BAREKENG: Journal of Mathematics and Its Application
December 2022 Volume 16 Issue 4 Page 1497–1504
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol16iss4pp1497-1504

THE USE OF PENALIZED WEIGHTED LEAST SQUARE TO OVERCOME CORRELATIONS BETWEEN TWO RESPONSES

Anna Islamiyati [1*], Anisa[2], Muhammad Zakir[3], Nasrah Sirajang[4], Ummi Sari[5], Fajar Affan[6], Muhammad Jayzul Usrah[7]

[1,2,4,6,7]Department of Statistics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Jl. Perintis Kemerdekaan KM. 10, Makassar, 90245, Indonesia
[3]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Jl. Perintis Kemerdekaan KM. 10, Makassar, 90245, Indonesia
[5]Hasanuddin University Teaching Hospital, Jl. Perintis Kemerdekaan KM. 10, Makassar, 90245, Indonesia

Corresponding author's e-mail: [1] annaislamiyati701@gmail.com*

Abstract. *The non-parametric regression model can consider two correlated responses. However, for these conditions, we cannot use the usual estimation process because there are violations of assumptions. To solve this problem, we simultaneously use a penalized weighted least square involving knots, smoothing parameters, and weighting in the estimation criteria. The estimation process involves a weighted criteria matrix $\frac{1}{n}$ in the estimation criteria. Estimation results show that the estimated two-response non-parametric regression function with penalized spline is a linear estimation class in y response observation and depends on the knot point and smoothing parameter. Furthermore, the use of the model on toddler growth data shows changes in the weight and height gain pattern. The pattern segmentation that experienced a gradual increase was age 7-43 months for weight and age 6-54 months for height.*

Keywords: *age, height, knots, two responses, weight, weighted penalized.*

Article info:
Submitted: 22nd August 2022 *Accepted: 21st November 2022*

How to cite this article:
A. Islamiyati, Anisa, M. Zakir, N. Sirajang, U. Sari and F. Affan, "THE USE OF PENALIZED WEIGHTED LEAST SQUARE TO OVERCOME CORRELATION BETWEEN TWO RESPONSES", *BAREKENG: J. Math. & App.*, vol. 16, iss. 4, pp. 1497-1504, Dec., 2022

1. INTRODUCTION

Nonparametric regression is used when we need more information about the relationship patterns that occur in the data. Usually, an easy step to detect the pattern of relationships between responses and predictors is the scatter plot. We can use a nonparametric regression approach when the data plot does not follow a parametric pattern [1]. Some estimators that can be used include spline [2], local polynomial [3], kernel [4], Fourier series [5], and wavelet [6]. The spline estimator has become one of the most developed estimators because it has high flexibility in modeling data. Some spline estimator developments include truncated spline [7], smoothing spline [8], penalized spline [9], robust spline [10], b spline [11], quantile spline [12], multiresponse spline [13], and others.

This article discusses the estimation method of regression parameters in a model that contains two correlated responses. This is very important to study because there are many real data that contain more than one response and correlate with each other. For example, measurements of blood sugar in diabetic patients during the day and night [14] and the child's weight and height [15]. We use penalized weighted least square (PWLS) estimation criteria that involve weighting, knot points, and smoothing parameters simultaneously. That is the development of penalized least square, which is a standard criterion in the smoothing process. The least square method is used when the data does not experience a violation of the assumption that only involves knots and smoothing parameters. PWLS criteria have been developed by researchers for their use in nonparametric regression models. Some use estimators for smoothing spline [16] and penalized spline [17]. Besides the difference estimator, usually with the use of PWLS by some researchers, it is different in the weighting value used. Some use a weighting of the covariance variance matrix [18], number of parameters [19], and bootstrap weighting [20]. However, in simpler conditions, we can use weights based on the number of samples, and that can provide accurate estimation results. Therefore, we use the PWLS criterion by weighting the sample size $\left(\frac{1}{n}\right)$ in estimating the spline regression coefficient [21].

