International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 5, Issue 3, 2019 Special Edition: *Science, Applied Science, Teaching and Education*

Comparison of Smoothing and Truncated Spline Estimators in Estimating Blood Pressure Models

Fatmawati[1], I Nyoman Budiantara[2], Budi Lestari[3]*

[1]Department of Mathematics, Faculty of Science and Technology, Airlangga University, Jalan Mulyorejo Kampus C-UNAIR, Surabaya 60115, Indonesia.

[2]Department of Statistics, Sepuluh Nopember Institute of Technology, Indonesia.

[3]Department of Mathematics, The University of Jember, Indonesia.

[3]Doctoral Study Program, Faculty of Science and Technology, Airlangga University, Indonesia.

Email: budi.lestari-2016@fst.unair.ac.id

Abstract

The functions, namely regression functions, which describe the relationship of more than one response variable observed at several values of the predictor variables in which there are correlations between responses can be estimated by using both smoothing spline and truncated spline estimators in multi-response non-parametric regression model that is as development of a uni-response non-parametric regression model. In this paper, we discuss estimating regression function of the multi-response non-parametric regression model by using both smoothing spline and truncated spline estimators with application to the association between blood pressures affected by body mass index. Results show that by comparing their mean squared error values, smoothing spline estimators give a better estimate of results than truncated spline estimators. It means that for a prediction need, smoothing spline estimators are better than truncated spline estimators.

Keywords: Blood Pressure, Body Mass Index, Multi-response Non-parametric Regression, Smoothing Spline, Truncated Spline.

1. Introduction

According to IOTF-WHO (2000), the risk of negative health consequences is correlated with increasing body mass index (BMI), and a BMI more than or equal to 23 kg/m^2 was categorised as overweight or obese. Since levels of overweight can be measured by BMI (body mass index), then an increase in BMI can also cause an increase in blood pressure (systolic and diastolic). An increase of BMI in someone who was less than 60 years old can cause an increase of systolic and diastolic blood pressure as shown by Brown et al. (2000). Next, Droyvold et al. (2005) and Lestari et al. (2019c) pointed out that increasing and decreasing systolic and diastolic blood pressure were significantly caused by increasing and decreasing of BMI for all sex and all ages. Further, Tesfaye et al. (2007) stated that BMI significantly influenced systolic and diastolic blood pressures of Ethiopian, Vietnamese, and Indonesian people. Also, Kumar et al. (2008) pointed out that BMI affects systolic and diastolic blood pressures of females and males. In addition, Nanaware et al. (2011) have shown that there was a positive correlation between BMI and both systolic and diastolic blood pressure of children aged between 8-16 years old. Then, Roka et al. (2015) pointed out that a high BMI, being overweight or obese can cause an increase of blood pressure (systolic and diastolic).

Statistical analysis often involves building mathematical models which examines the relationship between response and predictor variables. Spline is a general class of powerful and flexible modeling techniques. Research on spline models has attracted a great deal of attention in recent years, and the methodology has been widely used in many areas. Spline estimator with its powerful and flexible properties is one of the most popular estimators used for estimating regression function of the non-parametric regression model. There are many researchers who have considered spline estimator for estimating regression function of the non-parametric regression model. Researchers in Kimeldorf & Wahba (1971), Craven & Wahba (1979), and Wahba (1990) used original spline estimators to estimate regression function of smooth data. M-type splines to overcome outliers in non-parametric regression were proposed by Cox (1983), and Cox & O'Sullivan (1996). Confidence intervals for original spline models by using the Bayesian approach has been constructed by Wahba (1983). Also, Wahba (1985) compared between generalised cross validation (GCV) and generalised maximum likelihood (GML) for choosing the smoothing parameter in the generalised spline smoothing problem. Relaxed spline and quantile spline have been introduced by Oehlert (1992) & Koenker et al. (1994). Next, Wang (1998) discussed smoothing spline models with correlated random errors. Some techniques for spline statistical model building by using reproducing kernel Hilbert spaces (RKHS) have been introduced by Wahba (2000). A method that combines smoothing spline estimates of different smoothness to form a final improved estimate was proposed by Lee (2004). Further, Cardot et al. (2007) gave the asymptotic property of smoothing splines estimators in functional linear regression with errors-in-variables. Smoothing spline estimation of variance functions has been studied by Liu *et al.* (2007). Also, Aydin (2007) showed goodness of spline estimator rather than kernel estimator in estimating non-parametric regression model for gross national product data. Next, Aydin et al. (2013) have studied the determination of an optimal smoothing parameter for non-parametric regression using smoothing spline. All these researchers studied spline estimators in the case of single response non-parametric regression models only.

In the real cases, we are frequently faced the problem in which two or more dependent variables are observed at several values of the independent variables, and there are correlations between the responses. Multi-response non-parametric regression models provide powerful tools to model the functions which represent the association of these variables. There are many researchers who have considered non-parametric models for multi-response data. Spline smoothing for estimating non-parametric functions from bivariate data with the same correlation of errors has been studied by Wang et al. (2000). Next, Fernandez and Opsomer (2005) proposed methods of estimating the non-parametric regression model with spatially correlated errors. Also, Lestari et al. (2009) and Lestari et al. (2010) have studied spline estimators in multi-response non-parametric regression models with equal correlation of errors and unequal correlation of errors, respectively. Then, Chamidah et al. (2012) used multi-response non-parametric regression model approach to design a child's growth chart. A mathematical statistics method for estimating the regression curve of the multi-response non-parametric regression model in case of heteroscedasticity of variance was proposed by Lestari et al. (2012). In addition, Chamidah & Lestari (2016) discussed estimating the regression curve of the homoscedastic multi-response non-parametric regression in which the number of observations were unbalanced. Smoothing spline estimators for estimating the multi-response non-parametric regression model by using RKHS has been proposed by Lestari et al. (2017b) and Lestari et al. (2018). Further, Lestari et al. (2017a), Lestari et al. (2018b) and Lestari et al. (2019a) discussed construction of covariance matrix in case of homoscedasticity of variances of errors, and estimating of both covariance matrix and optimal smoothing parameter. But, these researchers have not discussed estimating of the smoothing parameter in the multi-response non-parametric regression model when the variances of errors are not the same for cross-section data. In addition, Lestari et al. (2019b) have discussed estimating of smoothing parameter in multi-response non-parametric regression model when the variances of errors are not the same for cross-section data but these researchers have not discussed application of the multi-response non-parametric regression model on the real case data. Therefore, in this paper, we discuss methods to estimate regression function of the multi-response non-parametric regression model that apply to real case data, i.e. data of blood pressures and BMI. Thus, the goals of this research are estimating a model of blood pressures affected by BMI by using both smoothing spline and truncated spline estimators, and comparing between smoothing spline and truncated spline estimators in estimating the blood pressures based on their mean squared errors values.

