MULTIRESPONSE SEMIPARAMETRIC REGRESSION MODEL APPROACH TO STANDARD GROWTH CHARTS DESIGN FOR ASSESSING NUTRITIONAL STATUS OF EAST JAVA TODDLERS

NUR CHAMIDAH1,2,*, BADRUS ZAMAN1, LAILATUL MUNIROH3, BUDI LESTARI2,4

1Department of Mathematics, Faculty of Science and Technology, Airlangga University, Indonesia
2Research Group of Statistical Modeling in Life Science, Fac. Sci. & Technol., Airlangga University, Indonesia
3Department of Health Nutrition, Faculty of Public Health, Airlangga University, Indonesia
4Department of Mathematics, Faculty of Mathematics and Natural Science, The University of Jember, Indonesia

Copyright © 2023 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: A nutritional status of toddlers characterized by a lack of weight based on anthropometric index weight-for-age (W/A) is called underweight. In Indonesia, the anthropometric index is recorded on a Card Towards Health called KMS (Kartu Menuju Sehat) which refers to WHO–2005 Standard Growth Charts (WHO–2005 SGC). Samples used to design the WHO–2005 SGC were toddlers taken from Brazil, Ghana, India, Norway, Oman, and USA, that physically they are different from the Indonesian toddlers especially toddlers from East Java. This study aims to design Percentile Standard Growth Charts (Percentile SGC) of W/A based on samples of toddlers taken from East Java by using both least square spline (LS-spline) and local linear estimators of a multiresponse semiparametric regression (MSR) model. Those obtained Percentile SGC would be used to assess the nutritional status of toddlers in East Java by comparing them with the WHO–2005 SGC. Results show that these Percentile SGC based on both LS-spline and

*Corresponding author
E-mail address: nur-c@fst.unair.ac.id
Received November 07, 2022
local linear estimators give a lower reference than WHO–2005 SGC such that percentages of malnutrition and less-nutrition status categories provided by these Percentile SGC are also less than those provided by WHO–2005 SGC.

In general, the Percentile SGC of W/A for boy and girl toddlers in East Java have the evaluator standard and have lower standard than the WHO–2005 SGC.

Keywords: multiresponse semiparametric regression; standard growth charts of toddlers; nutritional status; percentile standard growth charts; WHO–2005 standard growth charts.

2020 AMS Subject Classification: 62F10, 62G05, 62G08, 62P10.

1. INTRODUCTION

An underweight status of toddler is a nutritional status of toddler which is characterized by a lack of weight based on anthropometric index weight-for-age (W/A). In terms of malnutrition, Indonesia ranked fifth in the world that is about 3.8 per cent of 87 millions of total Indonesian toddlers. Nationally, the second position after East Nusa Tenggara province for cases of malnutrition in toddlers aged under five years old is East Java province. According to Health Department of East Java province, Indonesia there had been an increase in cases of malnutrition by 31.36 per cent that is from 4.716 cases to 6.195 cases in 2017 [1].

A Card Towards Health (KMS) is an instrument containing normal growth curves for toddlers based on the anthropometric index weight-for-age (W/A). As we know that up to now, Indonesia uses KMS based on WHO–2005 anthropometric standards. The use of KMS in Indonesia is based on the Z–Score standard growth charts of weight-for-age. Designing WHO–2005 standard growth charts used samples of toddlers aged 0–60 months who took from Brazil, Ghana, India, Norway, Oman, and USA. These samples are considered to represent regions of the world that are recommended as an assessment of global nutritional status [2]. However, there are different physical characteristics between Indonesian toddlers and those toddlers who used by WHO–2005. This, of course, makes a difference in chart patterns between the WHO–2005 standard growth charts and Indonesian standard growth charts that we propose in this study. Therefore, an effort that can be done to overcome the discrepancy is to design a KMS chart locally using data on toddlers aged under five years old whose physical condition is in accordance with Indonesian
Multiresponse semiparametric regression model approach

toddlers. Additionally, the growth charts for toddlers every age show different patterns at each stage [3]. In usual, the pattern does not form a linear curve or a particular shape, so that the appropriate model approach for this case is nonparametric regression model [4]. However, in several cases, some parts of toddlers growth charts form a linear pattern while other parts form nonlinear pattern, and there is correlation between responses, then for these cases, a multiresponse semiparametric regression (MSR) model approach is very suitable to use for designing the standard growth charts of toddlers [5]. Furthermore, to estimate these nonparametric regression models and semiparametric regression models, we use smoothing techniques. There are several smoothing techniques in nonparametric regression and semiparametric regression, for examples local linear [6–11], local polynomial [12–17], spline [4,5,18–34], and kernel [24,25,30,35–38], and two smoothing techniques of which are the LS-spline and local linear with the advantage of being able to overcome data patterns that show a sharp rise or fall with the help of knots, the resulting curve is relatively smooth [29,34], and able to determine the localness properties of data [6,7,10].

A study on designing standard growth charts of toddlers locally has been carried out by [39] in Padang City, but the underweight samples in this study were not differentiated by sex. Further, researches using local linear estimators have been done by [8–10] that have differentiated the sex of toddlers and found that the results of the East Java toddlers growth chart design curves lie in lower position than the WHO–2005 standard. In common, the standard growth charts of toddlers are designed based on percentile values. The advantage of using the percentile is that the calculation results are more accurate and able to be compared for each age group and anthropometric index. Calculating the percentile can indicate the problem of malnutrition and stunting which is more suitable compared to conventional systems [40].