In this article, the theoretical study we developed is applied to data on the growth of children under five in South Sulawesi Province. The use of nonparametric regression with PWLS aims to model children's weight and height, which are correlated with each other based on age. Modeling of weight and height has been previously investigated using local polynomial regression models [22] and kernel [23] for data on children under five in East Java Province. This article is divided into four parts. The second part describes the method used as the theoretical basis of this research. It is related to a nonparametric regression model for two correlated responses. The regression model was estimated using a penalized spline containing knot points and smoothing parameters in controlling the smoothing of the regression curve. The third part contains the study's results, which include the estimation of nonparametric regression parameters for two responses with PWLS and the application of the model to toddler growth data. In addition to toddler data, the theoretical study that we make here is expected to be one of the reference methods for studying real data that has conditions or assumptions that are in accordance with the method. Furthermore, the fourth section contains a brief conclusion from the results of this study.

2. RESEARCH METHODS

The regression model that contains two responses means that there are paired data $\left(y_{1.i}, y_{2.i}\right)$ which are assumed to be correlated with each other, with the predictor variable $\left(x_1, x_2, \ldots, x_n\right)$ which is an unknown form of a relationship pattern. The relationship between variables $y_{r.i}$ and x_i follows the nonparametric regression model as follows.

$$y_{r.i} = f_r(x_i) + \varepsilon_{r.i} \tag{1}$$

where $i = 1, 2, \ldots, n$ is the observed subject, $r = 1, 2$ is the number of response variables, $f_r(x_i)$ is a regression function contained in a Sobolev space, ε is the measurement residual which is assumed to be independent of the mean 0 and variance ε, which ε_1 and ε_2 are correlated.

The function $f_r(x_i)$ in Equation (1) is approximated using segmented polynomial functions like the following equation.

$$f_r(x_i)=\sum_{m=0}^{p+k}\beta_{r.m}g_{r.m}(x_i) \tag{2}$$

where a variable $\beta_{r.m}=(\beta_{r.0},\beta_{r.1},...,\beta_{r.(p+k)})^T$ is a regression parameter vector, p is a regression order, k is the number of knots and $g_{r.m}=(g_{r.0},g_{r.1},...,g_{r.(p+k)})$ is a function that has a definition like the following.

$$g_{r.m}(x_i)=\begin{cases} x_i^m & ; \quad 0\le m\le p \\ (x_i-\tau_{r.(m-p)})_+^p & ; \quad p+1\le m\le p+k \end{cases} \tag{3}$$

and

$$(x_i-\tau_{r.(m-p)})_+^p=\begin{cases} (x_i-\tau)^p & ; \quad x_i\ge\tau \\ 0 & ; \quad x_i<\tau \end{cases} \tag{4}$$

Based on Equations (3) and (4), the function in Equation (2) can be stated as follows.

$$f_r(x_i)=\sum_{m=0}^{p}\beta_{r.m}x_i^m+\sum_{l=1}^{k}\beta_{\text{r}.(p+l)}(x_i-\tau_{r.l})_+^p \tag{5}$$

where $\tau_{r.l}$ is the knot points of $l=1,2,\ldots,k$ in the r^{th} response.

Equation (5) on each response r can be stated as follows.

$$\underset{\sim}{f}=\mathbf{X}\underset{\sim}{\beta} \tag{6}$$

where $\underset{\sim}{f}=\left(\underset{\sim}{f}_1,\underset{\sim}{f}_2\right)^T$ and $\underset{\sim}{\beta}=\left(\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2\right)^T$. The vector $\underset{\sim}{f}$ is a regression function in the 1st and 2nd responses, with $\underset{\sim}{f}_1=\left(f_1(x_1),f_1(x_2),\ldots,f_1(x_n)\right)^T$ and $\underset{\sim}{f}_2=\left(f_2(x_1),f_2(x_2),\ldots,f_2(x_n)\right)^T$. In the regression coefficient vector $\underset{\sim}{\beta}$, it also contains the regression coefficient vector for the 1st and 2nd responses with $\underset{\sim}{\beta}_1=\left(\beta_{1.0},\beta_{1.1},\ldots,\beta_{1.(p+k)}\right)^T$ and $\underset{\sim}{\beta}_2=\left(\beta_{2.0},\beta_{2.1},\ldots,\beta_{2.(p+k)}\right)^T$. For matrix $\mathbf{X}=\text{diag}(\mathbf{X}_1,\mathbf{X}_2)$ is a predictor matrix that contains knots at the 1st and 2nd responses.

