International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 13, Issue 1, 2020

Designing Local Standard Growth Charts of Children in East Java Province Using a Local Linear Estimator

Nur Chamidah[a*], Badrus Zaman[b], Lailatul Muniroh[c], Budi Lestari[d],
[a,b]Department of Mathematics, Faculty of Sci. and Tech., Airlangga University, [c]Department of Health Nutrition, Faculty of Public Health, Airlangga University Campus C-UNAIR, Mulyorejo Street, Surabaya 60115, Indonesia, [d]Mathematics Department, Fac. of Math. and Natural Sci., The University of Jember, Indonesia, Email: [a*]nur-c@fst.unair.ac.id

In this study we propose new designs of standard growth charts of children called a local standard growth chart (LSGC) which is significantly more suitable for children aged up to five years old in the East Java province of Indonesia than the WHO-2005 standard growth chart (WSGC), to assess their nutrition status. The proposed designs are LSGC of weight, LSGC of height and LSGC of BMI which are based on age by using a sample of children aged up to five years from East Java province. Children around one year old grow up rapidly then slowly. Also, growth patterns of boys and girls are different. Hence, we use predictor variables that consist of sex as a parametric component and age as a nonparametric component, leading to a semiparametric local linear model approach for constructing LSGC. Results show that the obtained median models of LSGC have satisfied the goodness of model fitting criterions, that include determining coefficient (R2) of 0.997 and mean squared error of 0.21. In addition, the percentage of normal status of nutrition for children is considered province based on LSGC which are higher than those based on WSGC. LSGC can be used by the Government for assessing the status of nutrition of children from East Java .

Key words: *Children aged up to five years, LSGC, Multi-response Local Linear Estimator, Nutritional Status.*

Introduction

Regression analysis is a method used to analyse the functional association between independent (predictor) variable and dependent (response) variable. There are two regression model approaches to estimate regression function which include parametric regression and nonparametric regression model. Parametric regression model is used when its regression function form is known, for examples, linear, quadratic, cubic, etc. Non-parametric regression model is used when its regression function form is unknown. Furthermore, by combining these models, we obtain a semiparametric regression model. There are many researchers who have used some estimators for approximating the function of regression in non-parametric and semiparametric regression models.

Nottingham & Cook (2011) used local linear regression method for estimating time series data. Chamidah et. al. (2018) have designed SGC for children in East Java by using local linear estimators. Aydin & Yilmaz (2018) have modified spline regression for right-censoring data randomly. Chamidah et. al. (2019a) have used local polynomial estimator to improve accuracy of classifications on tumour and cyst. Darnah et. al. (2019) used local linear estimator in Poisson regression to model cases of maternal and infant mortalities in East Kalimantan province. Ana et. al. (2019) used local a linear estimator in additive non-parametric logistic regression to model risk factors of hypertension. Murbarani et. al. (2019) predicted percentage of people in East Java province who suffer AIDS by applying a truncated spline estimator of non-parametric regression. Chamidah et. al. (2019b) estimated median SGC for the height of children in East Java by using a P-Spline estimator. Ramadan et. al. (2019) determined the wasting status of nutrition of children in East Java by using a semiparametric LS-Spline estimator. Puspitawati & Chamidah (2019) used local linear methods for classifying neovascularisation of choroid on images of retina fundus. Massaid et. al. (2019) modelled the percentage of poverty per capita expenditures of non-food in Indonesia by using an LS-Spline estimator. However, these researchers only discussed certain estimators in nonparametric and semiparametric regressions models for nonresponses.

In some actual cases, researchers frequently meet problems where more than one variables of responses are observed by several values of predictor variables. Multi-response nonparametric and semiparametric regression models provide potential tools to determine the functions that represent the relationship between each variable. Some researchers used certain approaches of these estimators of these models in their research. Wang et. al. (2000) used spline estimator of bivariate nonparametric regression model with the same errors correlations. Chamidah et. al. (2012) and Chamidah & Saifudin (2013) designed the growth chart of children aged up to five years old in Surabaya using local multi-response polynomial and kernel estimators. Chamidah & Eridani (2015) used P-Spline estimator of response semiparametric regression model to design SGC of boys and girls in Surabaya for both weight and height.

Chamidah & Rifada (2016a; 2016b) have studied the local linear estimator of response of the semiparametric model in estimating median SGC of boys and girls up to the age of two . Chamidah & Lestari (2016) have discussed estimating homoscedastic multi-response non-parametric regression model with equal errors variance using a spline estimator. Lestari et. al. (2018; 2019) used spline and kernel estimators for determining regression functions and optimal smoothing parameters in models of multi-response nonparametric regression. Lestari et. al. (2019a) gave a simulation study of smoothing spline estimator in estimating a multi-response nonparametric regression model. Lestari et. al. (2019b) discussed the theoretical use of smoothing a spline estimator to predict blood pressure and pulse. Lestari et. al. (2019c) discussed smoothing spline estimator in predicting blood pressure and heart rate. Chamidah et. al. (2019) used a spline estimator to predict heart rate and blood pressure affected by stress level. Chamidah & Lestari (2019) estimated a covariance matrix by using a local multi-response polynomial estimator. Hidayati et. al. (2019) used semiparametric truncated spline to model a computer based national exam. Nidhomuddin et. al. (2019) used a local linear model for modelling bivariate longitudinal data.