2. Methods for Estimating Blood Pressures Models

Firstly, for estimating blood pressures models by using smoothing spline, we consider a multi-response non-parametric regression model given by Lestari et al. (2019b) as follows:

$$y_{ki} = f_k(t_{ki}) + \varepsilon_{ki}\,;\; k = 1,2,\ldots,p\,;\; i = 1,2,\ldots,n_k\,,\text{ where } Var(\varepsilon_{ki}) = \sigma_{ki}^2 \tag{1}$$

Next, by putting $k = 1,2$ and $i = 1,2,\ldots,n$, we apply model (1) to the data of blood pressures affected by BMI such that we have a blood pressures model as follows:

$$y_{ki} = f_k(t_{ki}) + \varepsilon_{ki}\,;\; k = 1,2\,;\; i = 1,2,\ldots,n \tag{2}$$

where $Var(\varepsilon_{ki}) = \sigma_{ki}^2$, y_{1i} and y_{2i} are response variables that represent the first response (i.e., systolic blood pressure), and the second response (i.e., diastolic blood pressure), respectively, $f_k(t_{ki})$, $k = 1,2$ are unknown regression functions which represent the function of predictor variables (i.e., BMI). Also, we construct a covariance matrix of errors, namely $Cov(\underset{\sim}{\varepsilon}) = \mathbf{W}^{-1}(\underset{\sim}{\sigma}^2)$. Next, by using reproducing kernel Hilbert space (RKHS), we take solutions to penaliszed weighted least square (PWLS) optimiszation:

$$\underset{f_1 \in W_2^m[a_1,b_1],\, f_2 \in W_2^m[a_2,b_2]}{Min} \left\{ \left(\underset{\sim}{y} - \underset{\sim}{f}\right)' W(\underset{\sim}{\sigma}^2)\left(\underset{\sim}{y} - \underset{\sim}{f}\right) + \sum_{k=1}^{2} \lambda_k \int_{a_k}^{b_k} \left[f_k^{(m)}(t_k) \right]^2 dt_k \right\} \tag{3}$$

for determining the estimation of regression function in model (2) that depends on selecting the optimal smoothing parameter (λ_{opt}). It can be obtained by taking solution of generaliszed cross validation $G(\lambda_k, \underset{\sim}{\sigma}^2)$ optimiszation:

$$\lambda_{opt} = \underset{\lambda_k}{Min}\left\{ G\left(\lambda_k, \underset{\sim}{\sigma}^2\right) \right\} = \underset{\lambda_k}{Min}\left\{ \frac{n^{-1} \left\| [I - H(\lambda_k)]\underset{\sim}{y} \right\|^2}{\left[n^{-1} trace\left(I - H(\lambda_k)\right) \right]^2} \right\},\; k = 1,2 \tag{4}$$

Secondly, for estimating blood pressures models by using truncated spline, we test correlation systolic and diastolic blood pressure. Next, we determine optimal smoothing parameters (λ_{opt}), optimal knot, and optimal order of truncated spline. Then, we determine estimated blood pressures model by applying truncated spline as well as given by Wahba (1990) as follows:

$$S(t)=\sum_{i=0}^{r-1}\alpha_i t^i+\sum_{j=1}^{n}\beta_j\left(t-k_j\right)_+^{r-1} \tag{5}$$

where $(r-1)$ is the order of spline, $k_1,k_2,...,k_n (a<k_1<k_2<...<k_n<b)$ are knots of spline, $t\in[a,b]$, α_i and β_j are real valued constants, and

$$\left(t-k_j\right)_+^{r-1}=\begin{cases}\left(t-k_j\right)^{r-1}, & for \;\; t\geq k_j\\ 0, & for \;\; t<k_j\end{cases} \tag{6}$$

3. Results and Discussion

3.1. Estimating Blood Pressures Using Smoothing Spline Estimator

We consider model (2) and suppose that $\underset{\sim}{y}=(\underset{\sim}{y}_1,\underset{\sim}{y}_2)'$, $\underset{\sim}{f}=(\underset{\sim}{f}_1,\underset{\sim}{f}_2)'$, $\underset{\sim}{\varepsilon}=(\underset{\sim}{\varepsilon}_1,\underset{\sim}{\varepsilon}_2)'$, and $\underset{\sim}{t}=(\underset{\sim}{t}_1,\underset{\sim}{t}_2)'$ where $\underset{\sim}{y}_k=(y_{k1},y_{k2},...,y_{kn})'$, $\underset{\sim}{f}_k=(f_k(t_{k1}),f_k(t_{k2}),...,f_k(t_{kn}))'$, $\underset{\sim}{\varepsilon}_k=(\varepsilon_{k1},\varepsilon_{k2},...,\varepsilon_{kn})'$, $\underset{\sim}{t}_k=(t_{k1},t_{k2},...,t_{kn})'$, $k=1,2$; $i=1,2,...,n$. Therefore, we can write equation (2) in matrix notation as follows:

$$\underset{\sim}{y}=\underset{\sim}{f}+\underset{\sim}{\varepsilon} \tag{7}$$

where $E(\underset{\sim}{\varepsilon})=\underset{\sim}{0}$, and $Cov(\underset{\sim}{\varepsilon})=\mathbf{W}^{-1}(\underset{\sim}{\sigma}^2)=\text{diag}(\mathbf{W}_1(\underset{\sim}{\sigma}_1^2),\mathbf{W}_2(\underset{\sim}{\sigma}_2^2))$. Estimating functions $\underset{\sim}{f}$ in (7) by using smoothing spline estimators appears as a solution to the penaliszed weighted least-square (PWLS) minimiszation problem, i.e., determine $\hat{\underset{\sim}{f}}$ that can make the following PWLS minimum:

$$\underset{f_1,f_2\in W_2^m}{Min}\{(\sum_{k=1}^{2}n_k)^{-1}(\underset{\sim}{y}_1-\underset{\sim}{f}_1)'\mathbf{W}_1(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}_1-\underset{\sim}{f}_1)+(\underset{\sim}{y}_2-\underset{\sim}{f}_2)'\mathbf{W}_2(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}_2-\underset{\sim}{f}_2)+$$

$$\lambda_1\int_{a_1}^{b_1}(f_1^{(2)}(t))^2dt+\lambda_2\int_{a_2}^{b_2}(f_2^{(2)}(t))^2dt\} \tag{8}$$

for pre-specified value $\underset{\sim}{\lambda}=(\lambda_1,\lambda_2)'$. Note that, in equation (8), the first term represents the sum squares of errors and it penaliszes the lack of fit. While, the second term which is weighted by $\underset{\sim}{\lambda}$ represents the roughness penalty and it imposes a penalty on roughness. It means that the curvature of $\underset{\sim}{f}$ is penaliszed by it. In equation (11), λ_k $(k=1,2)$ is called as the smoothing parameter. The solution will be variedy from interpolation to a linear model, if λ_k varies from 0 to $+\infty$. So that, if $\lambda_k \to +\infty$, the roughness penalty will dominate in (8), and the smoothing spline estimate will be forced to be a constant. If $\lambda_k \to 0$, the roughness penalty will disappear in (8), and the spline estimate will interpolate the data. Thus, the trade-off between the goodness of fit given by $(\sum_{k=1}^{2} n_k)^{-1}(\underset{\sim}{y}_1-\underset{\sim}{f}_1)'\mathbf{W}_1(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}_1-\underset{\sim}{f}_1)+(\underset{\sim}{y}_2-\underset{\sim}{f}_2)'\mathbf{W}_2(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}_2-\underset{\sim}{f}_2)$ and smoothness of the estimate given by $\lambda_1\int_{a_1}^{b_1}(f_1^{(2)}(t))^2\,dt+\lambda_2\int_{a_2}^{b_2}(f_2^{(2)}(t))^2\,dt$ is controlled by the smoothing parameter λ_k $(k=1,2)$. The solution for minimiszation of the problem in (8) is a smoothing spline estimator where its function basis is a "natural cubic spline" with $t_1,t_2,\ldots,t_{n_k}$ $(k=1,2)$ as its knots. Based on this concept, a particular structured spline interpolation that depends on selection of the smoothing parameter λ_k $(k=1,2)$ value becomes an appropriate approach of the functions f_k $(k=1,2)$ in model (2). Let $\underset{\sim}{f}=(\underset{\sim}{f}_1,\underset{\sim}{f}_2)'$ where $\underset{\sim}{f}_k=(f_k(t_{k1}),f_k(t_{k2}),\ldots,f_k(t_{kn}))'$, $k=1,2$, be the vector of values of function f_k $(k=1,2)$ at the knot points $t_1,t_2,\ldots,t_{n_k}$ $(k=1,2)$. If we express the model of paired data set into a general smoothing spline regression model, we will get the following expression:

$$y_{ki}=L_{t_k}f_k+\varepsilon_{ki}\,;\; i=1,2,\ldots,n_k\,;\; k=1,2 \tag{9}$$

where $f_k \in \mathcal{H}_k$ ($\mathcal{H}_k$ represents Hilbert space) is an unknown smooth function, and $L_{t_k} \in \mathcal{H}_k$ is a bounded linear functional.