Furthermore, the non-parametric two response regression model in Equation (1) can also be expressed in the form of a vector referring to equation (6), namely:

$$\underset{\sim}{y}=\mathbf{X}\underset{\sim}{\beta}+\underset{\sim}{\varepsilon} \tag{7}$$

Equation (7) can also be written in the form of a matrix as follows.

$$\begin{bmatrix}\underset{\sim}{y}_1\\ \underset{\sim}{y}_2\end{bmatrix}=\begin{bmatrix}\mathbf{X_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{X_2}\end{bmatrix}\begin{bmatrix}\underset{\sim}{\beta}_1\\ \underset{\sim}{\beta}_2\end{bmatrix}+\begin{bmatrix}\underset{\sim}{\varepsilon}_1\\ \underset{\sim}{\varepsilon}_2\end{bmatrix} \tag{8}$$

Each element in Equation (8) can be explained by $\underset{\sim}{y}_1=(y_{1.1},y_{1.2},...,y_{1.n})^T$ is the 1st response vector, $\underset{\sim}{y}_2=(y_{2.1},y_{2.2},...,y_{2.n})^T$ is the 2nd response vector, $\mathbf{X}_1$ is the $\mathbf{X}$ matrix for the 1st response and $\mathbf{X}_2$ is the $\mathbf{X}$ matrix for the 2nd response.

$$\mathbf{X_1} = \begin{bmatrix} 1 & x_1^1 & x_1^2 & \cdots & x_1^p & (x_1-\tau_{1.1})_+^p & \cdots & (x_1-\tau_{1.k})_+^p \\ 1 & x_2^1 & x_2^2 & \cdots & x_2^p & (x_2-\tau_{1.1})_+^p & \cdots & (x_2-\tau_{1.k})_+^p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n^1 & x_n^2 & \cdots & x_n^p & (x_n-\tau_{1.1})_+^p & \cdots & (x_n-\tau_{1.k})_+^p \end{bmatrix},$$

$$\mathbf{X_2} = \begin{bmatrix} 1 & x_1^1 & x_1^2 & \cdots & x_1^p & (x_1-\tau_{2.1})_+^p & \cdots & (x_1-\tau_{2.k})_+^p \\ 1 & x_2^1 & x_2^2 & \cdots & x_2^p & (x_2-\tau_{2.1})_+^p & \cdots & (x_2-\tau_{2.k})_+^p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n^1 & x_n^2 & \cdots & x_n^p & (x_n-\tau_{2.1})_+^p & \cdots & (x_n-\tau_{2.k})_+^p \end{bmatrix}.$$

Furthermore, $\underset{\sim}{\beta}_1 = \left(\beta_{1.0}, \beta_{1.1}, \ldots, \beta_{1.(p+k)}\right)^T$ is a regression coefficient vector in the 1st response and $\underset{\sim}{\beta}_2 = \left(\beta_{2.0}, \beta_{2.1}, \ldots, \beta_{2.(p+k)}\right)^T$ is a regression coefficient vector in the 2nd response, $\underset{\sim}{\varepsilon}_1 = (\varepsilon_{1.1}, \varepsilon_{1.2}, \ldots, \varepsilon_{1.n})^T$ is an error vector in the 1st response and $\underset{\sim}{\varepsilon}_2 = (\varepsilon_{2.1}, \varepsilon_{2.2}, \ldots, \varepsilon_{2.n})^T$ is an error vector in the 2nd response.