According to the Global Nutrition Report (2014), nutritional problems in Indonesia require further attention. Indonesia is the 17th country in the world that still has nutritional problems for infants including 37.2% of infants being short, 12.1% are wasting and 11.9% are overweight. The Health Department of East Java province, Indonesia pointed out that the percentage of malnutrition and nutrient lacking in children in East Java province in 2014 was 10.3% where it was greater than 2013 which was 9.3%. Indonesia uses a health card which is called *Kartu Menuju Sehat* (KMS) for assessing nutritional status based on weight for age. To measure children' growth, we use anthropometry measurements not only by weight but also height and body mass index (BMI). In this case, since there are significant correlations between weight, height and BMI, they are better modelled in multi-response than single response model (Chamidah et. al., 2018). In fact, there are different patterns of growth charts for boys and girls so that gender is involved in the model as a parametric component and age as a nonparametric component. Therefore, children's growth chart can be designed by using the semiparametric regression model approach. Although Chamidah et. al. (2018) have discussed children's growth charts by using spline estimators in multi-response models of semiparametric regression, they discussed the estimation of median (percentile 50) growth charts only. According to WHO-MGRS (2006), in order to assess the status of nutrition for children under five, we need to estimate 3rd, 15th, 50th, 85th, 97th percentiles of growth charts. Furthermore, besides the spline estimator, we can also use a local linear estimator, as it is effective in accomodating behavioural change of data at particular intervals. In other words, this estimator can overcome local data patterns . Therefore, we discuss estimating 3$^{rd.}$, 15th, 50th, 85th, 97th percentiles of growth charts for assessing the nutritional status of children in East Java province, Indonesia by applying a local linear estimator of multi-response semiparametric regression model.

Methods

Suppose $(y_{ij}^{(k)}, x_{ij}, t_{ij})$, $k = 1, 2, ..., r$; $i = 1, 2, ..., n$; and $j = 1, 2, \ldots, m_i$ are paired observations data set that follow a multi-response semiparametric regression model:

$$\mathbf{y}_{ij} = \mathbf{X}_{ij}\boldsymbol{\beta} + \mathbf{f}(\mathrm{t}_{ij}) + \boldsymbol{\varepsilon}_{ij} \qquad (1)$$

where $y_{ij}^{(k)}$ represents k^{th} response variable in i^{th} subject and in j^{th} observation, x_{ij} represents predictor variable of parametric component in i^{th} subject and in j^{th} observation, and t_{ij} represents predictor variable of nonparametric component in i^{th} subject and in j^{th} observation.

Here $\mathbf{y}_{ij} = \begin{pmatrix} y_{ij}^{(1)} & y_{ij}^{(2)} & \cdots & y_{ij}^{(r)} \end{pmatrix}^T$ is a ($r \times 1$)-vector; $\mathbf{X}_{ij} = diag(\mathbf{x}_{ij}, \mathbf{x}_{ij}, ..., \mathbf{x}_{ij})$ is a ($r \times 2r$)-matrix, where $\mathbf{x}_{ij} = \begin{pmatrix} 1 & x_{ij} \end{pmatrix}$, and $\boldsymbol{\beta} = \begin{pmatrix} \beta_0^{(1)} & \beta_1^{(1)} & \beta_0^{(2)} & \beta_1^{(2)} & \cdots & \beta_0^{(r)} & \beta_1^{(r)} \end{pmatrix}^T$ is a ($2r \times 1$)-vector of parameters of parametric component $\mathbf{X}_{ij}$. $\mathbf{f}(\mathrm{t}_{ij}) = \begin{pmatrix} f_1(t_{ij}) & f_2(t_{ij}) & \cdots & f_r(t_{ij}) \end{pmatrix}^T$ is a ($r \times 1$)-vector of regression functions of i^{th} subject and j^{th} observation that assumed to be smooth; $\boldsymbol{\varepsilon}_{ij} = \begin{pmatrix} \varepsilon_{ij}^{(1)} & \varepsilon_{ij}^{(2)} & \cdots & \varepsilon_{ij}^{(r)} \end{pmatrix}^T$ is a ($r \times 1$)-vector of random errors of i^{th} subject and j^{th} observation with mean $\mathbf{0}$ and variance $\boldsymbol{\Sigma}_i$.

Consequently, to obtain the estimated regression function based on local linear estimator, we firstly use WLS (weighted least squared) method, that is, by carrying out WLS optimisation. Secondly, to obtain the solution of WLS, that is, the value that minimises WLS, we can take the partial difference of WLS with respect to their parameters, after which they equal to zero. We use the data set of children's growth for ages 0 until 60 months obtained from the integrated health service centre recorded from 16 cities in East Java Province, Indonesia during 2018. We designed standard growth charts (SGC) of children based on local linear method in the multi-response semiparametric regression model which has three responses, including weight (in kg), height (in cm), and BMI (in kg/m^2), and a parametric predictor, that is gender, and a nonparametric predictor, which is age (in month). Generalised cross validation (GCV) criterion is used to determine optimal bandwidth (h). Finally, we created a code on open source software (OSS)-R to estimate children's growth charts.

Results and Discussion

In this section, we present the estimation of multi-response semiparametric regression model, data and steps of analysis, estimating LSGC for children and assessing children's nutritional status in the East Java province.