Next, suppose that $\mathcal{H}_k$ can be decomposed into two subspaces $\mathbf{U}_k$ and $\mathbf{W}_k$ as follows:

$$\mathcal{H}_k = \mathcal{U}_k \oplus \mathcal{W}_k$$

where $\mathcal{U}_k$ is orthogonal to $\mathcal{W}_k$, $k = 1,2$. Suppose that $\{u_{k1}, u_{k2}, ..., u_{km_k}\}$ and $\{\omega_{k1}, \omega_{k2}, ..., \omega_{kn_k}\}$ are bases of spaces $\mathcal{U}_k$ and $\mathcal{W}_k$, respectively. Then, we can express every function $f_k \in \mathcal{H}_k$ ($k = 1,2$) into the following expression:

$$f_k = g_k + h_k$$

where $g_k \in \mathcal{U}_k$ and $h_k \in \mathcal{W}_k$. Since $\{u_{k1}, u_{k2}, ..., u_{km_k}\}$ is basis of space $\mathcal{U}_k$ and $\{\omega_{k1}, \omega_{k2}, ..., \omega_{kn_k}\}$ is basis of space $\mathcal{W}_k$, then for every $f_k \in \mathcal{H}_k$ ($k = 1,2$) follows:

$$f_k = \sum_{j=1}^{m_k} d_{kj} u_{kj} + \sum_{i=1}^{n_k} c_{ki} \omega_{ki} = \underset{\sim}{u}'_k \underset{\sim}{d}_k + \underset{\sim}{\omega}'_k \underset{\sim}{c}_k;\ k = 1,2;\ d_{kj} \in \mathbb{R};\ c_{ki} \in \mathbb{R} \tag{10}$$

where $\underset{\sim}{u}_k = (u_{k1}, u_{k2}, ..., u_{km_k})'$, $\underset{\sim}{d}_k = (d_{k1}, d_{k2}, ..., d_{km_k})'$, $\underset{\sim}{\omega}_k = (\omega_{k1}, \omega_{k2}, ..., \omega_{kn_k})'$, and $\underset{\sim}{c}_k = (c_{k1}, c_{k2}, ..., c_{kn_k})'$. Furthermore, since $L_{t_{ki}}$ is a function which is bounded and linear in $\mathcal{H}_k$, and $f_k \in \mathcal{H}_k$, $k = 1,2$ then we have:

$$L_{t_{ki}} f_k = L_{t_{ki}}(g_k + h_k) = g_k(t_{ki}) + h_k(t_{ki}) = f_k(t_{ki}). \tag{11}$$

Based on (11), and by applying the Riesz representation theorem Wang (2011), and because of $L_{t_{ki}} \in \mathcal{H}_k$ is bounded linear functional, then according to [36] there is a representer $\xi_{ki} \in \mathcal{H}_k$ of $L_{t_{ki}}$ which follows:

$$L_{t_{ki}} f_k = \langle \xi_{ki}, f_k \rangle = f_k(t_{ki}),\ f_k \in \mathcal{H}_k \tag{12}$$

where $\langle \cdot,\cdot \rangle$ denotes an inner product. Based on (6) and by applying the properties of the inner product, we get:

$$f_k(t_{ki}) = \langle \xi_{ki}, \underset{\sim}{u}'_k \underset{\sim}{d}_k + \underset{\sim}{\omega}'_k \underset{\sim}{c}_k \rangle = \langle \xi_{ki}, \underset{\sim}{u}'_k \underset{\sim}{d}_k \rangle + \langle \xi_{ki}, \underset{\sim}{\omega}'_k \underset{\sim}{c}_k \rangle. \tag{13}$$

Next, by applying equation (13), for $k = 1$ we have:

$$f_1(t_{1i}) = \langle \xi_{1i}, \underset{\sim}{u}_1' \underset{\sim}{d}_1 \rangle + \langle \xi_{1i}, \underset{\sim}{\omega}_1' \underset{\sim}{c}_1 \rangle,\ i = 1, 2, \ldots, n_1;$$

and for $i = 1, 2, 3, \ldots, n_1$ we have:

$$\underset{\sim}{f}_1(t_1) = (f_1(t_{11}), f_1(t_{12}), \ldots, f_1(t_{1n_1}))' = K_1 \underset{\sim}{d}_1 + \Sigma_1 \underset{\sim}{c}_1, \tag{14}$$

where:

$$K_1 = \begin{bmatrix} \langle \xi_{11}, u_{11} \rangle & \langle \xi_{11}, u_{12} \rangle & \cdots & \langle \xi_{11}, u_{1m_1} \rangle \\ \langle \xi_{12}, u_{11} \rangle & \langle \xi_{12}, u_{12} \rangle & \cdots & \langle \xi_{12}, u_{1m_1} \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \xi_{1n_1}, u_{11} \rangle & \langle \xi_{1n_1}, u_{12} \rangle & \cdots & \langle \xi_{1n_1}, u_{1m_1} \rangle \end{bmatrix},\ \underset{\sim}{d}_1 = (d_{11}, d_{12}, \ldots, d_{1m_1})',$$

$$\Sigma_1 = \begin{bmatrix} \langle \xi_{11}, \omega_{11} \rangle & \langle \xi_{11}, \omega_{12} \rangle & \cdots & \langle \xi_{11}, \omega_{1n_1} \rangle \\ \langle \xi_{12}, \omega_{11} \rangle & \langle \xi_{12}, \omega_{12} \rangle & \cdots & \langle \xi_{12}, \omega_{1n_1} \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \xi_{1n_1}, \omega_{11} \rangle & \langle \xi_{1n_1}, \omega_{12} \rangle & \cdots & \langle \xi_{1n_1}, \omega_{1n_1} \rangle \end{bmatrix},\ \text{and}\ \underset{\sim}{c}_1 = (c_{11}, c_{12}, \ldots, c_{1n_1})'.$$

Similarly, we obtain: $\underset{\sim}{f}_2(t_2) = \mathbf{K}_2 \underset{\sim}{d}_2 + \mathbf{\Sigma}_2 \underset{\sim}{c}_2$. Therefore, the regression function $\underset{\sim}{f}(t)$ can be expressed as:

$$\begin{aligned} \underset{\sim}{f}(t) &= (\underset{\sim}{f}_1(t_1), \underset{\sim}{f}_2(t))' = (\mathbf{K}_1 \underset{\sim}{d}_1, \mathbf{K}_2 \underset{\sim}{d}_2)' + (\mathbf{\Sigma}_1 \underset{\sim}{c}_1, \mathbf{\Sigma}_2 \underset{\sim}{c}_2)' \\ &= \mathrm{diag}(\mathbf{K}_1, \mathbf{K}_2)(\underset{\sim}{d}_1, \underset{\sim}{d}_2)' + \mathrm{diag}(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2)(\underset{\sim}{c}_1, \underset{\sim}{c}_2)' = \mathbf{K}\underset{\sim}{d} + \mathbf{\Sigma}\underset{\sim}{c}. \end{aligned} \tag{15}$$