3. RESULTS AND DISCUSSION

3.1. Estimation of the Regression Coefficient with Penalized Weighted Least Square

Estimation of parameter $\underset{\sim}{\beta}$ as a coefficient in the two-response nonparametric regression model through Penalized Weighted Least Square (PWLS) optimization. The PWLS equation contains the goodness of fit and the penalty functions which can be stated as follows.

$$\mathbf{L} = (2n)^{-1}(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta})^T \mathbf{W}(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta}) + \lambda \underset{\sim}{\beta}^T \mathbf{D}\underset{\sim}{\beta} \tag{9}$$

The goodness of fit function is given by $(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta})^T(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta})$ and the penalty function is given by $\lambda \underset{\sim}{\beta}^T \mathbf{D}\underset{\sim}{\beta}$. Furthermore, the **W** matrix is a weighted of matrix given by $\mathbf{W} = \mathrm{diag}\left(\frac{1}{2n}\right)$ of size $2n \times 2n$ or in the form of a matrix as follows:

$$\mathbf{W} = \begin{bmatrix} \frac{1}{2n}\mathbf{I} & | & \mathbf{0} \\ -- & | & -- \\ \mathbf{0} & | & \frac{1}{2n}\mathbf{I} \end{bmatrix}$$

In PWLS, there is a matrix **D** as a diagonal matrix (0,1), i.e. $\mathbf{D} = \begin{bmatrix} \mathbf{D_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D_2} \end{bmatrix}$, where $\mathbf{D}_1 = \mathrm{diag}\left(a_{1.11}, \ldots, a_{1.pp}, a_{1.(p+1)(p+1)}, a_{1.(p+2)(p+2)}, \ldots, a_{1.(p+k+1)(p+k+1)}\right)$ is the **D** matrix in the 1st response and $\mathbf{D}_2 = \mathrm{diag}\left(a_{2.11}, \ldots, a_{2.pp}, a_{2.(p+1)(p+1)}, a_{2.(p+2)(p+2)}, \ldots, a_{2.(p+k+1)(p+k+1)}\right)$ is the **D** matrix in the 2nd response. The value of $a_{11} = \ldots = a_{pp} = a_{(p+1)(p+1)} = 0$, and $a_{(p+2)(p+2)} = \ldots = a_{(p+k+1)(p+k+1)} = 1$.

The estimation process $\underset{\sim}{\beta}$ is done by differentiating **L** for $\underset{\sim}{\beta}$.

$$\begin{aligned} \mathbf{L} &= (2n)^{-1}(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta})^T \mathbf{W}(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta}) + \lambda \underset{\sim}{\beta}^T \mathbf{D}\underset{\sim}{\beta} \\ &= (2n)^{-1}(\underset{\sim}{y}^T - \underset{\sim}{\beta}^T\mathbf{X}^T)\mathbf{W}(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta}) + \lambda \underset{\sim}{\beta}^T \mathbf{D}\underset{\sim}{\beta} \\ &= (2n)^{-1}(\underset{\sim}{y}^T\mathbf{W}\underset{\sim}{y} - 2\underset{\sim}{y}^T\mathbf{W}\mathbf{X}\underset{\sim}{\beta} + \underset{\sim}{\beta}\mathbf{X}^T\mathbf{W}\mathbf{X}\underset{\sim}{\beta}^T) + \lambda \underset{\sim}{\beta}^T \mathbf{D}\underset{\sim}{\beta} \end{aligned}$$

The minimum value of **L** in Equation (9) is obtained when $\frac{\partial \mathbf{L}}{\partial \underset{\sim}{\beta}} = \mathbf{0}$:

$$-2\mathbf{X}^{\mathrm{T}}\mathbf{W}\underset{\sim}{y} + 2\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\underset{\sim}{\beta} + 2\lambda\mathbf{D}\underset{\sim}{\beta} = \mathbf{0}$$

$$\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}\underset{\sim}{\beta} + \lambda\mathbf{D}\underset{\sim}{\beta} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\underset{\sim}{y} \tag{10}$$

$$\hat{\underset{\sim}{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X} + \lambda\mathbf{D})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}\underset{\sim}{y}$$

As a result, the form of nonparametric regression estimation of two responses with penalized spline becomes as follows:

$$\hat{\underset{\sim}{f}} = \mathbf{X}\hat{\underset{\sim}{\beta}}$$

Based on Equation (10), it can be written:

$$\hat{\underset{\sim}{f}} = \mathbf{X}((\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X} + \lambda\mathbf{D})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}\underset{\sim}{y}) \tag{11}$$