Estimating Multi-response Semiparametric Regression Model Using Local Linear Estimator

Firstly, we express the multi-response semiparametric regression model in (1) in the following matrix notation:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}(\mathbf{t}) + \boldsymbol{\varepsilon} \qquad (2)$$

where $\mathbf{y} = \begin{pmatrix} \mathbf{y}^{(1)} & \mathbf{y}^{(2)} & \mathbf{y}^{(3)} & \cdots & \mathbf{y}^{(r)} \end{pmatrix}^T$ is a (rM×1)-vector of responses; and $\boldsymbol{X} = diag(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, ..., \mathbf{X}^{(r)})$ is a (rM×2)-matrix of the parametric component predictor, where $\mathbf{X}^{(k)} = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & \cdots & 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1m_1} & x_{21} & x_{22} & \cdots & x_{2m_2} & \cdots & x_{n1} & x_{n2} & \cdots & x_{nm_n} \end{pmatrix}^T$ and $\boldsymbol{\beta} = \begin{pmatrix} \beta_0^{(1)} & \beta_1^{(1)} & \beta_0^{(2)} & \beta_1^{(2)} & \cdots & \beta_0^{(r)} & \beta_1^{(r)} \end{pmatrix}^T$ is a ($2r$×1)-vector of parameters of the parametric component; $\mathbf{f}(\mathbf{t}) = \begin{pmatrix} \mathbf{f}_1(\mathbf{t}) & \mathbf{f}_2(\mathbf{t}) & \mathbf{f}_3(\mathbf{t}) & \cdots & \mathbf{f}_r(\mathbf{t}) \end{pmatrix}^T$ is a (rM×1)-vector of the nonparametric component predictor, where $\mathbf{f}_k(\mathbf{t}) = (\mathbf{f}_k(t_{11}) \ \cdots \ \mathbf{f}_k(t_{1m_1}) \ \mathbf{f}_k(t_{21}) \ \cdots \ \mathbf{f}_k(t_{2m_2}) \ \cdots \ \mathbf{f}_k(t_{n1}) \ \cdots \ \mathbf{f}_k(t_{nm_n}))^T$ where $k = 1, 2, ..., r$; $M = \sum_{i=1}^{n} m_i$. By assuming that $\boldsymbol{\beta}$ is given, we can express the model in (2):

$$\mathbf{y} - \mathbf{X}\boldsymbol{\beta} = \mathbf{f}(t) + \boldsymbol{\varepsilon} . \qquad (3)$$

If we let $\mathbf{y} - \mathbf{X}\boldsymbol{\beta} = \mathbf{y}^*$ then equation (3) can be written as follows:

$$\mathbf{y}^* = \mathbf{f}(\mathrm{t}) + \boldsymbol{\varepsilon} . \qquad (4)$$

The nonparametric regression functions $f(t)$ for the k^{th} response ($k = 1, 2, ..., r$) in each subject *i* (*i*=1,2,…,*n)* and each observation *j* (*j* =1,2, … ,*m*) is a smooth continuous function and has (d+1)$^{\text{th}}$ derivatives on an interval about $t = t_0$ where t_0 is a fixed point and *d* represents the degree of polynomial. By applying Taylor's series about t_0 and $t \in (t_0 - h, t_0 + h)$ where *h* is bandwidth, we have the function as follows:

$$f_k(t) = f_k(t_0) + (t - t_0)^1 \frac{(f_k^1(t_0))}{1!} + (t - t_0)^2 \frac{(f_k^2(t_0))}{2!} + ... + (t - t_0)^{p_k} \frac{(f_k^{p_k}(t_0))}{p_k!} \qquad (5)$$

Suppose $\eta_d^{(k)}(t_0) = \dfrac{f_k^d(t_0)}{d!}$, $d = 0, 1, 2, ..., p_k$, then for $t \in (t_0 - h, t_0 + h)$, equation (5) can be written as follows:

$$f_k(t)=\eta_0^{(k)}(t_0)+(t-t_0)^1\eta_1^{(k)}(t_0)+(t-t_0)^2\eta_2^{(k)}(t_0)+\ldots+(t-t_0)^{p_k}\eta_{p_k}^{(k)}(t_0) \quad (6)$$

Smooth nonparametric regression function $f_k(t)$ in equation (5) is approached by local linear estimator. We will get a local linear estimator if the degree of polynomial equals to one (i.e., d=1). Therefore, for $t\in(t_0-h,t_0+h)$, equation (6) can be written as:

$$f_k(t)=\eta_0^{(k)}(t_0)+(t-t_0)\eta_1^{(k)}(t_0) \quad (7)$$

Function $\underset{\sim}{f}(t)$ in equation (4) can be expressed in matrix notation as follows:

$$\underset{\sim}{f}(t)=\mathbf{Z}(t_0)\underset{\sim}{\eta}(t_0) \quad (8)$$

where

$$\mathbf{Z}(t_0)=\begin{pmatrix}1 & t-t_0 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & 1 & t-t_0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & 1 & t-t_0\end{pmatrix};$$

$$\underset{\sim}{\eta}^{(\mathrm{k})}=\left(\eta_0^{(\mathrm{k})}(t_0) \quad \eta_1^{(\mathrm{k})}(t_0)\right)^T$$

