Standard Growth Chart of Weight for Height to Determine Wasting Nutritional Status in East Java Based on Semiparametric Least Square Spline Estimator

W Ramadan¹, N Chamidah¹*, B Zaman¹, L Muniroh², and B Lestari³

¹Department of Mathematics, Faculty of Sciences and Technology, Airlangga University, Surabaya, Indonesia
²Department of Health Nutrition, Faculty of Public Health, Airlangga University, Surabaya, Indonesia
³Department of Mathematics, Faculty of Mathematics and Natural Sciences, The University of Jember, Indonesia

Corresponding author’s email: nur-c@fst.unair.ac.id

Abstract. Wasting is a condition of a children characterized by a lack of weight by measuring weight for height. Currently, to monitor the growth conditions for childrens in Indonesia, we use the Towards Healthy Card called as Kartu Menuju Sehat (KMS) which is guided by WHO 2005. The samples used to design WHO-2005 standard charts are children from Brazil, Ghana, India, USA, Norway, and Oman that have different physical conditions from children in Indonesia. Therefore, the using of standards growth charts from other countries cause incompatibility with Indonesian’s children growth. To illustrate the growth patterns of children in East Java, we use the semiparametric least square spline estimator that gives more flexible pattern. In this study we used weight (kg) as a response variable, height (cm) as a predictor variable for nonparametric component, and gender as a predictor variable for parametric component. The results show the semiparametric least square spline estimator can explain the growth patterns of children well because it has determination coefficient (R²) of 99.78% and mean square error (MSE) of 0.0353. The standard chart of weight for height of boy is higher than that of girl and percentage of wasting nutritional status of girl greater than that of boy.

1. Introduction

Indonesia is the fourth largest country in the world with a total of 260 million residents. Group the most dominant age is at the age of five with 23.85 million people or 9.17% of the population. Recently, Indonesia is launching one of the Indonesia Emas 2045 program that has vision is to create highly competitive human resources. Some things that can be done is to solve problems in the field children nutrition like wasting because in 2045 the age of children’s is currently at productive age. Wasting is a condition of a children characterized by lack body weight with measurements of weight for height. Prevalence of children’s national wasting in 2016 is 11.1% and it is classified as a serious problem because the prevalence is between 10-14% according to WHO 2010. Wasting conditions in East Java in 2017 is 6.9% and it is classified as acute nutritional problems because the prevalence is ≥ 5% according to WHO 1997 [1].

Nutrition problems in Indonesia still need special attention from the government. based on the latest data states that Indonesia is the 17th country with three nutritional problems in infants such as 37.2%
short, 12.1% wasting, and 11.9% overweight in infants [2]. The latest data from of East Java governent showed that the percentage of malnutrition in East Java was 10.3% in 2014 where this percentage had increased from 2013 where previously only 9.3% [3]. According to [4], children in the age range of two years grow rapidly, then decline slowly with increasing age of children. Monitoring the growth conditions for children’s in Indonesia is currently using the Towards Healthy Card (KMS) which is guided by WHO 2005. The fact is that the physical condition of children in Indonesia is different from the sample used by WHO to compile a standard chart so that it allows the KMS standard by WHO not to be in accordance with the condition of children under five in Indonesia. Previous research carried out by [5] in Europe explained that the results of their national growth measurements are considered more accurate. So, the estimating of children growth charts is more appropriate modeled by using semiparametric regression method through the least square spline estimator approach.

Regression analysis method can be used to determine the relationship between predictor variables and response variables. There are two approaches to estimating the regression function, namely parametric and nonparametric approaches. We can use the parametric regression analysis approach when the researcher assumes the form of the regression function is known. However, we cannot assume especially the form of the regression function is unknown, for this case, we can use a nonparametric regression analysis approach. Furthermore, if we combine these approaches, we will obtain an analysis method called as semiparametric regression analysis. There are many estimators that can be used to estimate regression functions in nonparametric and semiparametric regressions models, for example, spline estimators, wavelet estimators, local polynomial estimators, etc. Among all these estimators, spline estimators have the best flexibility for nonparametric estimates. and semiparametric regressions functions. The spline estimator for estimating nonparametric functions from bivariate data with the same error correlation was studied by [6]. Designing of children's growth chart for up to five years based on multi-response local polynomial model and smoothing of the kernel has been provided by [7]. Next, [8] have studied the P-Spline estimator in the bi-response semiparametric regression applied to design growth charts for boys’ and girls' references to weight and height in Surabaya, Indonesia. Homoscedastic multiresponse nonparametric regression model has been discussed by [9]. Also, [10] propose a bi-response local semiparametric linear estimator to estimate the growth rate chart of boys and girls up to two years old, in Surabaya which concludes the average growth chart for age and length for age for boys higher than girls. In addition, [11] and [12] discussed estimations of covariance matrix and regression function in the multiresponse nonparametric regression models, respectively. The local linear estimator has been used by [13] to design a standard growth chart for toddlers in region. Furthermore, [14] and [15] determined regression functions of the multiresponse and biresponse nonparametric regressions models, respectively, by using spline and kernel estimators. Finally, [16] using least square spline estimator to estimate median growth of children.