Equation (11) can also be simplified as follows.

$$\hat{\underset{\sim}{f}} = \mathbf{H}(\lambda)\underset{\sim}{y} \text{ , where } \mathbf{H}(\lambda) = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X} + 2\lambda\mathbf{D})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}$$

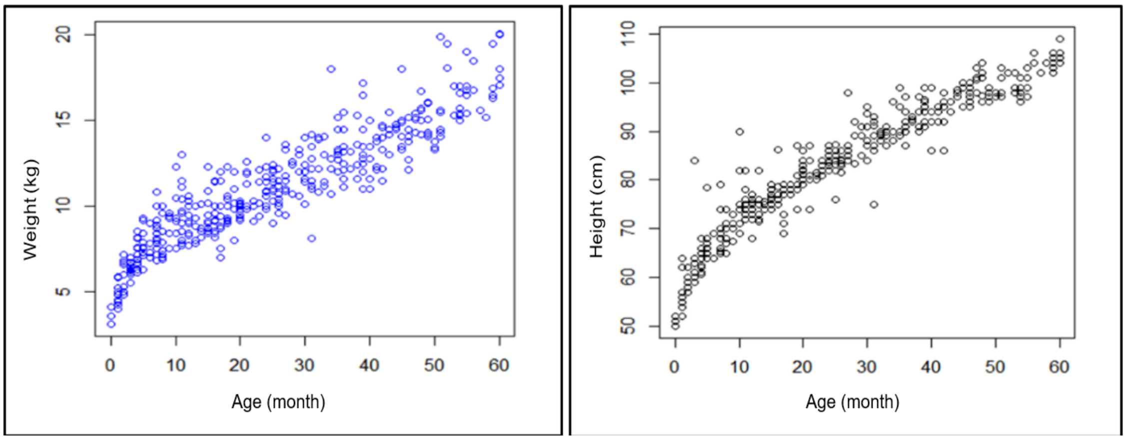
The knot point in matrix **X** is a combination of curves that shows the pattern of changes in curve behavior at different intervals. The lambda parameter λ is a smoothing parameter that controls the smoothness of the function of the data. If λ is large, then the estimation of the function obtained is smoother, and preferably if λ is small, then the estimate of the function obtained is even more rougher [20]. The choice of knot points and λ parameters is made by trial and error by selecting knot points that are at data intervals in the predictor variables. The best model is a model that comes from optimal knots and lambda. One method used to select the optimal knot point and lambda is the Generalized Cross Validation (GCV) method. The optimum smoothing parameter λ is determined based on the minimum Generalized Cross Validation (GCV) value. The GCV method can be defined as follows.

$$GCV(\tau,\lambda) = \frac{MSE(\tau,\lambda)}{(2n^{-1}tr[\mathbf{I}-\mathbf{H}(\tau,\lambda)])^{2}} \tag{12}$$

where $MSE(\tau,\lambda) = 2n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ is the mean square error and $\mathbf{H}(\lambda) = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X} + 2\lambda\mathbf{D})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}$ is the hat matrix.

3.2 Applications on Toddler Growth Data

The nonparametric regression model with PWLS was applied to the growth data of children under five as measured by weight and height based on the age of the toddler. Figure 1 shows a scatter plot of weight and height data by age.



(a) **(b)**

Figure 1. Scatter Plot of Weight by Age (a) and Height by Age (b)

Based on Figure 1, we can see that both data plots of weight and height tend to increase with age. We use nonparametric regression to be more flexible in modeling the data. This is due to the placement of knot points that can be selected according to the data. In addition, the involvement of smoothing parameters can also result in a smoother and more accurate estimation of the regression curve.

Next, we tested the correlation between body weight and height and obtained a correlation of 0.8. This means that there is a very strong correlation between the weight and height of toddlers. Therefore, these data were analyzed using PWLS nonparametric regression. The optimal model selection is based on the minimum GCV value and the optimal model has obtained in the second order with 2-knot points. The bi-response nonparametric regression equation with PWLS for the weight factor is as follows.