To estimate $\underset{\sim}{\eta}(t_0)$ in equation (8), we use WLS method by determining value that minimises the following function:

$$Q(t_0)=\left(\underset{\sim}{Y}^*-\mathbf{Z}(t_0)\underset{\sim}{\eta}(t_0)\right)^T\mathbf{V}^{-1}\mathbf{K}_h(t_0)\left(\underset{\sim}{Y}^*-\mathbf{Z}(t_0)\underset{\sim}{\eta}(t_0)\right) \quad (9)$$

where $\mathbf{V}^{-1}$ is weighted matrix that is formed from invert of covariance matrix of error with dimension $(rM\times rM)$, and $\mathbf{K_h}$ is kernel weight matrix as follows:

$$\begin{aligned}\mathbf{K}_h=diag(&K_h(t_{11}-t_0),K_h(t_{12}-t_0),\ldots,K_h(t_{1n(1)}-t_0),K_h(t_{21}-t_0),K_h(t_{22}-t_0),\ldots,\\ &K_h(t_{2n(2)}-t_0),K_h(t_{m1}-t_0),K_h(t_{m2}-t_0),\ldots,K_h(t_{mn(m)}-t_0),K_h(t_{11}-t_0),K_h(t_{12}-t_0),\ldots,\\ &K_h(t_{1n(1)}-t_0),K_h(t_{21}-t_0),K_h(t_{22}-t_0),\ldots,K_h(t_{2n(2)}-t_0),K_h(t_{m2}-t_0),\ldots,K_h(t_{mn(m)}-t_0))\end{aligned}$$

and $K_h(\cdot)$ is kernel function with optimum bandwidth h. The estimation of $\underset{\sim}{\eta}(t_0)$ is $\hat{\underset{\sim}{\eta}}(t_0)$ that is derived from $\dfrac{\partial Q(t_0)}{\partial \underset{\sim}{\eta}(\mathrm{t}_0)}=\underset{\sim}{0}$, and we get the estimator as follows:

$$\underset{\sim}{\hat{\eta}}(t_0)=\left(\mathbf{Z}^T(t_0)\mathbf{V}^{-1}\mathbf{K_h}(t_0)\mathbf{Z}(t_0)\right)^{-1}\mathbf{Z}^T(t_0)\mathbf{V}^{-1}\mathbf{K_h}(t_0)\underset{\sim}{y}^* \quad (10)$$

So, we get the estimation of $\underset{\sim}{f}(t)$ based on local linear estimator as follows:

$$\underset{\sim}{\hat{f}}(t)=\mathbf{A}_h(\underset{\sim}{t})\underset{\sim}{y}^* . \quad (11)$$

Based on equation (5) and equation (11), we can express the local linear estimator for $\underset{\sim}{\hat{f}}(t)$ as:

$$\underset{\sim}{\hat{f}}(\underset{\sim}{t})=\mathbf{A}_h(\underset{\sim}{t})((\underset{\sim}{y}-\mathbf{X}\underset{\sim}{\beta})) \quad (12)$$

Therefore, the function sum of squared error of equation (1) is given by:

$$S=[\underset{\sim}{y}-\mathbf{X}\underset{\sim}{\beta}-(\mathbf{A}_h(\underset{\sim}{t})(\underset{\sim}{y}-\mathbf{X}\underset{\sim}{\beta}))]^T[\underset{\sim}{y}-\mathbf{X}\underset{\sim}{\beta}-(\mathbf{A}_h(\underset{\sim}{t})(\underset{\sim}{y}-\mathbf{X}\underset{\sim}{\beta}))] \quad (13)$$

By minimising S in equation (13), we obtain estimation of β, that is, $\underset{\sim}{\hat{\beta}}$ as follows:

$$\underset{\sim}{\hat{\beta}}=[\mathbf{X}^T(\mathrm{I}-\mathbf{A}_h(\underset{\sim}{t}))^T(\mathrm{I}-\mathbf{A}_h(\underset{\sim}{t}))\mathbf{X}]^{-1}\mathbf{X}^T(\mathrm{I}-\mathbf{A}_h(\underset{\sim}{t}))^T(\mathrm{I}-\mathbf{A}_h(\underset{\sim}{t}))\underset{\sim}{y} \quad (14)$$

By substituting estimator $\underset{\sim}{\hat{\beta}}$ in equation (14) into equation (13), we have:

$$\underset{\sim}{\hat{f}}(\underset{\sim}{t})=\mathbf{A}_h(\underset{\sim}{t})[\mathbf{I}-\mathbf{X}(\mathbf{X}^T\mathbf{S}(\underset{\sim}{t})\mathbf{X})^{-1}\mathbf{X}^T\mathbf{S}(\underset{\sim}{t})]\underset{\sim}{y} \quad (15)$$

where $\mathbf{S}(\underset{\sim}{t})=(\mathbf{I}-\mathbf{A}_h(\underset{\sim}{t}))^T(\mathbf{I}-\mathbf{A}_h(\underset{\sim}{t}))$.