In equation (15), $\mathbf{K}$ is a $(\mathrm{N} \times \mathrm{M})$-matrix and $\underset{\sim}{d}$ is a vector of parameters with dimension $(\mathbf{M} \times 1)$ (where $\mathrm{N} = \sum_{k=1}^{2} n_k = 2n$, $\mathbf{M} = \sum_{k=1}^{2} m_k = 2m$) that are expressed as:

$\mathbf{K} = diag(\mathbf{K}_1, \mathrm{K}_2)$, and $\underset{\sim}{d} = (\underset{\sim}{d}_1', \underset{\sim}{d}_2')'$, respectively.

Also, $\mathbf{\Sigma}$ is a $(\mathrm{N} \times \mathrm{N})$-matrix, and $\underset{\sim}{c}$ is a $(\mathrm{N} \times 1)$-vector of parameters which are expressed as:

$\Sigma = diag(\Sigma_1, \Sigma_2)$, and $\underset{\sim}{c} = (\underset{\sim}{c}_1', \underset{\sim}{c}_2')'$, respectively.

Therefore, we can write a model in (9) as follows:

$$\underset{\sim}{y} = \mathbf{K}\underset{\sim}{d} + \mathbf{\Sigma}\underset{\sim}{c} + \underset{\sim}{\varepsilon}.$$

We use the RKHS method to obtain the estimation of $\underset{\sim}{f}$, by solving the following optimiszation:

$$\underset{\substack{f_k \in \mathcal{H}_k \\ k=1,2}}{Min}\left\{\left\|\mathbf{W}^{\frac{1}{2}}(\underset{\sim}{\sigma}^2)\underset{\sim}{\varepsilon}\right\|^2\right\} = \underset{\substack{f_k \in \mathcal{H}_k \\ k=1,2}}{Min}\left\{\left\|\mathbf{W}^{\frac{1}{2}}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\underset{\sim}{f})\right\|^2\right\}, \tag{16}$$

With the constraint:

$$\int_{a_k}^{b_k}[f_k^{(m)}(t_k)]^2\,dt_k < \gamma_k \;,\; \gamma_k \geq 0 \;,\quad k=1,2\,. \tag{17}$$

To solve the optimisation (16) with constraint (17) is equivalent to solve the optimisation PWLS:

$$\underset{\substack{f_k \in W_2^m[a_k,b_k] \\ k=1,2}}{Min}\left\{\mathrm{N}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})'\mathbf{W}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\underset{\sim}{f})+\sum_{k=1}^{2}\lambda_k\int_{a_k}^{b_k}[f_k^{(m)}(t_k)]^2\,dt_k\right\}, \tag{18}$$

where λ_k, $k=1,2$ are smoothing parameters that control trade-off between goodness of fit represented by $\mathrm{N}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})'\mathbf{W}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\underset{\sim}{f})$ and the roughness penalty measured by $\lambda_1\int_{a_1}^{b_1}[f_1^{(m)}(t_1)]^2\,dt_1+\lambda_2\int_{a_2}^{b_2}[f_2^{(m)}(t_2)]^2\,dt_2$. To get the solution to (18), we first decompose the roughness penalty as follows:

$$\int_{a_1}^{b_1}[f_1^{(m)}(t_1)]^2\,dt_1 = \|Pf_1\|^2 = \langle Pf_1, Pf_1\rangle = \left\langle \underset{\sim}{\omega}_1'\underset{\sim}{c}_1, \underset{\sim}{\omega}_1'\underset{\sim}{c}_1\right\rangle = \underset{\sim}{c}_1'(\underset{\sim}{\omega}_1\underset{\sim}{\omega}_1')\underset{\sim}{c}_1 = \underset{\sim}{c}_1'\Sigma_1\underset{\sim}{c}_1$$

It implies:

$$\lambda_1\int_{a_1}^{b_1}[f_1^{(m)}(t_1)]^2\,dt_1 = \lambda_1\underset{\sim}{c}_1'\Sigma_1\underset{\sim}{c}_1, \text{ and } \lambda_2\int_{a_2}^{b_2}[f_2^{(m)}(t_2)]^2\,dt_2 = \lambda_2\underset{\sim}{c}_2'\Sigma_2\underset{\sim}{c}_2 \tag{19}$$

Based on (19), we have the penalty:

$$\sum_{k=1}^{2}\lambda_k\int_{a_k}^{b_k}[f_k^{(m)}(t_k)]^2\,dt_k\} = \underset{\sim}{c}'\lambda\Sigma\underset{\sim}{c} \tag{20}$$

where $\Lambda = diag(\lambda_1\mathbf{I}_{n_1}, \lambda_2\mathbf{I}_{n_2})$. We can express the goodness of fit in (18) as follows:

$$\mathrm{N}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})'\mathbf{W}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\underset{\sim}{f}) = \mathrm{N}^{-1}(\underset{\sim}{y}-\mathbf{K}\underset{\sim}{d}-\mathbf{\Sigma}\underset{\sim}{c})'\mathbf{W}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\mathbf{K}\underset{\sim}{d}-\mathbf{\Sigma}\underset{\sim}{c})\;.$$

If we combine the goodness of fit and the roughness penalty, we will have optimisation PWLS:

$$\underset{\substack{\underset{\sim}{c}\in R^{2n}\\ \underset{\sim}{d}\in R^{2m}}}{Min}\left\{(\underset{\sim}{y}-\mathbf{K}\underset{\sim}{d}-\mathbf{\Sigma}\underset{\sim}{c})'\mathbf{W}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\mathbf{K}\underset{\sim}{d}-\mathbf{\Sigma}\underset{\sim}{c})+\underset{\sim}{c}'\mathrm{N}\Lambda\mathbf{\Sigma}\underset{\sim}{c}\right\}=\underset{\substack{\underset{\sim}{c}\in R^{2n}\\ \underset{\sim}{d}\in R^{2m}}}{Min}\left\{Q(\underset{\sim}{c},\underset{\sim}{d})\right\} \tag{21}$$

To get the solution to (21), firstly we must take the partially differential of $Q(\underset{\sim}{c},\underset{\sim}{d})$ and then their results are to be equal to zeros as follows:

$$\partial Q(\underset{\sim}{c},\underset{\sim}{d})/\partial\underset{\sim}{c}=\underset{\sim}{0}\;\Leftrightarrow\;\hat{\underset{\sim}{c}}=\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\mathbf{K}\underset{\sim}{d}). \tag{22}$$

$$\partial Q(\underset{\sim}{c},\underset{\sim}{d})/\partial\underset{\sim}{d}=\underset{\sim}{0}\;\Leftrightarrow\;\hat{\underset{\sim}{d}}=[\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)\mathbf{K}]^{-1}\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)\underset{\sim}{y}. \tag{23}$$

Next, if we substitute (23) into (22), we obtain:

$$\hat{\underset{\sim}{c}}=\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)[\mathbf{I}-\mathbf{K}(\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)\mathbf{K})^{-1}\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)]\underset{\sim}{y}. \tag{24}$$