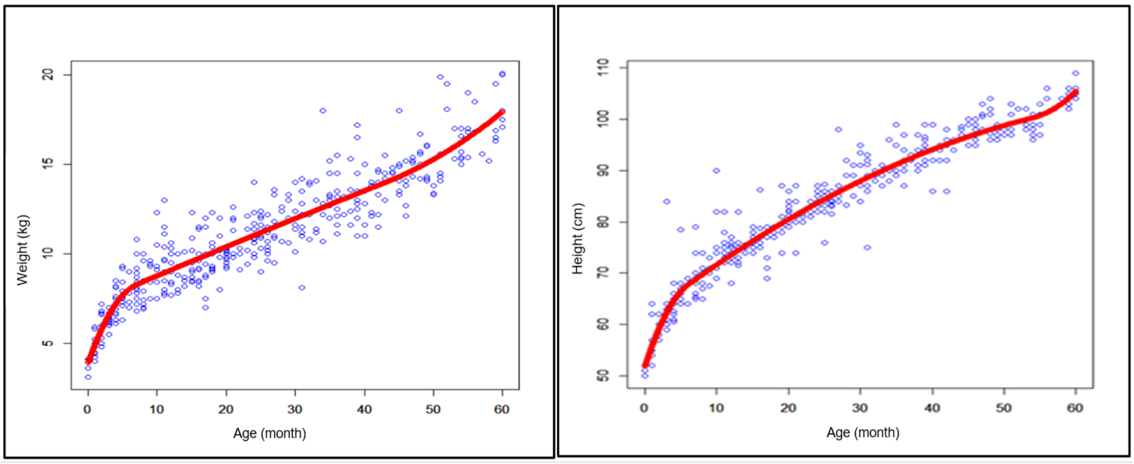
$$\hat{y}_1 = 3{,}8967 + 1{,}0871x - 0{,}0659x^2 + 0{,}0657(x-7)_+^2 + 0{,}0049(x-43)_+^2$$

For the height factor, the nonparametric two response regression equation with PWLS is obtained as follows.

$$\hat{y}_2 = 51{,}8304 + 4{,}3016x - 0{,}2745x^2 + 0{,}2676(x-6)_+^2 + 0{,}0925(x-54)_+^2$$

The PWLS biresponse nonparametric regression equation for weight and height corresponds to the estimated regression curve shown in Figure 2 below.

Based on Figure 2 (a) and the nonparametric two-response regression model with PWLS, the results show that the average body weight increases with the age of the toddler. In general, weight gain in infancy (0-60 months) follows 3 different patterns. For the age of less than 7 months, toddlers experience a very rapid weight gain, then it increases slowly at the age of 7-43 months, and again increases rapidly after the age of 43 months. Furthermore, in Figure 2 (b) for height, the average height also increases with the age of the toddler. The increase in height during toddlerhood follows 3 patterns, namely at the age of less than 6 months, toddlers experience a very rapid increase in height and tend to slow down their growth at the age of 6 to less than 54 months. Furthermore, the height again quickly increased at the age of 54 months and over. These results indicate that changes in weight and height gain are at different age intervals. This indicates that a child's weight and height have different trends at their growing age. Therefore, in the effort to prevent stunting, it is necessary to work on each pattern segmentation that is considered slowing child growth and development.



(a) (b)

Figure 2. Estimation of two response nonparametric regression curve with PWLS on toddler growth data a) weight and b) height

4. CONCLUSIONS

Data that contains two responses and correlate with each other cannot be modeled using the usual approach in non-parametric regression. This will be very much obtained in real data so that the error will correlate. In this article, we use PWLS criteria in estimating the regression parameters. PWLS considers the goodness of fit and penalty functions. The involvement of weighting, knot point and smoothing parameters can overcome the correlation that occurs in the data. The weighting we use in the goodness of fit function is $\frac{1}{2n}$. Of course, other weighting can also be used for other research material. The results of estimating the regression parameters through PWLS show that the model is very dependent on knots, smoothing parameters and weighting used. Therefore, we need to find the optimal parameter values based on the minimum GCV value. In addition to the GCV method, there are several other methods known in nonparametric regression to find optimal parameter values, including Generalized Maximum Likelihood (GML), Unbiassed Risk (UBR), and Cross Validation (CV).