For creating a standard growth chart of weight-for-age (W/A) for baby under five years old by using the nonparametric regression model approach, there are two variables used, namely toddler weight (kg) and toddler age (month) for each sex. While, if use gender as a variable in constructing the model then the appropriate model approach is semiparametric regression model. Semiparametric regression model is formed if in a regression model there are components of the
model in which some components can be estimated parametrically and remain components are estimated non-parametrically. The use of the semiparametric regression models based on the LS-spline estimator were also carried out by [29,41] that were to design the standard growth charts for East Java toddlers aged under five years old as a determinant of wasting nutrition status. The growth of toddlers that are different for each age will be appropriate if they are approached by semiparametric regression model where the toddlers gender variable is a parametric component which is as a dummy variable. Based on these facts, in this study we proposed an approach, namely multiresponse semiparametric regression (MSR) model, to design Percentile Standard Growth Charts (Percentile SGC) of weight-for-age (W/A) for East Java toddlers based on both LS-spline and local linear estimators. Next, the obtained these Percentile SGC designs will be compared with WHO–2005 Standard Growth Charts (WHO–2005 SGC) which can then be used as a reference for determining underweight or malnutrition status of East Java toddlers aged under five years old which is one of the indicators of stunting.

2. MATERIALS AND METHODS

In this section, we provide a brief overview of materials and methods used in this study such as multiresponse semiparametric regression model, curve of percentile, LS-spline and local linear estimators, generalized cross validation and coefficient of determination, and description of data.

2.1. Multiresponse Semiparametric Regression Model

In a regression model, if there is a combination of parametric regression and nonparametric regression then it is called semiparametric regression model [29]. Furthermore, if this semiparametric regression model draws a functional relationship between two or more response variables and two or more predictor variables where correlations between responses are occurred then this model is called Multiresponse Semiparametric Regression (MSR) model. Next, suppose a paired data \((x_{ri}, t_{ri}, y_{ri})\) has a functional relationship pattern which meets the MSR model as follows:

\[
y_{ri} = x_{ri}^T \beta_r + f_r(t_{ri}) + \varepsilon_{ri} , \quad i = 1,2,...,n ; r = 1,2,...,R
\]

where \(y_{ri}\) is the \(i^{th}\)-observation value of the \(r^{th}\)-response which has a parametric relationship
MULTIRESPONSE SEMIPARAMETRIC REGRESSION MODEL APPROACH

with the predictor variable $x_{ri}^T$ and a nonparametric relationship with predictor variables $t_{ri}$; $\beta_r$ is vector of unknown parameters of the $r^{th}$-response; $f_r$ is unknown regression function of the $r^{th}$-response; and $\varepsilon_{ri}$ is a zero mean independent random error with variance $\sigma^2$ of the $r^{th}$-response [5,32].

Hereinafter, based on the model presented by Eq.(1), it is easy to show that for every response $r = 1, 2, ..., R$ the following equation meets a Uniresponse semiparametric regression (USR) model [22,29,34]:

$$y_i = x_i^T \beta + f(t_i) + \varepsilon_i, \ i = 1, 2, ..., n.$$  

Next, in the following section we provide a brief overview of Percentile curve that will be used to design standard growth charts of East Java toddlers.

2.2. Curve of Percentile

For calculating values of standard deviation ($SD_i(t)$) on the Percentile curve of an observation, it involves mathematical calculations on normally distributed data from measurements that describe the population. The formula used is as follows [2]:

$$SD_i(t) = M(t)(1 + L(t) \times S(t) \times i)^{1/L(t)}, \ i = -3, -2, -1, 0, 1, 2, 3$$

where $t = 0, 1, 2, ..., 60$.

According to Cole [40], these L, M, and S values in Eq.(1) can be calculated through the following steps: (a). calculating the mean and standard deviation values of a natural logarithm of data, namely geometric mean ($M_g$) and geometric CV ($S_g$), respectively; (b). calculating the mean and standard deviation values of data, namely arithmetic mean ($M_a$) and arithmetic CV ($S_a$); (c). calculating mean and standard deviation values of inverse data, namely harmonic mean ($M_h$) and harmonic CV ($S_h$), respectively; (d). substituting $S_a$, $S_g$, and $S_h$ values into $A = \log(S_a/S_h)$ and $B = \log(S_a/S_g)$; (e). calculating estimation value of Box-Cox power ($L$), i.e., $L = -\frac{A}{2B}$; (f). calculating generalized coefficient of variance (S), namely $S = S_g \exp\left(\frac{AL}{4}\right)$; (g). calculating value of generalized mean (M), namely $M = M_g + \frac{(M_a-M_h)L}{2} + \frac{(M_a-2M_g+M_h)\cdot 2}{2}$. Next, in the following section we provide a brief overview of LS-spline and local linear estimators that will be used to
CHAMIDAH, ZAMAN, MUNIROH, LESTARI

design standard growth charts of East Java toddlers.