These previous researches have showed that the spline estimator approach is an appropriate approach for estimating the functions of the nonparametric and semiparametric regressions. Because it has a very good ability to overcome data problems that change at certain points. Because of its flexible nature and it can calculate knots and smoothing parameters simultaneously, therefore, in this paper, we use semiparametric least square spline estimator approach for describing the growth patterns of children and determining the wasting nutritional status in East Java.

2. Semiparametric Regression Model and Detail of Children Data
In this section we describe the semiparametric regression model based on truncated spline estimator and the data of children used in this analysis.

2.1. Semiparametric Regression Model
Suppose that paired data \( (y, x_1, t_1) \) follows the semiparametric regression model based on truncated spline estimator as follows:
\[ y_i = \alpha_{0i} + \alpha_{1i}x_{1i} + \cdots + \alpha_{pi}x_{pi} + \sum_{h=1}^{m} g(t_{ih}) + \varepsilon_i \]  

(1)

where \( y_i \) is \( i \)-th (\( i = 1, 2, \ldots, n \)) response variable; \( \sum_{j=0}^{p} \alpha_{ji}x_{ji} \) is parametric component that consists of parameter \( \alpha \) and variable \( x \); \( \sum_{h=1}^{m} g(t_{ih}) \) is nonparametric component that is sum of functions of variable \( t \); and \( \varepsilon_i \) is random error. We can write model (1) in matrix notation as follows:

\[ y = X'\alpha + T'\beta + \varepsilon \]

(2)

The parametric component of model (2) is \( X'\alpha \), where \( \alpha = (\alpha_{01}, \alpha_{11}, \ldots, \alpha_{pn})' \) and 

\[ X = \begin{bmatrix} 
1 & x_{11} & x_{12} & \cdots & x_{1p1} \\
1 & x_{12} & x_{22} & \cdots & x_{p2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_{1n} & x_{2n} & \cdots & x_{pm} 
\end{bmatrix} \]

The nonparametric component of model (2) is \( T'\beta \), where 

\[ \beta = (\beta_{01}, \beta_{11}, \beta_{21}, \ldots, \beta_{d1}, \beta_{(d+1)1}, \ldots, \beta_{(d+m)n})' \]  

and 

\[ T = \begin{bmatrix} 
1 & (t_{11})^2 & \cdots & (t_{11})^d & (t_{11}-K_1)^2 & \cdots & (t_{11}-K_1)^d \\
1 & (t_{12})^2 & \cdots & (t_{12})^d & (t_{12}-K_1)^2 & \cdots & (t_{12}-K_1)^d \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
1 & (t_{1n})^2 & \cdots & (t_{1n})^d & (t_{1n}-K_1)^2 & \cdots & (t_{1n}-K_1)^d \\
1 & (t_{2n})^2 & \cdots & (t_{2n})^d & (t_{2n}-K_1)^2 & \cdots & (t_{2n}-K_1)^d \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
1 & (t_{mn})^2 & \cdots & (t_{mn})^d & (t_{mn}-K_1)^2 & \cdots & (t_{mn}-K_1)^d 
\end{bmatrix} \]

The random error of model (2) is \( \varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)' \).