Finally, based on Eq. (14) and Eq. (15), we obtain the following estimated model:

$$\underset{\sim}{\hat{y}}=(\mathbf{A}_{\mathrm{par}}+\mathbf{A}_{\mathrm{nonpar}})\underset{\sim}{y}=\mathbf{A}_h(\underset{\sim}{t})\underset{\sim}{y} \quad (16)$$

where $\mathbf{A}_h=\mathbf{A}_{\mathrm{par}}+\mathbf{A}_{\mathrm{nonpar}}$, $\mathbf{A}_{\mathrm{par}}=\mathbf{X}[\mathbf{X}^T\mathbf{S}(\underset{\sim}{t})\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{S}(\underset{\sim}{t})$, and
$\mathbf{A}_{\mathrm{nonpar}}=\mathbf{A}_h(\underset{\sim}{t})[\mathbf{I}-\mathbf{X}(\mathbf{X}^T\mathbf{S}(\underset{\sim}{t})\mathbf{X})^{-1}\mathbf{X}^T\mathbf{S}(\underset{\sim}{t})]$

Since in this study we use a linear regression approach, there is a parameter called bandwidth (h) which controls the roughness of fit and also affects trade-off of biased-variance. Optimal bandwidth (h) value is determined based on GCV criterion where GCV is given as follows:

$$GCV(h) = \frac{(\frac{1}{rM})\mathbf{y}^T(\mathbf{I}-\mathbf{S}(\underset{\sim}{t}))^T(\mathbf{I}-\mathbf{S}(\underset{\sim}{t}))\mathbf{y}}{[(\frac{1}{rM})trace(\mathbf{I}-\mathbf{A}_h)]^2} \quad (17)$$

Data and Analysis Steps

The data set used in this study includes children aged up to five years of age in East Java province. The data set consists of 30,490 boys and 28,680 girls recorded from 20 cities (Lamongan, Sampang, Madiun, Jember, Banyuwangi, Gresik, Probolinggo, Mojokerto, Nganjuk, Bojonegoro, Jombang, Surabaya, Tulungagung, Sidoarjo, Kediri, Malang, Lumajang, Ngawi, Tuban and Pamekasan) during 2018.

In this case, we use the multi-response semiparametric regression model approach that consists of the first response (y_1), i.e., weight, second response (y_2), i.e., height, third response (y_3), i.e., body mass index (BMI), and the first predictor (t) as a nonparametric component, i.e., age, and the second predictor (x) as a parametric component, i.e., gender. We need to create OSS-R code to analyse data. By designing the LSGC (locally standard growth charts) of children aged up to five years for weight, height and BMI based on age, and by considering WHO-MGRS (2006), we get the chart area which is useful for assessing children's nutritional status as outlined in Table 1.

Table 1: Chart Area for Assessing Nutritional Status

Area of Chart	Nutritional Status		
	Weight for Age	Height for Age	BMI for Age
$< 3^{rd}$ percentile	Severely Underweight	Severely Stunted	Severely Wasted
$\geq 3^{rd}$ percentile and $< 15^{th}$ percentile	Underweight	Stunted	Wasted
$\geq 15^{th}$ percentile and $\leq 85^{th}$ percentile	Normal	Normal	Normal
$> 85^{th}$ percentile	Overweight	Tall	Risk Overweight

Estimation of LSGC for Boys and Girls

The result of analysis for median (P_{50}) of data shows that Pearson's correlation between height and weight is 0.99. Based on equation (16) and (17), we get optimal bandwidth value of 1.2, minimum value of GCV of 0.5030. Based on optimal bandwidth, we estimate the median growth charts of weight/age, height/age and BMI/age for boys and girls by using multi-response semiparametric regression based on a local linear estimator that provides a coefficient of determination (R^2) of 0.997 and the mean squared error of 0.21. The estimated median growth charts of weight/age, height/age and BMI/age for boys and girls are provided in Figures 1-3. Figures 1-3 show that the median growth of boys in East Java are higher than those of girls. Their average differences are 0.40 kg (weight-for-age), 0.96 cm (height-for-age) and 0.24 kg/m^2 (BMI-for-age).