2.3. LS-Spline and Local Linear Estimators

If regression function \( f(t_i) \) in Eq.(2) is approximated by using the LS-spline estimator where \( p \) is the order of spline and \( m \) is the number of knots, then we can express the model in Eq.(2) as follows [34]:

\[
y_i = \alpha_0 + \gamma_1 x_i + \sum_{j=1}^{p} \beta_j t_i^j + \sum_{k=1}^{m} \beta_{(p+k)}(t_i - \varphi_k)^p + \varepsilon_i , \quad i = 1, 2, ..., n
\]

where \( \alpha_0 = \gamma_0 + \beta_0 \) and \( (t_i - \varphi_k)^p \) meets the following equation:

\[
(t_i - \varphi_k)^p = \begin{cases} (t_i - \varphi_k)^p, & t_i \geq \varphi_k \\ 0, & t_i < \varphi_k \end{cases}
\]

Hence, the semiparametric regression model in Eq.(4) based on the LS-spline estimator can be presented into a matrix notation as follows:

\[
y = X\gamma + T\beta + \varepsilon
\]

where \( X = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}^T \), \( \beta = (\beta_1, \ldots, \beta_p, \beta_{p+1}, \ldots, \beta_{p+m})^T \), \( \gamma = (\alpha_0, \gamma_1)^T \), \( \varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T \) and \( T = \begin{pmatrix} t_1 & \cdots & t_1^p & (t_1 - \varphi_1)^p & \cdots & (t_1 - \varphi_m)^p \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ t_n & \cdots & t_n^p & (t_n - \varphi_1)^p & \cdots & (t_n - \varphi_m)^p \end{pmatrix} \).

Therefore, semiparametric regression model presented in (5) can be stated into:

\[
y = C\omega + \varepsilon
\]

where \( C = (X \quad T) \) and \( \omega = (\alpha_0 \quad \gamma_1 \quad \beta)^T \).

Next, estimation of \( \omega \) in model (6) can be obtained by solving the following optimization:

\[
\text{Min} \{ Q = \sum_{i=1}^{n} \varepsilon_i^2 \} = \text{Min} \{ \varepsilon^T \varepsilon \} = \text{Min} \{ (y - C\omega)^T(y - C\omega) \}.
\]

Estimation of \( \omega \), namely \( \hat{\omega} \), is obtained by differentiating \( Q \) with respect to parameter \( \omega \) that is \( \frac{\partial Q}{\partial \omega} = 0 \). It would give the solution to optimization in (7) that is \( \hat{\omega} = (C^T C)^{-1} C^T y \). In this step, \( \hat{\omega} \) consists of estimation values of constant \( \alpha_0 \), parameter \( \gamma_1 \) of parametric component and parameter \( \beta \) of nonparametric component, so that we obtain the estimation of response variable
MULTI RESPONSE SEMIPARAMETRIC REGRESSION MODEL APPROACH

of semiparametric regression model as follows [5,32,34]:

\[ \hat{y} = C(C^T C)^{-1} C^T y = A y \]

where \( A = C(C^T C)^{-1} C^T \).

Furthermore, if the regression function \( f(t) \) in model (2) is estimated by using local linear estimator, then it can be expressed as follows [7]:

\[ \hat{f}(t) = t(t_0) \hat{\beta}(t_0) \]

where \( \hat{\beta}(t_0) \) is estimator for \( \beta(t_0) \) that obtained by minimizing the weighted least square function as follows:

\[ Q(t_0) = (y - Z(t_0) \beta(t_0))^T K_h(t_0) (y - Z(t_0) \beta(t_0)) \]

and we obtain:

\[ \hat{\beta}(t_0) = (Z^T(t_0) K_h(t_0) Z(t_0))^{-1} Z^T(t_0) K_h(t_0) y. \]

Hence, by considering Eq.(9) and Eq.(11), we obtain the estimation of regression function based on local linear estimator as follows [7]:

\[ \hat{f}(t) = t(t_0) (Z^T(t_0) K_h(t_0) Z(t_0))^{-1} Z^T(t_0) K_h(t_0) y. \]

Next, in the following section we provide a brief overview of generalized cross validation and coefficient of determination that will be used to design standard growth charts of East Java toddlers.

2.4. Generalized Cross Validation and Coefficient of Determination

In LS-spline regression it is very important to calculate the optimum knot point. Based on optimum knots, the best spline function is obtained. There are several methods for calculating the optimum knot point, one of which is the GCV (Generalized Cross Validation) method. The GCV for the MSR model is defined as follows [26,27,30,31]:

\[ GCV(\lambda) = \frac{\text{MSE}(\lambda)}{\text{tr}(I-A)^2} = \frac{\frac{1}{n} y^T (I-A)y}{\text{tr}(I-A)^2} \]

where \( \lambda = (p, \varphi_1, \varphi_2, ..., \varphi_m), \ A = C(C^T C)^{-1} C^T \), \( p \) is order of spline and \( m \) is the number of knots. Optimal knot values are determined from a minimum GCV value from a combination of observation points which are assumed to begin to change behavior patterns.

Next, the coefficient of determination notated \( R^2 \) is a measure of the accuracy of the regression curve [41,42]. The purpose of calculating the \( R^2 \) value is to determine the variation of the response.
variable \((y)\) that can be explained by the predictor variable \((x)\) together. The coefficient of determination can be calculated by the following formula:

\[
R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{\sum_{i=1}^{n}(y_i - \bar{y})^2}
\]

where \(SSE = \sum_{i=1}^{n}(y_i - \hat{f}(t_i))^2\) and \(0 \leq R^2 \leq 1\).