If \( C[X, T] \) represents the predictor variables of model (2) and \( \hat{\delta}[\alpha, \beta] \) represents the parameters of model (2), then model (2) can be expressed as follows:

\[ y = C'\hat{\delta} + \varepsilon \]

(3)

The estimated model of (3) can be obtained by minimizing the following least square:

\[ \text{Min} Q(\hat{\delta}) = \text{Min}[(y - C'\hat{\delta})(y - C'\hat{\delta})'] \]

(4)

The estimated parameter of \( \hat{\delta} \) can be obtained by taking partially derivation \( Q(\hat{\delta}) \) with respect to \( \hat{\delta} \) as follows:

\[ \frac{\partial Q(\hat{\delta})}{\partial \hat{\delta}} = 0 \]

(5)

such that we obtain the estimated parameter of \( \hat{\delta} \) as follows:
\[ \hat{\beta} = (C'C)^{-1}C'y \]

(6)

We use truncated spline approach then we need the smoothing parameter including knot and order. Smoothing parameter serves to control the smoothness of the regression curve. In this study we use generalized cross validation (GCV) criterion, i.e., by selecting the minimum value of the following GCV:

\[ GCV(K) = \frac{MSE(K)}{1 - \frac{\text{trace}(I - A(K))}{n}} \]

(7)

2.2. Detail of Children Data

The data used in this analysis to determine the wasting nutritional status is secondary data that consists of observation data of body weight (kg), height (cm), and sex of children from 2016 to 2018. There are 30,491 observation data for male children and 28,681 observation data for female children obtained from Integrated Service Centers (Posyandu) and Centers of Public Health (Puskesmas). The population came from 21 regencies and cities in East Java. The sampling technique for retrieving data used non-probability sampling with purposive sampling technique because of taking the sample was determined by the researcher.

3. Results and Discussion

The percentile value is used to get the best model estimation. The best model can be obtained by looking for the minimum GCV value, the smallest MSE, and the largest value of \( R^2 \) at each point value of knots.

The results of analysis on the data give the number of knots, the optimum knots, and minimum GCV values for percentiles as given in Table 1.

The estimated model of median (\( P_{50} \)) of weight for height growth for children is obtained as follows:

\[ \hat{y}_{P_{50}} = -12.27 + 0.18x + 0.32t - 0.13(t - 55)^2 + 0.04(t - 83)^2 + 0.07(t - 105) \]

(8)

The estimated of median of weight for height given in (8) can be described again in order to make it easier to understand it by using the form of the piecewise function. The form of the piecewise function can be stated as follows:

\[ \hat{y}_{P_{50}} = \begin{cases} 
-12.27 + 0.18x + 0.32t, & t < 55 \\
-5.12 + 0.18x + 0.19t, & 55 \leq t < 83 \\
-8.44 + 0.18x + 0.23t, & 83 \leq t < 105 \\
-15.79 + 0.18x + 0.3t, & t \geq 105 
\end{cases} \]

(9)

The estimated and observed plot of median growth charts for boy and girl are given in Figure 1 and Figure 2.
Table 1. The number of knots, the optimum knots, and minimum GCV values for percentiles

| Percentile | Number of Knot | Optimum Knot | Minimum GCV Value |
|------------|----------------|--------------|-------------------|
| 3          | 1              | 62           | 0.05045457        |
|            | 2              | 70, 75       | 0.04837877        |
|            | 3              | 67, 79, 97   | 0.04718184        |
| 15         | 1              | 52           | 0.0332645         |
|            | 2              | 64, 79       | 0.02900991        |
|            | 3              | 63, 81, 109  | 0.02733034        |
| 50         | 1              | 94           | 0.06634452        |
|            | 2              | 54, 89       | 0.05006562        |
|            | 3              | 55, 83, 105  | 0.04656288        |
| 85         | 1              | 96           | 0.1102755         |
|            | 2              | 51, 95       | 0.09447514        |
|            | 3              | 93, 109, 111 | 0.08565596        |
| 79         | 1              | 95           | 0.2869668         |
|            | 2              | 93, 103      | 0.2795403         |
|            | 3              | 71, 93, 103  | 0.2796111         |

Figure 1. Estimation median of boys in East Java.
Based on the WHO 2005 design standard which is also applied to the growth chart in Indonesia, the range of height for boy and girl is 45-120 cm, so it is necessary to calculate the estimated value in under-in-sample children height categories and the estimated chart can be obtained as follows:

Figure 2. Estimation median of girls in East Java.

Figure 3. Standard chart for boys in East Java.
Figure 4. Standard chart for girls in East Java.

