel Hilbert Space for Penalized Regression Multi Predictors: Case in Longitudinal Data. **International Journal of Mathematical Analysis**, 8, 1951-1961.0

[13] FERNANDES, A. A. R., BUDIANTARA, I. N., OTOK, B. W., and SUHARTONO (2015): Spline Estimator for Bi-Response and Muti-Predictors Nonparametric Regression Model in Case of Longitudinal Data. **Journal of Mathematics and Statistics**, 11, 61-69.

[14] FILHO, C. M., and YAO, F. (2009): Nonparametric Regression Estimation With General Parametric Error Covariance. **Journal of Multivariate Analysis**, 100, 309-333.

[15] GALTCHOUK, L., and PERGAMENSHCHIKOV, S. (2009: Adaptive Asymptotically Efficient Estimation in Heteroscedastic Nonparametric Regression. **Journal of the Korean Statistical Society**, 38, 305-322.

[16] HE, H., and HUANG, L. S. (2009): Double Smoothing for Bias Reduction in Local Linier Regression. **Journal of Statistical Planning and Inference**, 139, 1056-1072.

[17] KAYRI, M., and ZIRHLIOGLU, G. (2009): Kernel Smoothing Function and Choosing Bandwidth for Nonparametric Regression Methods. **Ocean Journal of Applied Sciences**, 2, 49-54.

[18] MERDEKAWATI, I.P., and BUDIANTARA, I.N. (2013): Pemodelan Regresi Spline *Truncated* Multivariabel pada Faktor-Faktor yang mempengaruhi Kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah. **Jurnal Sains dan Seni POMITS**, 2, 19-24.

[19] OKUMURA, H., and NAITO, K. (2006): Non-Parametric Kernel Regression for Multinomial Data. **Journal of Multivariate Analysis**, 97, 2009-2022.

[20] PANE, R., BUDIANTARA, I. N., ZAIN, I., and OTOK, B. W. O. (2014): Parametric and Nonparametric Estimators in Fourier Series Semiparametric Regression and Their Characteristics. **Journal of Applied Mathematical Sciences**, 8, 5053-5064.

[21] QINGGUO, T. (2010): L1-Estimation in Semiparametric Model with Longitudinal Data. **Journal of Statistical Planning and Inference**, 140, 393-405.

[22] RATNASARI, V., BUDIANTARA, I. N., ZAIN, I. RATNA, M. and MARIATI, N.P.A.M. (2015): Comparison Truncated Spline and Fourier Series in Multivariable Nonparametric Regresssion Model (Application: Data of Poverty in Papua, Indonesia). **International Journal of Basic & Applied Sciences IJBASIJENS**, 15, 9-12.

[23] RATNASARI, V., BUDIANTARA, I N., RATNA, M., and ZAIN, I. (2016): Estimation of Nonparametric Regression Curve using Mixed Estimator of Multivariable Truncated Spline and Multivariable Kernel. **Global Journal of Pure and Applied Mathematics**, 12, 5047-5057

[24] SU, L., and ULLAH, A. (2008): Local Polynomial Estimation of Nonparametric Simultaneous Equations Models. **Journal of Econometrics**, 144, 193-218.

[25] TAYLOR, L. W. (2009): Using the Haar Wavelet Transform in the Semiparametric Specification of Time Series. **Economic Modeling**, 26, 392-403.

[26] WAHBA, G. (1990): **Spline Model for Observational Data**, SIAM Pensylvania.

[27] WIBOWO, W., HARYATMI, S., and BUDIANTARA, I. N. (2012): On Multiresponse Semiparametric Regression Model. **Journal of Mathematics and Statistics**, 8, 489-499.

[28] YAO, F. (2007): Asymptotic Distribution of Nonparametric Regression Estimators for Longitudinal or Functional Data. **Journal of Multivariate Analysis**, 98, 40-56.