Finally, based on (15), (23) and (24), we get the estimated regression function based on the smoothing spline estimator which can be expressed as follows:

$$\hat{\underset{\sim}{f}}_{\lambda}(\underset{\sim}{t})=\begin{pmatrix}\hat{\underset{\sim}{f}}_{1,\lambda_1}(\underset{\sim}{t}_1)\\ \hat{\underset{\sim}{f}}_{2,\lambda_2}(\underset{\sim}{t}_2)\end{pmatrix}=\mathbf{K}\hat{\underset{\sim}{d}}+\mathbf{\Sigma}\hat{\underset{\sim}{c}}=\mathbf{H}(\underset{\sim}{\lambda})\underset{\sim}{y} \tag{25}$$

where

$$\mathbf{H}(\underset{\sim}{\lambda})=\mathbf{K}[\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)\mathbf{K}]^{-1}\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)+\mathbf{\Sigma}\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)\times[\mathbf{I}-\mathbf{K}(\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)\mathbf{K})^{-1}\mathbf{K}'\mathbf{M}^{-1}\mathbf{W}(\underset{\sim}{\sigma}^2)],$$

and $\hat{\underset{\sim}{f}}_{\lambda}(\underset{\sim}{t})$ is smoothing spline with a natural cubic spline as a basis function with knots at $t_1,t_2,\ldots,t_{n_k}$ $(k=1,2)$, for a fixed smoothing parameter $\underset{\sim}{\lambda}>\underset{\sim}{0}$. $\mathbf{H}(\underset{\sim}{\lambda})$ is a positive-definite (symmetrical) smoother matrix that depends on smoothing parameter $\underset{\sim}{\lambda}$ and the knot points $t_1,t_2,\ldots,t_{n_k}$ $(k=1,2)$. Yet, it does not depend on $\underset{\sim}{y}$. Further discussion about this estimator can be obtained on Watson (1964), Hardle (1991), Oehlert (1992), Eubank (1999), and Schimek (2000).

Related to application of multi-response non-parametric regression on the real data, in this research we use secondary data obtained from the District General Hospital of Trenggalek city, East Java Province, Indonesia. Data consists of systolic and diastolic blood pressures, and BMI of 99 patients who took medical care in that hospital in 2018. Data is provided in Table 1.

Table 1. Systolic and Diastolic Blood Pressures, and Body Mass Index (BMI) of 99 Patients

Patient	Systolic (mmHg) (y_1)	Diastolic (mmHg) (y_2)	BMI (kg/m²) (t)	Patient	Systolic (mmHg) (y_1)	Diastolic (mmHg) (y_2)	BMI (kg/m²) (t)	Patient	Systolic (mmHg) (y_1)	Diastolic (mmHg) (y_2)	BMI (kg/m²) (t)
1	110	70	18.4	34	126	96	23.2	67	130	90	25.7
2	120	78	18.8	35	120	80	23.2	68	115	80	25.8
3	130	90	18.9	36	110	75	23.3	69	150	104	26.1
4	115	75	19.0	37	160	110	23.4	70	122	94	26.1
5	120	80	19.3	38	120	88	23.4	71	125	88	26.2
6	110	70	19.4	39	114	76	23.5	72	140	95	26.3
7	148	90	19.5	40	120	70	23.6	73	190	112	26.3
8	140	90	19.7	41	120	74	23.6	74	130	80	26.3
9	120	80	20.0	42	130	90	23.6	75	130	88	26.4
10	135	90	20.2	43	120	90	23.7	76	138	94	26.5
11	108	70	20.4	44	132	90	23.8	77	120	90	26.6
12	122	78	20.7	45	138	82	23.8	78	115	80	26.7
13	115	70	20.7	46	140	90	23.9	79	130	90	26.8
14	114	80	20.9	47	130	80	23.9	80	114	80	27.1
15	140	90	21.0	48	110	75	24.1	81	120	86	27.1
16	110	75	21.1	49	115	65	24.2	82	124	80	27.3
17	118	88	21.1	50	122	82	24.2	83	130	90	27.3
18	110	70	21.3	51	110	80	24.3	84	130	80	27.4
19	120	80	21.3	52	120	80	24.3	85	155	104	27.4
20	130	90	21.4	53	130	80	24.4	86	125	90	27.5
21	110	60	22.0	54	125	90	24.5	87	120	85	27.5
22	125	80	22.0	55	130	94	24.5	88	112	90	27.6
23	120	80	22.2	56	130	88	24.6	89	136	82	27.6
24	138	90	22.2	57	112	80	24.7	90	190	110	27.7
25	110	80	22.3	58	150	100	24.7	91	120	80	27.7
26	120	82	22.5	59	156	90	24.8	92	125	90	27.9

27	100	70	22.6	60	112	78	24.9	93	160	100	27.9
28	140	90	22.6	61	130	75	25.0	94	141	100	28.5
29	120	70	22.6	62	120	85	25.1	95	130	90	28.6
30	154	80	22.8	63	120	82	25.2	96	168	105	29.2
31	108	68	23.0	64	128	92	25.3	97	160	90	29.6
32	134	90	23.0	65	120	80	25.4	98	170	100	30.3
33	110	70	23.1	66	160	90	25.5	99	190	100	30.5

Source: District General Hospital of Trenggalek city, East Java Province, Indonesia, Year 2018.

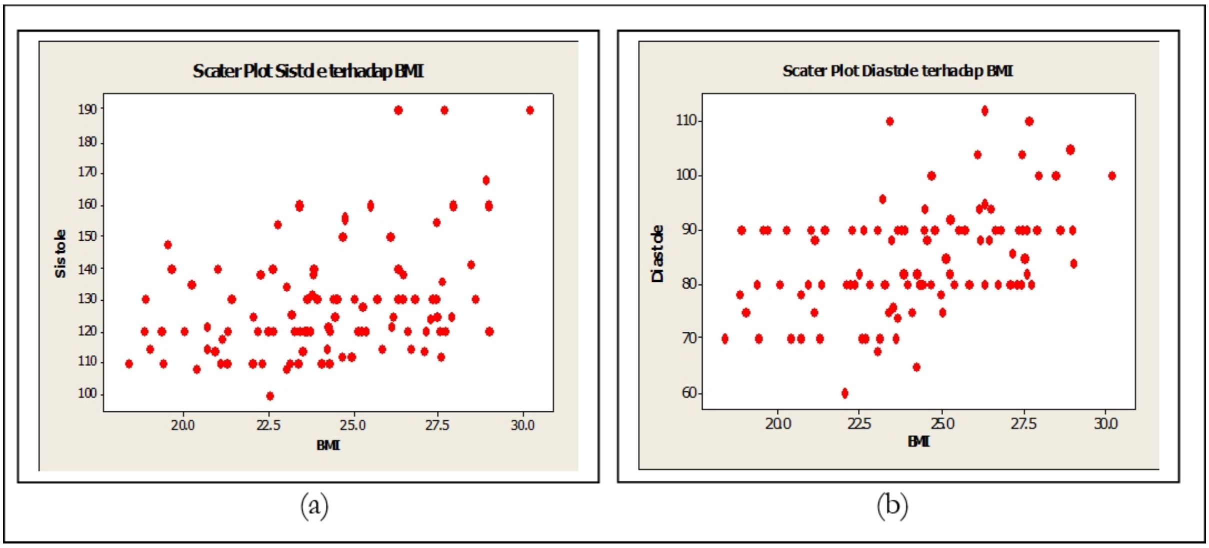


Figure 1. Plots of systolic blood pressure versus BMI (a) and plots of diastolic blood pressure versus BMI (b)

Next, we use the estimated regression function in (25) to estimate systolic and diastolic blood pressures affected by BMI. Firstly, we make scatter plots for systolic and diastolic blood pressure versus BMI as given in Fig. 1. The figure is used to look whether the non-parametric regression approach is appropriate to analyse the data. Plots given in Fig. 1 show that plots of systolic blood pressure versus BMI, and plot of diastolic blood pressure versus BMI do not follow paterns of certain regression functions as well as owned by parametric regression. Therefore, to estimate the regression functions which draw association paterns between blood pressures (systolic and diastolic) and BMI for this data, we use non-parametric regression model approach. Also, we did hypothesis testing to test correlations between responses by using the following hypothesis:

H_0 : There is no correlation between systolic and diastolic blood pressures

H_1 : There is correlation between systolic and diastolic blood pressures.