Hereinafter, in the next section we give a brief presentation on description of data that will be used to determine standard growth charts design which is suitable for determining the nutritional status of East Java toddlers.

2.5. Description of Data

Here we use a secondary data that contains observations on gender and toddlers’ body weight of toddlers aged 0–60 months which have been recorded from twenty three districts and cities in East Java province of Indonesia. Data collection was carried out in 2017-2018 at the Public Health Centre that is in Indonesia called as PUSKESMAS and Integrated Service Centre that is in Indonesia called as POSYANDU using nonprobability sampling with purposive sampling technique because the sampling is determined by the researcher. The samples used to design standard growth charts are toddlers who have passed the screening process according to WHO-2005 standards based on criteria for the condition of mothers and toddlers, environmental conditions, and economic conditions. The data obtained was in the form of cross-sectional data totaling 21,081 observations consisting of 10,859 observations for boys under five years of age and 10,222 observations for girls under five years of age.

3. RESULTS AND DISCUSSION

In this section we provide results and discussions on estimation of Percentile values based on both LS-spline estimator and local linear estimator.

3.1. Estimating Percentile Values Using LS-Spline Estimator

The best estimation model for creating a standard growth chart based on the anthropometric index weight-for-age \((W/A)\) is obtained by estimating the percentile values at each age based on the LS-spline estimator of the MSR model approach. In the MSR model approach, the optimal order of
splines and the optimal knots for each percentile value, namely 3-rd percentile (\(P_3\)); 15-th percentile (\(P_{15}\)); 50-th percentile (\(P_{50}\)); 85-th percentile (\(P_{85}\)); and 97-th percentile (\(P_{97}\)), are estimated based on GCV criterion namely the minimum GCV value. The estimation results which include order of spline, the optimal knots, minimum GCV values, MSE values, and coefficient of determination (\(R^2\)) for each percentile are presented in Table 1.

**Table 1. The Estimation Results Using LS-Spline Estimator for Each Percentile.**

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Order</th>
<th>Optimal Knots</th>
<th>Minimum GCV</th>
<th>MSE</th>
<th>(R^2) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_3)</td>
<td>2</td>
<td>6 ; 12</td>
<td>0.0893</td>
<td>0.0752</td>
<td>98.098</td>
</tr>
<tr>
<td>(P_{15})</td>
<td>2</td>
<td>6 ; 12</td>
<td>0.0451</td>
<td>0.038</td>
<td>99.274</td>
</tr>
<tr>
<td>(P_{50})</td>
<td>2</td>
<td>6 ; 12 ; 18</td>
<td>0.02027</td>
<td>0.0171</td>
<td>99.807</td>
</tr>
<tr>
<td>(P_{85})</td>
<td>2</td>
<td>6 ; 24</td>
<td>0.1968</td>
<td>0.1659</td>
<td>98.837</td>
</tr>
<tr>
<td>(P_{97})</td>
<td>2</td>
<td>6</td>
<td>0.4393</td>
<td>0.3836</td>
<td>97.837</td>
</tr>
</tbody>
</table>

Table 1 shows that the SGC W/A percentile for toddlers in East Java using the LS-spline estimator of the MSR model can explain the pattern of growth in body weight for toddlers in East Java well. Further, this is also supported by the mean value of the coefficient of determination (\(R^2\)) of 98.78 percent which is close to 100 percent and the mean value of the mean square error (MSE) of 0.1359 which is close to zero. This R2 value indicates that the MSR model with the LS-spline estimator used to design the SGC W/A percentile for East Java toddlers can explain the average variation in weight variables in East Java of 98.78 percent.

Next, by using Eq. (1) and Eq. (4) we obtain the estimation result of MSR model with the LS-spline estimator used on the median or 50-th percentile (\(P_{50}\)) of the first response, namely weight-for-age (W/A), for boy toddlers as follows:

\[
\hat{y}_{(1)\text{Boy}} = 8.021 + 0.691x + 0.537t - 0.012t^2 + 0.008(t - 6)^2 + 0.008(t - 12)^2 + 0.008(t - 18)^2
\]

where subscript “(1)” represents the first response, namely weight-for-age (W/A). Hence, we can express the Eq.(15) in the form of truncated function as follows:

\[
\hat{y}_{(1)\text{Boy}} = \begin{cases} 
8.021 + 0.691x + 0.537t - 0.012t^2, & 0 \leq t < 6 \\
8.312 + 0.691x + 0.441t - 0.004t^2, & 6 \leq t < 12 \\
9.461 + 0.691x + 0.249t + 0.004t^2, & 12 \leq t < 18 \\
12.053 + 0.691x - 0.039t + 0.012t^2, & 18 \leq t \leq 24 
\end{cases}
\]
Based on truncated function presented in Eq.(16), we obtained the highest average baby weight gain in East Java which occurred at the age interval \( 0 \leq t < 6 \). This means that, every one month the average weight gain of the baby increases by 0.477 kg while the lowest average weight gain occurs in the interval \( 6 \leq t < 12 \), and every increase of one month the average weight gain is 0.373 kg.