Figure 1. Estimated median chart of weight for boys and girls

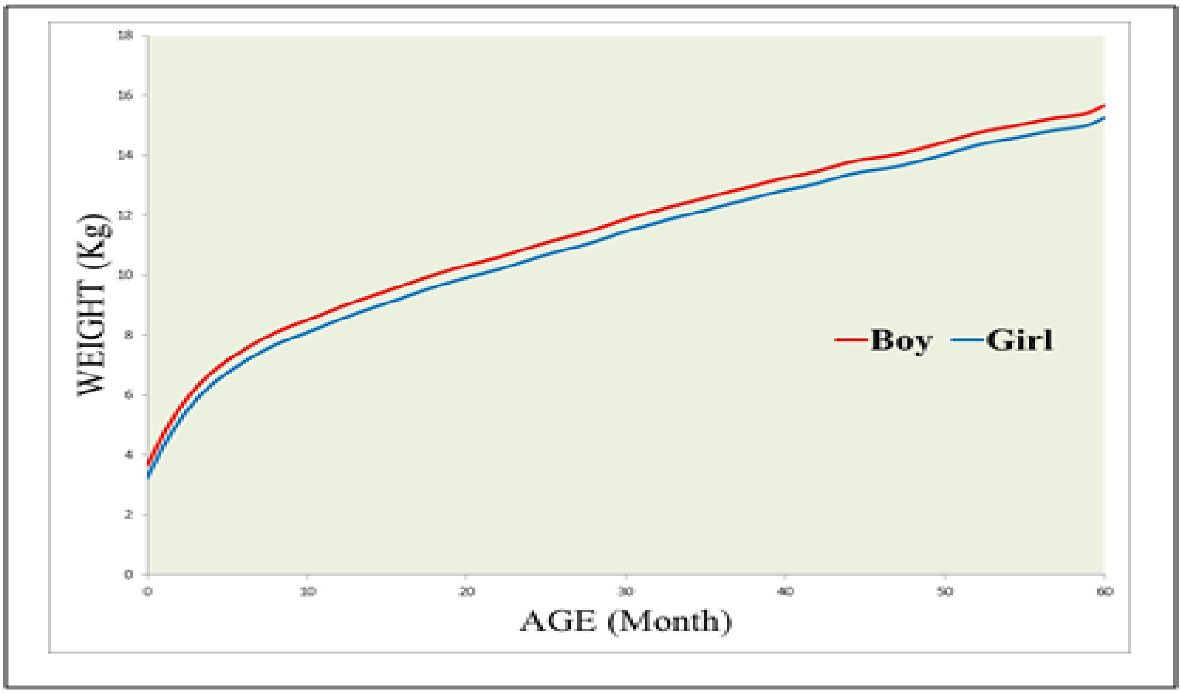


Figure 2. Estimated median chart of height for boys and girls

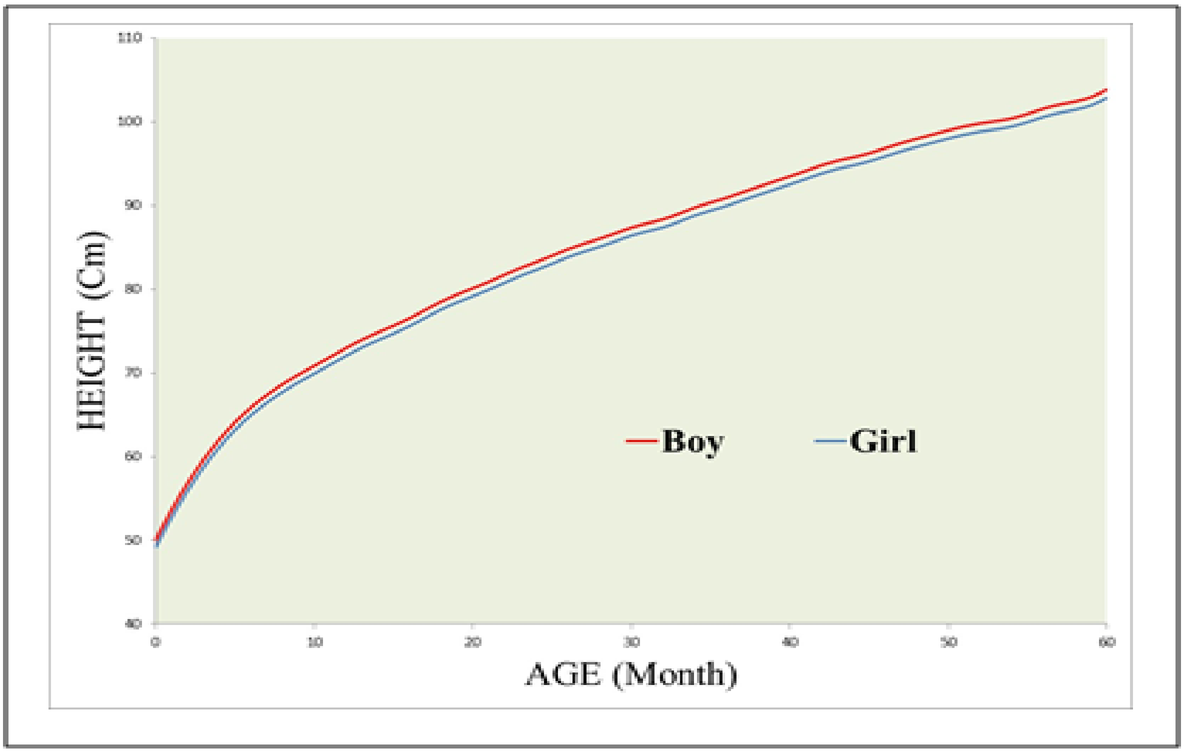
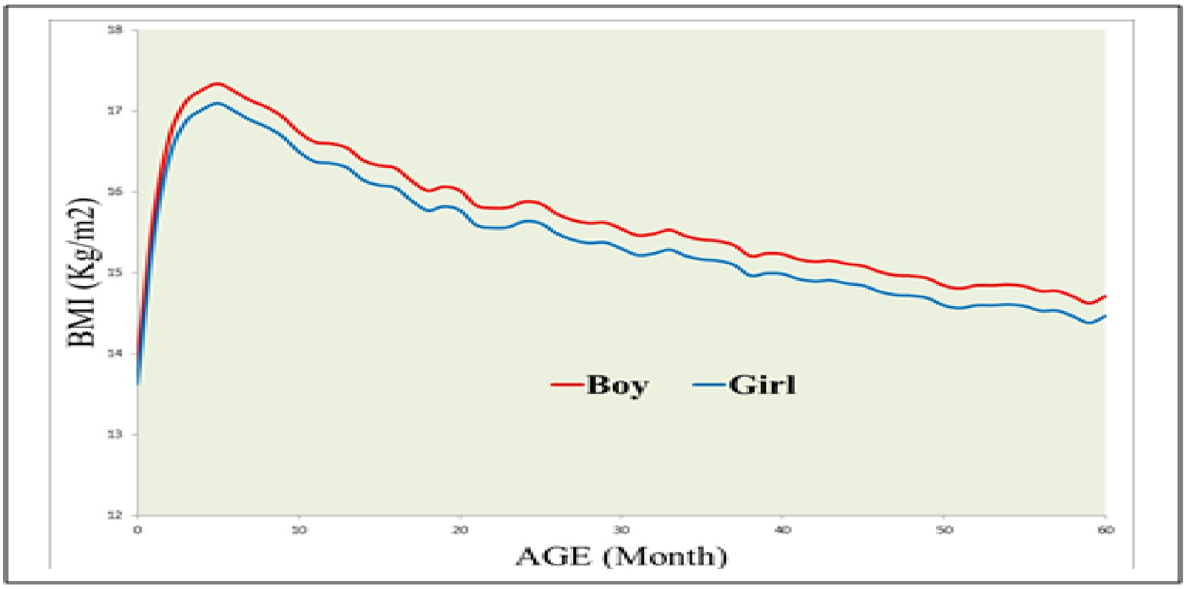