Based on this testing, we get a p-value < 0.05. It means that there is correlation between systolic and diastolic blood pressures, i.e., correlation value equals to 0.786. Therefore, to estimate the regression functions which draw association paterns between blood pressures (systolic and diastolic) and BMI for this data set, we use multi-response non-parametric regression model approaches.

Table 2. Estimated Results of Systolic and Diastolic Blood Pressure, Optimal Smoothing Parameters, Minimum GCV Value, dan MSE Value

i	$\hat{y}_{1i}$	$\hat{y}_{2i}$	i	$\hat{y}_{1i}$	$\hat{y}_{2i}$	i	$\hat{y}_{1i}$	$\hat{y}_{2i}$
1	118.4454	76.96665	34	125.2724	82.44138	67	129.2187	87.14209
2	118.9152	77.36970	35	125.2724	82.44138	68	129.4103	87.35517
3	119.0338	77.47020	36	125.4170	82.60042	69	130.0692	87.99828
4	119.1531	77.57053	37	125.5656	82.76186	70	130.0692	87.99828
5	119.5096	77.87094	38	125.5656	82.76186	71	130.3240	88.21327
6	119.6258	77.97103	39	125.7184	82.92567	72	130.5978	88.42841
7	119.7392	78.07117	40	125.8739	83.09193	73	130.5978	88.42841
8	119.9566	78.27174	41	125.8739	83.09193	74	130.5978	88.42841
9	120.2686	78.57368	42	125.8739	83.09193	75	130.8909	88.64366
10	120.4779	78.77624	43	126.0306	83.26075	76	131.2065	88.85896
11	120.6955	78.98047	44	126.1877	83.43215	77	131.5482	89.07420
12	121.0481	79.29169	45	126.1877	83.43215	78	131.9195	89.28934
13	121.0481	79.29169	46	126.3435	83.60619	79	132.3230	89.50434
14	121.3090	79.50330	47	126.3435	83.60619	80	133.7430	90.14840
15	121.4512	79.61041	48	126.6494	83.96239	81	133.7430	90.14840
16	121.6027	79.71840	49	126.8005	84.14457	82	134.8767	90.57689
17	121.6027	79.71840	50	126.8005	84.14457	83	134.8767	90.57689
18	121.9354	79.93743	51	126.9512	84.32943	84	135.5007	90.79085
19	121.9354	79.93743	52	126.9512	84.32943	85	135.5007	90.79085
20	122.1147	80.04877	53	127.1027	84.51685	86	136.1628	91.00456
21	123.2762	80.75327	54	127.2559	84.70666	87	136.1628	91.00456

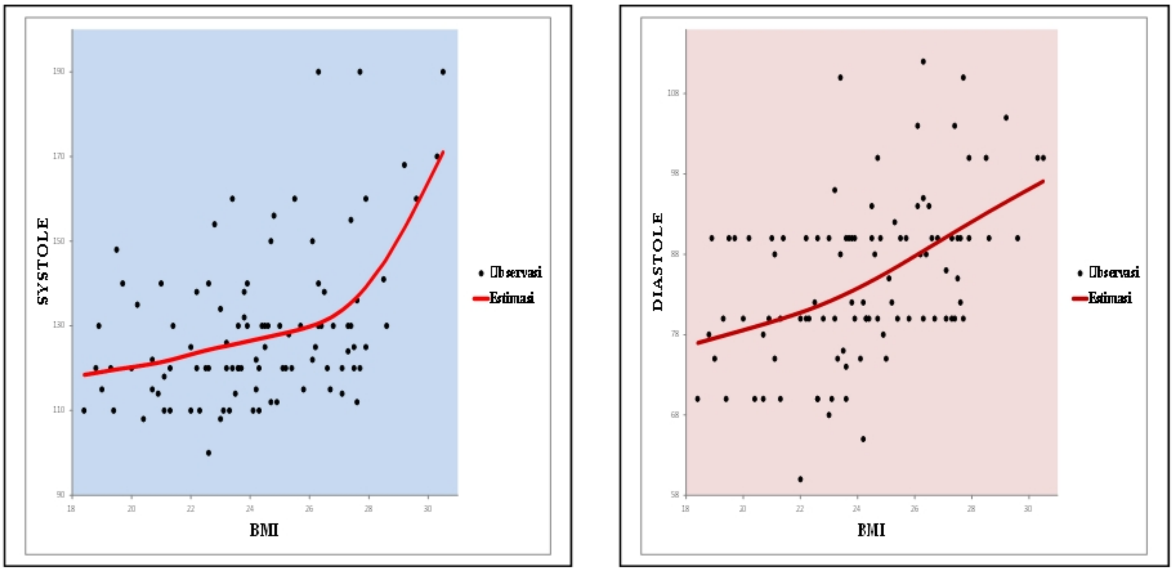
22	123.2762	80.75327	55	127.2559	84.70666	88	136.8636	91.21796
23	123.6553	81.00646	56	127.4116	84.89868	89	136.8636	91.21796
24	123.6553	81.00646	57	127.5694	85.09286	90	137.6023	91.43099
25	123.8399	81.13699	58	127.5694	85.09286	91	137.6023	91.43099
26	124.1991	81.40620	59	127.7280	85.28917	92	139.1887	91.85573
27	124.3715	81.54509	60	127.8862	85.48762	93	139.1887	91.85573
28	124.3715	81.54509	61	128.0442	85.68819	94	144.8514	93.11266
29	124.3715	81.54509	62	128.2023	85.89083	95	145.9310	93.31879
30	124.6951	81.83195	63	128.3621	86.09537	96	153.1253	94.53487
31	124.9894	82.13099	64	128.5250	86.30169	97	158.4223	95.33041
32	124.9894	82.13099	65	128.6916	86.50963	98	168.1741	96.70919
33	125.1308	82.28484	66	128.8621	86.71911	99	171.0127	97.10159

Optimal Smoothing Parameters :

$$\underset{\sim}{\lambda}_{(opt)} = (\lambda_{1(opt)}, \lambda_{2(opt)})' = (0.000682558, 0.02318668)'$$

Minimum GCV (generalised cross-validation) : 168.3557
MSE (mean squared error) : 155.1437

Further, we determine estimation values based on smoothing spline estimator for systolic blood pressure ($\hat{y}_1$) and for diastolic blood pressure ($\hat{y}_2$) based on minimum GCV value. The results of estimations are given in Table 2. Table 2 provides estimation values of systolic blood pressure ($\hat{y}_{1i}$), estimation values of diastolic blood pressure ($\hat{y}_{2i}$), optimal smoothing parameter values, i.e., $\lambda_{1(opt)} = 0.000682558$ and $\lambda_{2(opt)} = 0.02318668$, minimum GCV (generaliszed cross-validation) value of 168.3557, and MSE (mean square error) values of 155.1437. Then, plots of estimated systolic blood pressure versus BMI and estimated diastolic blood pressure versus BMI are given in Fig. 2.