Similarly, by using Eq. (1) and Eq. (4) we obtain the estimation result of MSR model with the LS-spline estimator used on the median or 50-th percentile \((P_{50})\) of the second response, namely height-for-age (H/A), for boy toddlers as follows:

\[
\hat{y}_{(2)\text{Boy}} = 54.997 + 0.691x + 2.110t - 0.856(t - 6)_+ - 0.571(t - 12)_+ + 0.026(t - 18)_+ 
\]

where subscript “(2)” represents the second response, namely height-for-age (H/A). Hence, we can express the Eq.(17) in the form of truncated function as follows:

\[
\hat{y}_{(2)\text{Boy}} = \begin{cases} 
54.997 + 0.691x + 2.110t, & 0 \leq t < 6 \\
60.133 + 0.691x + 1.254t, & 6 \leq t < 12 \\
66.985 + 9.691x + 0.683t, & 12 \leq t < 18 \\
66.517 + 0.691x + 0.709t, & 18 \leq t \leq 24 
\end{cases} 
\]

Based on truncated function presented in Eq.(18), we obtained the highest average baby height gain in East Java which occurs in the age interval \( 0 \leq t < 6 \). That is, every one month the average baby's height increases by 2.11 cm, while the lowest average height gain occurs at the interval \( 18 \leq t \leq 24 \), and every one month increase, the average height gain is 0.709 cm. Furthermore, based on the estimation results on each percentile value \( P_3; P_{15}; P_{50}; P_{85}; \) and \( P_{97} \) and by substituting \( x = 1 \) (for boy toddlers) into Eq. (16) and Eq. (18), we obtained the Percentiles SGC of weight-for-age (W/A) and height-for-age (H/A) for East Java boy toddlers using the MSR model LS-spline estimator as shown in Figure 1.

Next, the estimation result of MSR model with the LS-spline estimator used on the median or 50-th percentile \((P_{50})\) of the first response, namely weight-for-age (W/A), for girl toddlers is as follows:

\[
\hat{y}_{(1)\text{Girl}} = 3.860 + 0.691x + 0.537t - 0.012t^2 + 0.008(t - 6)_+ 
\]
MULTIRESPONSE SEMIPARAMETRIC REGRESSION MODEL APPROACH

\[ 0.008(t - 12)^2 + 0.008(t - 18)^2 \]

where subscript “(1)” represents the first response, namely weight-for-age (W/A).

\( \hat{y}_{(1)}^{\text{Girl}} = \begin{cases} 
3.860 + 0.691x + 0.537t - 0.012t^2, & 0 \leq t < 6 \\
4.148 + 0.691x + 0.441t - 0.004t^2, & 6 \leq t < 12 \\
5.300 + 0.691x + 0.249t + 0.004t^2, & 12 \leq t < 18 \\
7.892 + 0.691x - 0.039t + 0.012t^2, & 18 \leq t \leq 24 
\end{cases} \)

Based on truncated function presented in Eq.(20), we obtained the highest mean baby weight gain in East Java which occurred at the age interval \( 0 \leq t < 6 \). That is, every one month the mean baby's weight gain increases by 0.477 kg, while the lowest mean weight gain occurs at the interval \( 6 \leq t < 12 \), and every increase of one month the mean weight gain is 0.373 kg.

Similarly, by using Eq. (1) and Eq. (4) we obtain the estimation result of MSR model with the LS-spline estimator used on the median or 50-th percentile \( (P_{50}) \) of the second response, namely height-for-age (H/A), for girl toddlers as follows:

\( \hat{y}_{(2)}^{\text{Girl}} = 53.370 + 0.691x + 2.110t - 0.856(t - 6)_+ - 0.571(t - 12)_+ + 0.026(t - 18)_+ \)
Next, we can express Eq.(21) into a truncated function as follows:

\[
\hat{y}_{(2)\text{Girl}} = \begin{cases} 
53.370 + 0.691x + 2.110t, & 0 \leq t < 6 \\
58.506 + 0.691x + 1.254t, & 6 \leq t < 12 \\
64.358 + 0.691x + 0.683t, & 12 \leq t < 18 \\
64.890 + 0.691x + 0.709t, & 18 \leq t \leq 24 
\end{cases}
\]

Based on truncated function presented in Eq.(22), we obtained the highest mean baby height gain in East Java which occurs in the age interval 0 \leq t < 6. This means that, every one month the mean baby's height increases by 2.11 cm while the lowest mean height gain occurs at the interval 18 \leq t \leq 24, and every one month increase, the mean height gain is 0.709 cm.

Furthermore, based on the estimation results on each percentile value of \( P_3; P_{15}; P_{50}; P_{85} \); and \( P_{97} \) and by substituting \( x = 0 \) (for girl toddlers) into Eq. (20) and Eq. (22), we obtained the Percentiles SGC of weight-for-age (W/A) and height-for-age (H/A) for East Java girl toddlers by using the MSR model LS-spline estimator as shown in Figure 2.
yellow area, whereas a light green area indicates a toddler who has a nutritional status with a threshold of more than $P_{85}$. Malnutrition status lies in the area below the red line with a Percentile threshold value of less than $P_3$. Next, comparison between Percentile SGC and WHO–2005 SGC of W/A boys and girls in East Java using LS-spline estimator are presented in Figure 3.