The application of the model to toddler growth data, namely weight and height based on the age of the baby, has a different upward trend in the three segments of the pattern of change. Weight increases drastically in newborns up to the age of before 7 months, then increases slowly at the age of 7-43 months, and again increases rapidly at the age of over 43 months. For height, it also increases dramatically in newborns up to the age of 6 months, then the baby's height gain slows down at the age of 6-54 months, and when the age is over 54 months, the toddler's height begins to increase rapidly. Therefore, it is necessary to pay special attention to each segment that shows slowing growth. To avoid stunting, it is recommended to give extra attention to babies aged 6-54 months.

ACKNOWLEDGEMENT

Many thanks to the Directorate of Research, Technology, and Community Service, Directorate General of Higher Education, Research and Technology of the Ministry of Education, Culture, Research, and Technology for National Competitive Basic Research with research contract No: 090/E5/PG .02.00.PT/2020 dated May 10, 2022.

REFERENCES

[1] A. Islamiyati, Raupong, and Anisa, “Use of penalized spline linear to identify change in pattern of blood sugar based on the weight of diabetes patients,” *Int. J. Acad. Appl. Res.*, vol. 3, no. 12, pp. 75–78, 2019.

[2] A. Islamiyati, Fatmawati, and N. Chamidah, “Changes in blood glucose 2 hours after meals in type 2 diabetes patients based on length of treatment at Hasanuddin University Hospital, Indonesia,” *Rawal Med. J.*, vol. 45, no. 1, pp. 31–34, 2020.

[3] N. Chamidah, K. H. Gusti, E. Tjahjono, and B. Lestari, “Improving of classification accuracy of cyst and tumor using local polynomial estimator,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 3, pp. 1492–1500, 2019, doi: 10.12928/TELKOMNIKA.V17I3.12240.

[4] I. N. Budiantara, V. Ratnasari, M. Ratna, W. Wibowo, N. Afifah, and D. P. Rahmawati, “Modeling percentage of poor people in Indonesia using kernel and fourier series mixed in nonparametric regression,” *Investig. Operacional*, vol. 40, no. 4, pp. 538–551, 2019.

[5] N. P. A. M. Mariati, I. N. Budiantara, and V. Ratnasari, “Modeling poverty percentages in the papua islands using fourier series in nonparametric regression multivariable,” *J. Phys. Conf. Ser.*, vol. 1397, no. 1, pp. 1–7, 2019, doi: 10.1088/1742-6596/1397/1/012071.

[6] X. C. Zhou and J. G. Lin, “Wavelet estimator in nonparametric regression model with dependent error’s structure,” *Commununication Stat. – Theory Methods*, vol. 43, no. 22, pp. 4707–4722, 2014.

[7] A. Islamiyati, A. Kalondeng, N. Sunusi, M. Zakir, and A. K. Amir, “Biresponse nonparametric regression model in principal component analysis with truncated spline estimator,” *J. King Saud Univ. - Sci.*, vol. 34, no. 3, pp. 1–9, 2022, doi: 10.1016/j.jksus.2022.101892.

[8] B. Lestari, Fatmawati, I. N. Budiantara, and N. Chamidah, “Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches,” *J. Phys. Conf. Ser.*, vol. 1397, no. 1, 2019, doi: 10.1088/1742-6596/1397/1/012064.

[9] A. Islamiyati, N. Sunusi, A. Kalondeng, F. Fatmawati, and N. Chamidah, “Use of two smoothing parameters in penalized spline estimator for bi-variate predictor non-parametric regression model,” *J. Sci. Islam. Repub. Iran*, vol. 31, no. 2, pp. 175–183, 2020, doi: 10.22059/JSCIENCES.2020.286949.1007435.