Figure 3. Estimated median chart of BMI of boys and girls



Furthermore, we compare East Javanese children's growth of weight, height and BMI between LSGC and WSGC as provided in Figures 4-9. The Figures 4 - 9 show that median growth for boys and girls of LSGC (blue line) are less than those of WSGC (red line). The average difference for boys are: 1.4 kg (weight-for-age), 4.25 cm (height-for-age) and 0.2 kg/m^2 (BMI-for-age), while for girls they are: 1.8 kg (weight-for-age), 4.07 cm (height-for-age) and 0.2 kg/m^2 (BMI-for-age). This is due to the fact that the samples used to design WSGC came from multi-ethnic and multi-states, including the US, Oman, Norway, India, Ghana and India (WHO-MGRS, 2006). In Figures 8-9, BMI medians jump at age 24 month. According to WHO-MGRS (2006) BMI is ratio weight (in kg)/recumbent length (birth to 24 month) or standing height (from 24 month to 60 month), so the resulting disjunction between two standards reflects the 0.7 cm difference between length and height. The optimal h, MSE and R^2 values of estimated model for percentiles 3rd, 15th, 50th, 85th, and 97th of weight, height and BMI are provided in Table 2. Based on optimal bandwidth (h_1, h_2, h_3) for percentiles 3rd, 15th, 50th, 85th, and 97th in Table 2, we construct local standard charts of weight/age, height/age, and BMI/age for boys and girls based on samples of East Javanese children as indicated in Figures 10-15, respectively.

Figure 4. Comparison of median weight for boys between East Java and WHO-2005 charts

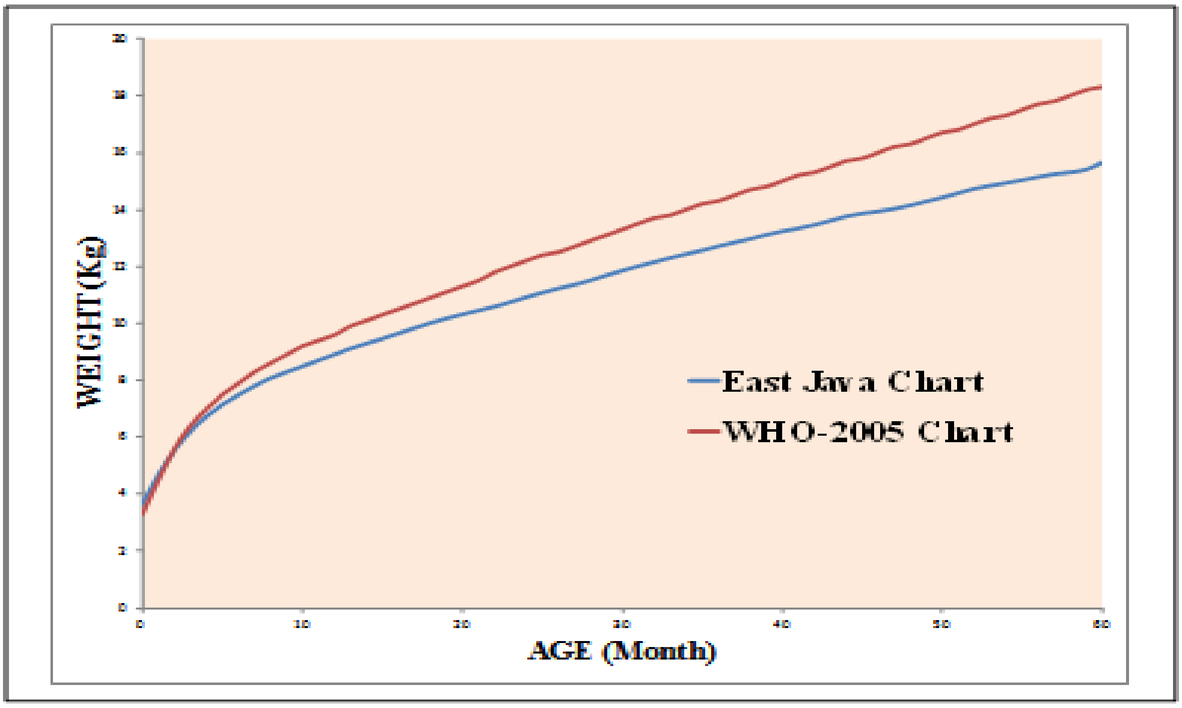


Figure 5. Comparison of median weight for girls between East Java and WHO-2005 charts