![Figure 3](image-url)  
**Figure 3.** Comparison of Percentile SGC and WHO–2005 SGC of W/A Boys (left) and W/A Girls (right) in East Java Using LS-Spline Estimator.

Based on results presented by Figure 3, in general the Percentile SGC of W/A for East Java toddlers lies under the WHO–2005 SGC. In average, the differences between the WHO–2005 SGC of W/A and the Percentile SGC of W/A for toddlers in East Java for boy toddlers and for girl toddlers are 1.2813 kg and 1.2059 kg, respectively.

The results of the comparison of the categorization (in per cent) of the nutritional status of under-fives based on the anthropometric index W/A between the SGC W/A under-fives and SGC WHO-2005 percentiles for each sex are presented in Table 2.

<table>
<thead>
<tr>
<th>Index of Anthropometric</th>
<th>Nutritional Status</th>
<th>Boys</th>
<th></th>
<th>Girls</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>W/A</td>
<td>Malnutrition</td>
<td>0.6209 %</td>
<td>3.6182 %</td>
<td>0.4781 %</td>
<td>3.0178 %</td>
</tr>
<tr>
<td></td>
<td>Underweight</td>
<td>2.9305 %</td>
<td>15.7873 %</td>
<td>3.1836 %</td>
<td>13.7288 %</td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>92.5065 %</td>
<td>78.9155 %</td>
<td>91.9412 %</td>
<td>81.5647 %</td>
</tr>
<tr>
<td></td>
<td>Overweight</td>
<td>3.9401 %</td>
<td>1.6805 %</td>
<td>4.3981 %</td>
<td>1.6879 %</td>
</tr>
</tbody>
</table>

Table 2 shows that the percentage of under-nutrition and malnutrition status based on the SGC of the weight-for-age percentile (W/A) of East Java toddlers is less than that based on the WHO-2005
SGC. Between Percentile SGC of the weight-for-age percentile (W/A) and WHO-2005 SGC there is the difference in the percentage of malnutrition status of 2.9972 per cent for boys and 2.5402 per cent for girls. While between Percentile SGC of the weight-for-age percentile (W/A) and WHO-2005 SGC there is the difference in the percentage of the underweight status of 12.8561 per cent for boys and 10.5463 per cent for girls. East Java toddlers who should have normal nutritional status based on the SGC Percentile will be categorized as having undernourished or severe nutritional status by the WHO-2005 SGC.

Based on results of the estimation of MSR model with LS-spline estimator used that have been discussed in this section, we obtain nutritional status of boy and girl toddlers based on W/A and H/A East Java SGC (Percentile SGC) compared to WHO-2005 SGC as presented in Table 3 and Table 4, respectively.

### Table 3. Nutritional Status for Toddlers Based on W/A of Percentile SGC and WHO-2005 SGC.

<table>
<thead>
<tr>
<th>Nutritional Status Based on W/A</th>
<th>Boy Toddlers</th>
<th>Girl Toddlers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentile SGC</td>
<td>WHO-2005 SGC</td>
</tr>
<tr>
<td>Very Underweight</td>
<td>3.04 %</td>
<td>13.11 %</td>
</tr>
<tr>
<td>Underweight</td>
<td>12.02 %</td>
<td>25.42 %</td>
</tr>
<tr>
<td>Normal</td>
<td>69.41 %</td>
<td>51.17 %</td>
</tr>
<tr>
<td>Overweight</td>
<td>15.53 %</td>
<td>10.3%</td>
</tr>
</tbody>
</table>

### Table 4. Nutritional Status for Toddlers Based on H/A of Percentile SGC and WHO-2005 SGC.

<table>
<thead>
<tr>
<th>Nutritional Status Based on H/A</th>
<th>Boy Toddlers</th>
<th>Girl Toddlers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentile SGC</td>
<td>WHO-2005 SGC</td>
</tr>
<tr>
<td>Very short</td>
<td>3.47 %</td>
<td>27.89%</td>
</tr>
<tr>
<td>Short</td>
<td>12.01%</td>
<td>27.65%</td>
</tr>
<tr>
<td>Normal</td>
<td>69.22%</td>
<td>37.08%</td>
</tr>
<tr>
<td>Tall</td>
<td>15.3%</td>
<td>7.38 %</td>
</tr>
</tbody>
</table>

#### 3.2. Estimating Percentile Values Using Local Linear Estimator

For designing Percentile SGC (also called local linear SGC) of W/A, the values of percentile curves, namely percentile values of 3-rd percentile ($P_3$), 15-th percentile ($P_{15}$); 50-th percentile...
MULTIRESPONSE SEMIPARAMETRIC REGRESSION MODEL APPROACH

($P_{50}$), 85-th percentile ($P_{85}$), and 97-th percentile ($P_{97}$) of toddlers weight in every age group should be estimated by using local linear estimator. Hence, the estimation of these percentile curves can be obtained by finding the optimal bandwidth ($h$) that meets cross-validation (CV) criterion which is the minimum value of CV. In this step we get the optimal bandwidth ($h$) values for every percentile which is shown in Table 5.