(a) (b)

Figure 2. Plot of estimated systolic blood pressure versus BMI (a) and plot of estimated diastolic blood pressure versus BMI (b)

3.2. Estimating Blood Pressures Using Truncated Spline Estimator

Based on data in Table 1 and by creating open source software R code, we determine optimal knots for systolic and diastolic, optimal smoothing parameter (λ_{opt}), and minimum GCV value for order one, order two, and order three splines. The results are given in Table 3. Table 3 shows that the most minimum GCV value (i.e. 0.8548374) is reached by order one spline with optimal knots of 19 and 28.5 for systolic blood pressure and 19 for diastolic blood pressure, and optimal smoothing parameter (λ_{opt}) of 0.005. It means that the blood pressures can be estimated by order one spline with knots 19 and 28.5 for systolic blood pressure and order one spline with knots 19 for diastolic blood pressure. By applying (5) and (6) we obtain truncated splines for estimating systolic blood pressure ($\hat{y}_1$) and diastolic blood pressure ($\hat{y}_2$) as follows:

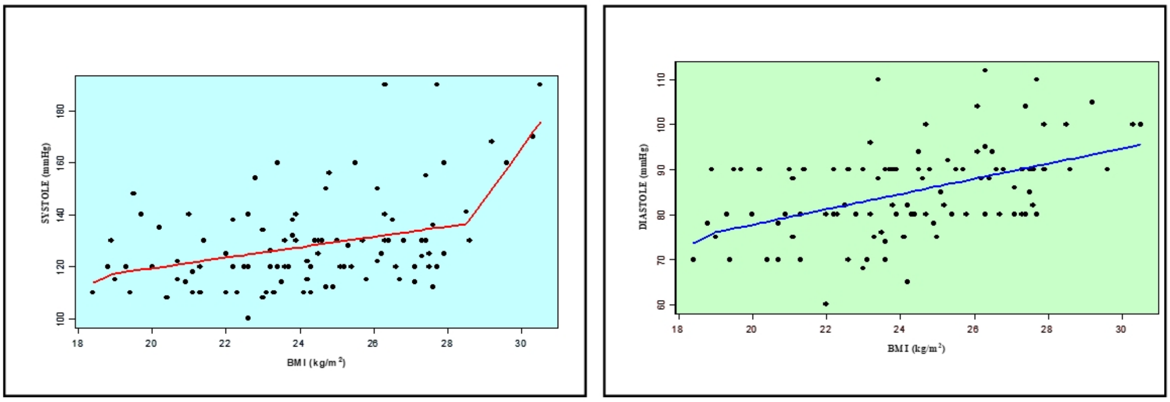
$$\hat{y}_1 = \begin{cases} 6.179029t & \text{for} \quad t < 19 \\ 79.266138 + 2.007127t & \text{for} \quad 19 \le t < 28.5 \\ -419.043537 + 19.055582t & \text{for} \quad t \ge 28.5 \end{cases}$$

$$\hat{y}_2 = \begin{cases} 4.003418t & \text{for} \quad t < 19 \\ 43.813905 + 1.697423t & \text{for} \quad 19 \le t < 28.5 \end{cases}$$

where t represents BMI value. Also, we obtain MSE (mean square error) of 155.6501, and plots of estimated systolic blood pressure ($\hat{y}_1$) and estimated diastolic blood pressure ($\hat{y}_2$) as given in Fig. 3. Finally, because of MSE (smoothing spline) = 155.1437 < 155.6501 = MSE (truncated spline) then for analyszing this data the use of multiresponsemulti-response nonparametricnon-parametric regression based on smoothing spline estimator is better than that based on truncated spline estimator.

Table 3. Order Splines, Optimal Knots, Optimal Smoothing Parameters, and Minimum GCV Values

Spline Order	Optimal Knots Systolic	Optimal Knots Diastolic	Optimal Smoothing Parameter	Minimum GCV
1	**19 ; 28.5**	**19**	**0.005**	**0.8548374**
2	28.5	19	0.013	0.8882841
3	19 ; 28.5	28.5	54	0.9034587



(a) (b)

Figure 3. Plots of estimated systolic (a) and diastolic (b) blood pressures versus BMI

4. Conclusion

Based on (25), we conclude that the estimated blood pressure model we obtained is a linear function in observation. It depends on the optimal smoothing parameter, for smoothing spline approach, and depends on optimal knot, optimal order of spline and optimal smoothing parameter, for truncated spline approach. In addition, since MSE values for smoothing spline approaches is less than that for the truncated spline approach, then smoothing spline estimators is better for estimating blood pressure models than truncated spline estimators.

5. Acknowledgment

Authors thank the Director of the Directorate of Research and Public Service, the Directorate General Reinforcing of Research and Development, the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia for financial support of this research through the doctoral dissertation research (*Penelitian Disertasi Doktor – PDD*) grant in the fiscal year 2019 with Contract Number: 702/UN3.14/LT/2019.

References

Aydin, D. (2007). A Comparison of the Non-parametric Regression Models Using Smoothing Spline and Kernel Regression. *World Acad. Sci. Eng. Technol.*, **26**, 730-734.

Aydin, D., Memmedhi, M., and Omay, R.E. (2013). Smoothing Parameter Selection for Non-parametric Regression Using Smoothing Spline. *European J. of Pure & Appl. Math.*, **6**, 222-238.

Brown, C.D., Higgins, M., Donato, K. A., Rohde, F. C., Garrison, R., Obarzanek, E., Ernst, N. D., Horan, M. (2000). Body Mass Index and the Prevalence of Hypertension and Dyslipidemia. *Obesity Research*, **8**(9), 605-619.

Cardot, Crambes, H.C., Kneip, A., and Sarda, P. (2007). Smoothing Splines Estimators in Functional Linear Regression with Errors-in-variables. *Computational Stat. Data Anal.*, **51,** 4832-4848.

Chamidah, N., Budiantara, I. N., Sunaryo, S., and Zain, I. (2012). Designing of Child Growth Chart Based on Multi-response Local Polynomial Modeling. *J. Math. Stat.*, **8**(3), 342-347.

Chamidah, N., and Lestari, B. (2016). Spline Estimator in Homoscedastic Multi-Response Non-parametric Regression Model in Case of Unbalanced Number of Observations. *Far East Journal of Mathematical Sciences* (*FJMS*), **100***,* 1433-1453.

Cox, D. D. (1983). Asymptotic for M-type Smoothing Spline. *Annals. Statistics*, **11**, 530-551.

Cox, D. D., and O'Sullivan, F. (1996). Penalized Likelihood Type Estimators for Generalized Non-parametric Regression. *J. Mult. Anal.*, **56**, 185-206.

Craven, P., and Wahba, G. (1979). Smoothing Noisy Data with Spline Function: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross Validation. *Numer. Math.*, **31**, 377-403.

Droyvold, W. B., Midthjell, K. , Nilsen, T. I. L., and Holmen, J. (2005). Change in Body Mass Index and Its Impact on Blood Pressure: A Prospective Population Study. *International Journal of Obesity*, **29**, 650-655.