[10] Musafirah, A. Islamiyati, and N. Sunusi, “Estimation of penalized spline linear regression models through robust M estimator,” *Int. J. Acad. Appl. Res.*, vol. 5, no. 11, pp. 166–168, 2021.

[11] Q. Zou and Z. Zhu, “M-estimators for single-index model using b-spline,” *Metrika*, vol. 77, pp. 225–246, 2014.

[12] W. N. A. Puteri, A. Islamiyati, and A. Anisa, “Penggunaan regresi kuantil multivariat pada perubahan trombosit pasien demam berdarah dengue,” *ESTIMASI J. Stat. Its Appl.*, vol. 1, no. 1, pp. 1–9, 2020, doi: 10.20956/ejsa.v1i1.9224.

[13] A. Islamiyati, “Spline longitudinal multi-response model for the detection of lifestyle-based changes in blood glucose of diabetic patients,” *Curr. Diabetes Rev.*, vol. 18, no. 7, pp. 98–104, 2021, doi: 10.2174/1573399818666211117113856.

[14] A. Islamiyati, Fatmawati, and N. Chamidah, “Penalized spline estimator with multi smoothing parameters in bi-response multi-predictor nonparametric regression model for longitudinal data,” *Songklanakarin J. Sci. Technol.*, vol. 42, no. 4, pp. 897–909,

2020.
[15] W. Ramadan, N. Chamidah, B. Zaman, L. Muniroh, and B. Lestari, “Standard growth chart of weight for height to determine wasting nutritional status in East Java based on semiparametric least square spline estimator,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 546, no. 5, pp. 1–8, 2019, doi: 10.1088/1757-899X/546/5/052063.
[16] R. Hidayat, I. N. Budiantara, B. W. Otok, and V. Ratnasari, “The regression curve estimation by using mixed smoothing spline and kernel (MsS-K) modeltle,” *Commun. Stat. - Theory Methods*, vol. 50, no. 17, pp. 3942–3953, 2021.
[17] A. Islamiyati, Fatmawati, and N. Chamidah, “Ability of covariance matrix in bi-response multi-prredictor penalized spline model through longitudinal data simulation,” *Int. J. Acad. Appl. Res.*, vol. 3, no. 3, pp. 8–11, 2019.
[18] A. Islamiyati, Fatmawati, and N. Chamidah, “Estimation of covariance matrix on bi-response longitudinal data analysis with penalized spline regression,” *J. Phys. Conf. Ser.*, vol. 979, no. 1, pp. 1–7, 2018, doi: 10.1088/1742-6596/979/1/012093.
[19] A. Islamiyati, A. Kalondeng, and U. Sari, “Estimating the confidence interval of the regression coefficient of the blood sugar model through a multivariable linear spline with known variance,” *Stat. Transit.*, vol. 23, no. 1, pp. 201–212, 2022, doi: 10.21307/stattrans-2022-012.
[20] V. Makarenkov, A. Boc, J. Xie, P. Peres-Neto, F. J. Lapointe, and P. Legendre, “Weighted bootstrapping: A correction method for assessing the robustness of phylogenetic trees,” *BMC Evol. Biol.*, vol. 10, no. 1, pp. 3–16, 2010, doi: 10.1186/1471-2148-10-250.
[21] A. Islamiyati, “Model regresi spline untuk data longitudinal dengan penalized likelihood pada pasien diabeted mellitus tipe II di Rumah Sakit Wahidin Sudirohusodo Makassar,” 2009.
[22] N. Chamidah, E. Tjahjono, A. R. Fadilah, and B. Lestari, “Standard growth charts for weight of children in East Java using local linear estimator,” *J. Phys. Conf. Ser.*, vol. 1097, no. 1, pp. 1–6, 2018, doi: 10.1088/1742-6596/1097/1/012092.
[23] N. Chamidah and T. Saifudin, “Estimation of children growth curve based on kernel smoothing in multi-response nonparametric regression,” *Appl. Math. Sci.*, vol. 7, no. 37–40, pp. 1839–1847, 2013, doi: 10.12988/ams.2013.13168.