**Table 5.** Values of Optimal Bandwidth, Minimum CV, MSE and $R^2$ for Every Percentile.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Percentile</th>
<th>Optimal Bandwidth</th>
<th>Minimum CV</th>
<th>MSE</th>
<th>$R^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy</td>
<td>$P_3$</td>
<td>1.26</td>
<td>0.0629</td>
<td>0.0264</td>
<td>99.55</td>
</tr>
<tr>
<td></td>
<td>$P_{15}$</td>
<td>0.86</td>
<td>0.0133</td>
<td>0.0038</td>
<td>99.95</td>
</tr>
<tr>
<td></td>
<td>$P_{50}$</td>
<td>1.81</td>
<td>0.0447</td>
<td>0.0185</td>
<td>99.79</td>
</tr>
<tr>
<td></td>
<td>$P_{85}$</td>
<td>1.50</td>
<td>0.0628</td>
<td>0.0288</td>
<td>99.77</td>
</tr>
<tr>
<td></td>
<td>$P_{97}$</td>
<td>1.83</td>
<td>0.1400</td>
<td>0.0707</td>
<td>99.55</td>
</tr>
<tr>
<td></td>
<td>$P_3$</td>
<td>2.44</td>
<td>0.0586</td>
<td>0.0326</td>
<td>99.48</td>
</tr>
<tr>
<td></td>
<td>$P_{15}$</td>
<td>0.23</td>
<td>0.0134</td>
<td>3.3196e-10</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>$P_{50}$</td>
<td>1.13</td>
<td>0.0213</td>
<td>0.0081</td>
<td>99.91</td>
</tr>
<tr>
<td></td>
<td>$P_{85}$</td>
<td>1.62</td>
<td>0.0699</td>
<td>0.0345</td>
<td>99.72</td>
</tr>
<tr>
<td></td>
<td>$P_{97}$</td>
<td>1.32</td>
<td>0.1644</td>
<td>0.0777</td>
<td>99.55</td>
</tr>
</tbody>
</table>

From the results presented in Table 5, we obtained the average $R^2$ values from the estimated weight gains for boys and girls in each percentile of 99.723 percent and 99.728 percent, respectively. Also the average MSE score is 0.027577 for boys and 0.029759 for girls. Estimation of MSR model for median ($P_{50}$) growth charts of weight at age 0, 6, 12, 24, 36, and 48 months for boy toddlers is given as follows:

\[
\hat{y}_{\text{Boy}} = \begin{cases} 
3.875 + 0.691x + 0.9121(t - t_0), & t_0 = 0 \\
7.782 + 0.691x + 0.3532(t - t_0), & t_0 = 6 \\
9.332 + 0.691x + 0.2158(t - t_0), & t_0 = 12 \\
11.332 + 0.691x + 0.1657(t - t_0), & t_0 = 24 \\
13.165 + 0.691x + 0.1044(t - t_0), & t_0 = 36 \\
14.748 + 0.691x + 0.0969(t - t_0), & t_0 = 48 
\end{cases}
\]

where $t_0 - 1.81 < t < t_0 + 1.81$.

Next, estimation of MSR model for median ($P_{50}$) growth charts of weight at age 0, 6, 12, 24, 36, and 48 months for girl toddlers is given as follows:
\[
\hat{y}_{\text{Girl}} = \begin{cases} 
3.527 + 0.691x + 0.9724(t - t_0), & t_0 = 0 \\
7.344 + 0.691x + 0.3202(t - t_0), & t_0 = 6 \\
8.760 + 0.691x + 0.2246(t - t_0), & t_0 = 12 \\
10.845 + 0.691x + 0.1758(t - t_0), & t_0 = 24 \\
12.792 + 0.691x + 0.2347(t - t_0), & t_0 = 36 \\
14.320 + 0.691x + 0.0613(t - t_0), & t_0 = 48 
\end{cases}
\]

where \(t_0 - 1.13 < t < t_0 + 1.13\).

Based on the estimation results on each percentile value of \(P_3\); \(P_{15}\); \(P_{85}\); and \(P_{97}\) and by substituting \(x = 0\) (for girl toddlers) and \(x = 1\) (for boy toddlers) into Eq.(23) and Eq.(24), it can be shown that the highest body weight growth rate is achieved at around one year of age and then decreases with age. Next, assessment of nutritional status of toddlers under five years of age in East Java, Indonesia based on Percentile SGC (or local linear SGC) and WHO–2005 SGC applied to 18,9608 boy toddlers and 17,9624 girl toddlers is presented in Table 6.

**Table 6. Nutritional Status of Toddlers Based on Percentile SGC and WHO–2005 SGC.**

<table>
<thead>
<tr>
<th>Nutritional Status of Toddlers for weight-for-age (W/A)</th>
<th>Boy Toddlers</th>
<th>Girl Toddlers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentile SGC</td>
<td>WHO-2005 SGC</td>
</tr>
<tr>
<td>Severely Underweight (Percentile &lt; (P_3))</td>
<td>3.05 %</td>
<td>16.95 %</td>
</tr>
<tr>
<td>Underweight ((P_3 \leq \text{Percentile} &lt; P_{15}))</td>
<td>12.20 %</td>
<td>26.13 %</td>
</tr>
<tr>
<td>Good ((P_{15} \leq \text{Percentile} \leq P_{85}))</td>
<td>69.43 %</td>
<td>49.67 %</td>
</tr>
<tr>
<td>Overweight ((\text{Percentile} &gt; P_{85}))</td>
<td>15.33 %</td>
<td>7.25 %</td>
</tr>
</tbody>
</table>