The comparison results of children nutritional status based on East Java Standard Chart and WHO-2005 Standard Chart in the 50-th percentile by sex (boy and girl) as given in Table 2.

| Nutritional Status | Boy          |            | Girl          |            |
|--------------------|--------------|------------|--------------|------------|
|                    | East Java    | WHO        | East Java    | WHO        |
| Severe Wasting     | 2.92 %       | 11.72 %    | 3.21 %       | 9.22 %     |
| Wasting            | 11.22 %      | 15.73 %    | 10.84 %      | 16.00 %    |
| Normal             | 73.70 %      | 51.35 %    | 73.71 %      | 63.37 %    |
| Obese              | 3.00 %       | 9.06 %     | 2.99 %       | 8.21 %     |

4. Conclusion
The standard growth chart of weight for height by using the semiparametric least square spline estimator can explain the growth patterns of children in East Java well because it has $R^2$ value of 99.78%, and MSE of 0.03539224. The median growth chart of weight for height of boy is higher than that of girl and the percentage of wasting nutritional status of girl greater than that of boy.

Acknowledgement
This Research was funded by the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia through Fundamental Research Grant fiscal year 2019.
References

[1] Atmarita 2008 Kaji ulang status gizi anak 0-59 bulan (berat badan menurut umur) Gizi Indonesia 31(1) 23-24

[2] Global Nutrition Report 2014 Actions and accountability to accelerate the world's progress on nutrition A-Peer-Reviewed Publication Washington DC

[3] Public Health Service 2016 Profil Kesehatan Provinsi Jawa Timur Tahun 2015. Dinas Kesehatan Provinsi Jawa Timur Retrieved from http://www.depkes.go.id/resources/download/profil/PROFIL_KES_PROVINSI_2015/15_Jatim_2015.pdf.

[4] Roumeliotis P 2012 Children’s health & wellness: growth and development Retrieved from http://www.drpaul.com/growth/index.html.

[5] Cameron N and Halwey N L 2009 Should the UK use WHO growth charts? Loughborough: Paediatrics and Child Health Journal 20 (4) 151-156

[6] Wang Y, Guo W and Brown M B 2000 Spline smoothing for bivariate data with applications to association between hormones Statistica Sinica 10 377-397

[7] Chamidah N and Saifudin T 2013 Estimation of children growth curve based on kernel smoothing in multi-response nonparametric regression Applied Mathematical Sciences 7(37) 1839-1847

[8] Chamidah N and Eridani 2015 Designing of growth reference chart by using birespon semiparametric regression approach based on P-Spline Estimator International Journal of Applied Mathematics and Statistics 3 (3) 150-158

[9] Chamidah N and Lestari B 2016 Spline estimator in homoscedastic multi-response nonparametric regression model in case of unbalanced number of observations Far East Journal of Mathematical Sciences (FJMS) 100(9) 1433-1453

[10] Chamidah N and Rifada M 2016 Local linear estimator in bi-response semiparametric regression model for estimating median growth chart of children Far East Journal of Mathematical Sciences (FJMS) 99(8) 1233-1244

[11] Lestari B, Anggraeni D and Saifudin T 2017 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, October 21, 2017, Airlangga University, Surabaya, Indonesia 238-242

[12] Lestari B, Fatmawati and Budiantara I N 2017 Estimasi fungsi regresi nonparametrik multirespon menggunakan reproducing kernel Hilbert space berdasarkan estimator smoothing spline. Proceeding of National Seminar on Mathematics and Its Applications, October 21, 2017, Airlangga University, Surabaya, Indonesia 243-250

[13] Chamidah N, Tjahjono E, Fadilah A R and Lestari B 2018 Standard growth charts for weight of children in East Java using local linear estimator Journal of Physics: Conference Series 1097 012092.

[14] Lestari B, Fatmawati, Budiantara I N and Chamidah N 2018 Estimation of regression function in multi-response nonparametric regression model using smoothing spline and kernel estimators Journal of Physics: Conference Series 1097 012091

[15] Lestari B, Chamidah N and Saifudin T 2019 Estimasi fungsi regresi dalam model reresi nonparametrik birespon menggunakan estimator smoothing spline dan estimator kernel Jurnal Matematika, Statistika, dan Komputasi 15(2) 20-24

[16] Chamidah N, Kurniawan A, Zaman B and Muniroh L 2018 Least square spline estimator in multi-response semiparametric regression model for estimating median growth charts of children in East Java, Indonesia Far East Journal of Mathematical Sciences (FJMS) 107(2) 295-307