Eubank, R. L. (1999). *Non-parametric Regression and Smoothing Spline*. Marcel Deker, New York.

Fernandez, F. M., and Opsomer, J. D. (2005). Smoothing Parameter Selection Methods for Non-parametric Regression with Spatially Correlated Errors. *Canadian J. Statistics*, **33**, 279-295.

Hardle, W. (1991). Applied Non-parametric Regression. Cambridge University Press. Cambridge.

IOTF-WHO (2000). *The Asia-Pasific Perspective: Redefining Obesity and Its Treatment*. Health Communications Australia Pty Limited, Sydney.

Kimeldorf, G., and Wahba, G. (1971). Some Result on Tchebycheffian Spline Functions. *J. of Math. Anal. and Appl.,* **33**, 82-95.

Koenker, R., Pin, N. G., and Portnoy, S. (1994). Quantile Smoothing Splines. *Biometrics*. **81**, 673-680.

Kumar, T. A., Sudhir, U., Srinivasan, G., and Punith, K. (2008). Association of Body Mass Index with Blood Pressure in the Elderly. *J. Indian Academy of Clinical Medicine*, **9(4)**, 274-276.

Lee, T., C., M. (2004). Improved Smoothing Spline Regression by Combines Estimates of Different Smoothness. *Stat. & Prob. Letters*, **67**, 133-140.

Lestari, B., Budiantara, I. N., Sunaryo, S., and Mashuri, M. (2009). Spline Estimator in Homoscedastic Multi-response Non-parametric Regression Model. *Proc. Indonesian Math. Soc. Int. Conf. on Math. & Its Appl.*, Oct. 12-13, Yogyakarta-Indonesia. (Available at: http://digilib.its.ac.id/ITS-Article-340312000750/18579).

Lestari, B., Budiantara, I. N., Sunaryo, S., and Mashuri, M. (2010). Spline Estimator in Multi-Response Non-parametric Regression Model with Unequal Correlation of Errors. *J. Math. Stat.*, **6**(3), 327-332.

Lestari, B., Budiantara, I. N., Sunaryo, S., and Mashuri, M. (2012). Spline Smoothing for Multi-response Non-parametric Regression Model in Case of Heteroscedasticity of Variance. *J. Math. and Stat.*, **8**, 377-384.

Lestari, B., Anggraeni, D., Saifudin, T. (2017a). Konstruksi dan Estimasi Matriks Kovariansi dalam Model Regresi Nonparametrik Multirespon Berdasarkan Estimator Smoothing Spline untuk Beberapa Kasus Ukuran Sampel. *Proceeding of National Seminar on Mathematics and Its Applications (SNMA), Faculty of Science and Technology*, *Airlangga University, Surabaya, Indonesia*, 243-250.

Lestari, B., Fatmawati, and Budiantara, I. N. (2017b). Estimasi Fungsi Regresi Nonparametrik Multirespon Menggunakan Reproducing Kernel Hilbert Space Berdasarkan Estimator Smoothing Spline. *Proceeding of National Seminar on Mathematics and Its Applications (SNMA) 2017*, *Faculty of Sciences and Technology, Airlangga University, Surabaya*, 243-250.

Lestari, B., Anggraeni, D., and Saifudin, T. (2018a). Estimation of Covariance Matrix Based on Spline Estimator in Homoscedastic Multi-Responses Non-parametric Regression Model in Case of Unbalance Number of Observations. *Far East Journal of Mathematical Sciences (FJMS)*, **108**(2), 341-355.

Lestari, B., Anggraeni, D., and Saifudin, T. (2018b). Estimation of Covariance Matrix Based on Spline Estimator in Homoscedastic Multi-Responses Non-parametric Regression Model in Case of Unbalance Number of Observations. *Far East Journal of Mathematical Sciences (FJMS)*, **108**(2), 341-355.

Lestari, B., Fatmawati, Budiantara, I. N., & Chamidah, N. (2018c). Estimation of Regression Function in Multi-Response Non-parametric Regression Model Using Smoothing Spline and Kernel Estimators. *Journal of Physics: Conference Series*, **1097**, 012091. DOI: 10.1088/1742-6596/1097/1/012091.

Lestari, B., Chamidah, N., & Saifudin, T. (2019a). Estimasi Fungsi Regresi dalam Model Regresi Nonparametrik Birespon Menggunakan Estimator Smoothing Spline dan Estimator Kernel. *Jurnal Matematika, Statistika & Komputasi (JMSK)*, **15**(2), 20-24. DOI: 10.20956/jmsk.v15i2.5565.

Lestari, B., Fatmawati, and Budiantara, I. N. (2019b). Spline Estimator and Its Asymptotic Properties in Multi-response Non-parametric Regression Model. *Songklanakarin Journal of Science and Technology*, In Press, (SJST-2018-0306.R1). http://rdo.psu.ac.th/sjstweb/ArticleInPress.php. 17 July 2019.

Lestari, B., Fatmawati, and Budiantara, I. N. (2019c). Smoothing Spline Estimator in Multi-response Non-parametric Regression for Predicting Blood Pressures and Heart Rate, *International Journal of Academic and Applied Research* (IJAAR), **3**(9), 1-8.

Liu, A., Tong, T., and Wang, Y. (2007). Smoothing Spline Estimation of Variance Functions. *Journal of Comput. Graphical Stat.*, **16**, 312-329.

Nanaware, N. L., Gavkare, A. M., and Surdi, A. D. (2011). Study of Correlation of Body Mass Index (BMI) with Blood Pressure in School Going Children and Adolescents. *International Journal of Recent Trends in Science and Technology*, **1**(1), 20-26.

Oehlert, G. W. (1992). Relaxed Boundary Smoothing Spline. *Annals. Statistics*, **20**, 146-160.

Roka, R., Michimi, A., dan Macy, G. (2015). Associations Between Hypertension and Body Mass Index and Waist Circumference in U.S. Adults: A Comparative Analysis by Gender. High Blood Pressure & Cardiovascular Prevention, **22**(3), 265-273.

Schimek, M. G., (2000). *Smoothing and Regression*, John Wiley & Sons. New York.

Tesfaye, F., Nawi, NG., Van Minh, H., Byass, P., Berhane, Y., Bonita, R., and Wall, S. (2007), Association Between Body Mass Index and Blood Pressure Across Three Populations in Africa and Asia. *Journal Human Hypertension*, **21**, 28-37.

Wahba, G. (1983). Bayesian Confidence Intervals for the Cross-Validated Smoothing Spline. *J. Royal Stat. Society, Series B.*, **45**, 133-150.

Wahba, G. (1985). A Comparison of GCV and GML for Choosing the Smoothing Parameter in the Generalized Spline Smoothing Problem. *Annals Stat.*, **13**, 1378-1402.

Wahba, G. (1990). *Spline Models for Observational Data*. SIAM Philadelphia. Pennsylvania.

Wahba, G. (2000). An Introduction to Model Building with Reproducing Kernel Hilbert Spaces. *Research Report*, Wisconsin University, USA.

Wang, Y. (1998). Smoothing Spline Models with Correlated Random Errors. *Journal of American Stat. Assoc.*, **93**, 341-348.

Wang, Y. (2001). Smoothing splines: Methods and Applications. CRC Press, New York.

Wang, Y., Guo, W., Brown, W. B. (2000). Spline Smoothing for Bivariate Data with Applications to Association Between Hormones. *Statistica Sinica*, **10**, 377-397.

Watson, G. S. (1964). Smooth Regression Analysis. *Sankhya Series A.*, **26**, 359-372.