Based on the results presented in Table 6, the percentage of undernourished children aged 5 years in East Java, Indonesia, for both boys and girls based on Percentiles SGC (or local linear SGC) is less than WHO–2005 SGC. In general, Percentile SGC (or local linear SGC) of body weight-for-age (W/A) for both boys and girls has a lower evaluator standard than WHO-2005 SGC. This condition occurs because the samples used to design the Percentile SGC are toddlers aged up to five years from East Java who physically have different characteristics from the samples used to design the WHO-2005 SGC, namely from Brazil, Ghana, India, Norway, Oman and United
States of America. In addition, some toddlers in East Java, especially those aged more than two years, have normal nutritional status based on the Percentile SGC but based on the WHO-2005 SGC, their nutritional status is underweight.

Furthermore, based on the estimation results on each percentile value of $P_3; P_{15}; P_{50}; P_{85};$ and $P_{97}$ and by substituting $x = 0$ (for girl toddlers) and $x = 1$ (for boy toddlers) into Eq.(23) and Eq.(24), we obtain the Percentile SGC of W/A for boy and girl toddlers in East Java province of Indonesia by using local linear estimator of MSR model as shown in Figure 4.

**Figure 4.** Percentile SGC of W/A for boy toddlers (left) and girl toddlers (right) in East Java.

Next, comparisons between Percentile SGC and WHO–2005 SGC of weight-for-age (W/A) for East Java toddlers (boys and girls) based on local linear estimator are presented in Figure 5.

**Figure 5.** Comparison of Percentile SGC and WHO–2005 SGC of W/A for Boys (left) and for Girls (right) in East Java Using Local Linear Estimator.

Furthermore, according to the estimation results of MSR model based on local linear estimator
that have been discussed in this section, we obtain nutritional status of boy and girl toddlers in East Java for height-for-age (H/A) based on Percentile SGC compared to WHO-2005 SGC as presented in Table 7.

**Table 7.** Nutritional Status Based on (H/A) for Boys and Girls in East Java.

<table>
<thead>
<tr>
<th>Nutritional Status Based on Height-for-Age (H/A)</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ ] Severely Stunted (Percentile &lt; ( P_3 ))</td>
<td>3.18 %</td>
<td>29.33%</td>
</tr>
<tr>
<td>[ ] Stunted (( P_3 \leq ) Percentile &lt; ( P_{15} ))</td>
<td>12.05 %</td>
<td>27.13 %</td>
</tr>
<tr>
<td>[ ] Normal (( P_{15} \leq ) Percentile ( \leq P_{85} ))</td>
<td>69.31 %</td>
<td>36.95 %</td>
</tr>
<tr>
<td>[ ] Tall (Percentile ( &gt;P_{85} ))</td>
<td>15.46 %</td>
<td>6.59 %</td>
</tr>
</tbody>
</table>

Next, based on the estimation results on each percentile value of \( P_3; P_{15}; P_{50}; P_{85}; \) and \( P_{97} \) and by substituting \( x = 0 \) (for girl toddlers) and \( x = 1 \) (for boy toddlers) into Eq.(23) and Eq.(24), we obtain the Percentile SGC of anthropometric index height-for-age (H/A) by using local linear estimator of MSR model for boys and girls in East Java as shown in Figure 6.

**Figure 6.** Percentile SGC H/A of toddlers boy (left) and toddlers girl (right) in East Java.
4. CONCLUSIONS

Estimation results based on the Percentile Standard Growth Charts (Percentile SGC) that uses LS-spline estimator and local linear estimator show that the Multiresponse Semiparametric Regression (MSR) model approach is a good fitting model because it satisfies the goodness of fit criteria namely the coefficient of determination ($R^2$) that tends to one and Mean Square Errors (MSE) that tends to zero. In addition, the Percentile SGC of anthropometric index weight-for-age (W/A) for toddlers in East Java yields a lower reference than the WHO–2005 SGC. In general, the Percentile SGC of anthropometric index W/A for boy toddlers and girl toddlers in East Java have the evaluator standard and have lower standard than the WHO–2005 Standard Growth Charts (WHO–2005 SGC).

ACKNOWLEDGMENTS

The authors thank the Directorate of Research and Public Service (Direktorat Riset dan Pengabdian kepada Masyarakat–DRPM), the Ministry of Research, Technology, and Higher Education, the Republic of Indonesia, for funding this research through the Featured Basic Research of Higher Education grant (Penelitian Dasar Unggulan Perguruan Tinggi–PDUPT) in the fiscal year 2018 with a contract number, No. 893/UN3/2018 dated 22 January 2018. Also, the authors thank Universitas Airlangga, Indonesia for technical support.

CONFLICT OF INTERESTS

The authors confirm that there is no conflict of interests.

REFERENCES


MULTI RESPONSE SEMIPARAMETRIC REGRESSION MODEL APPROACH


[22] B. Lestari, Semiparametric modeling of consumer price index, Majalah Ilmiah Matematika dan Statistika. 16 (2016), 79–90.