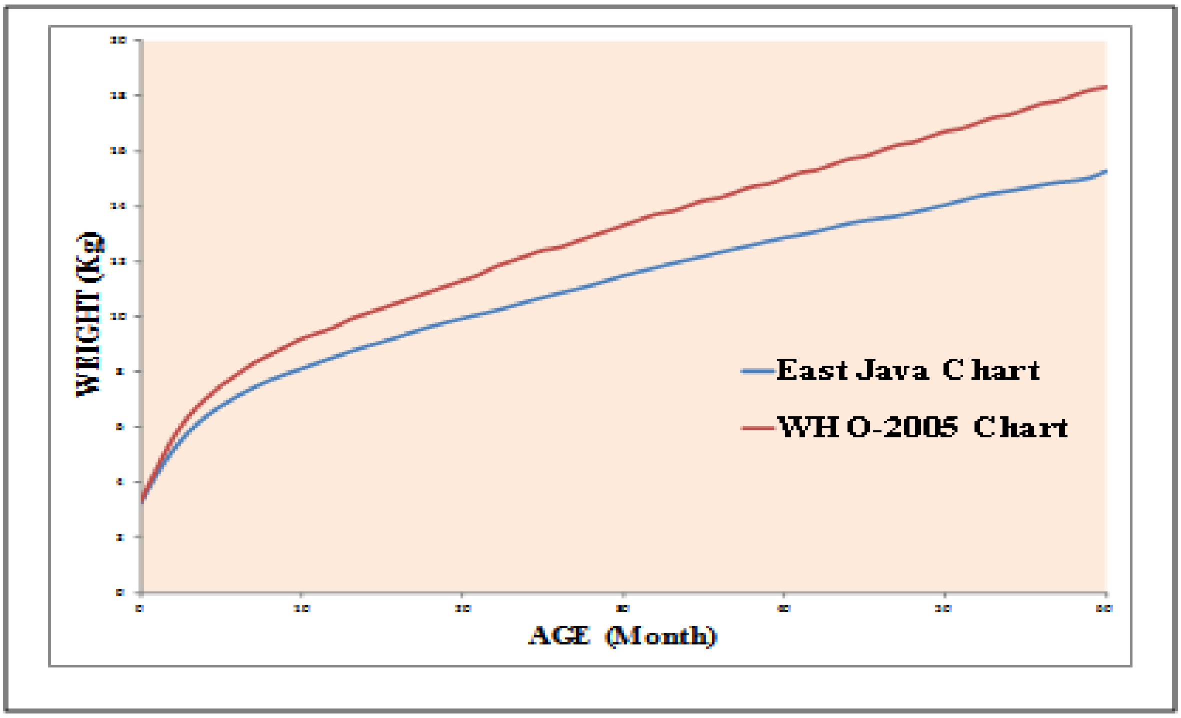


Figure 6. Comparison of median height for boys between East Java and WHO-2005 charts

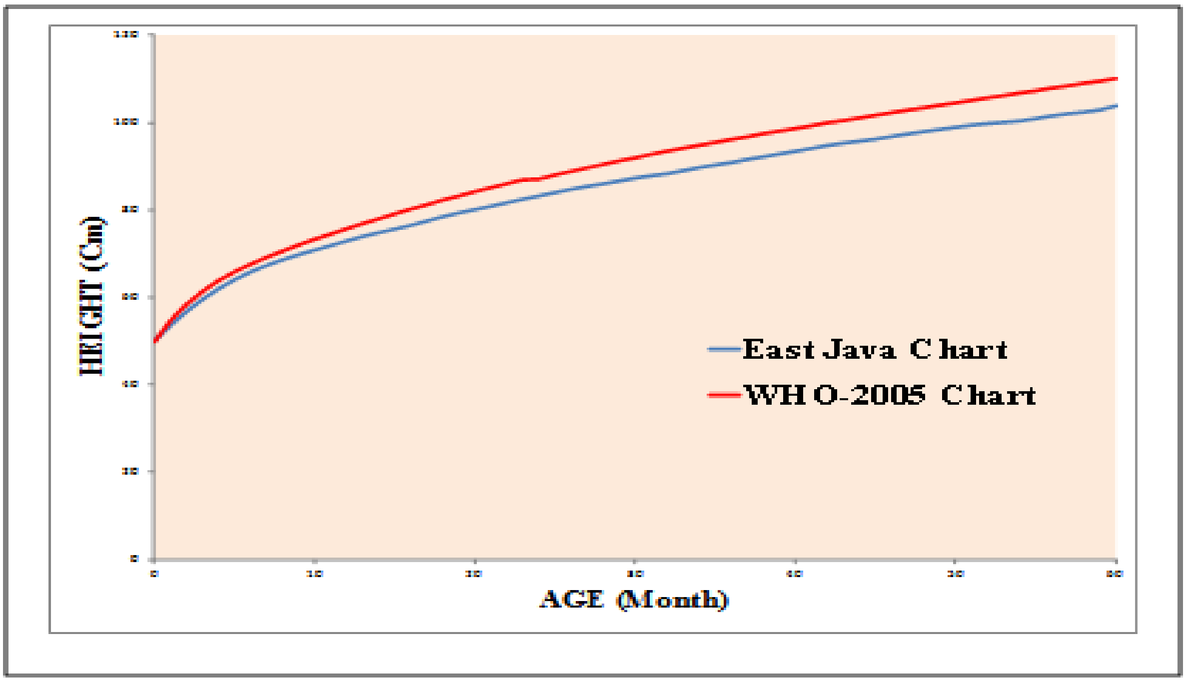


Figure 7. Comparison of median height for girls between East Java and WHO-2005 charts

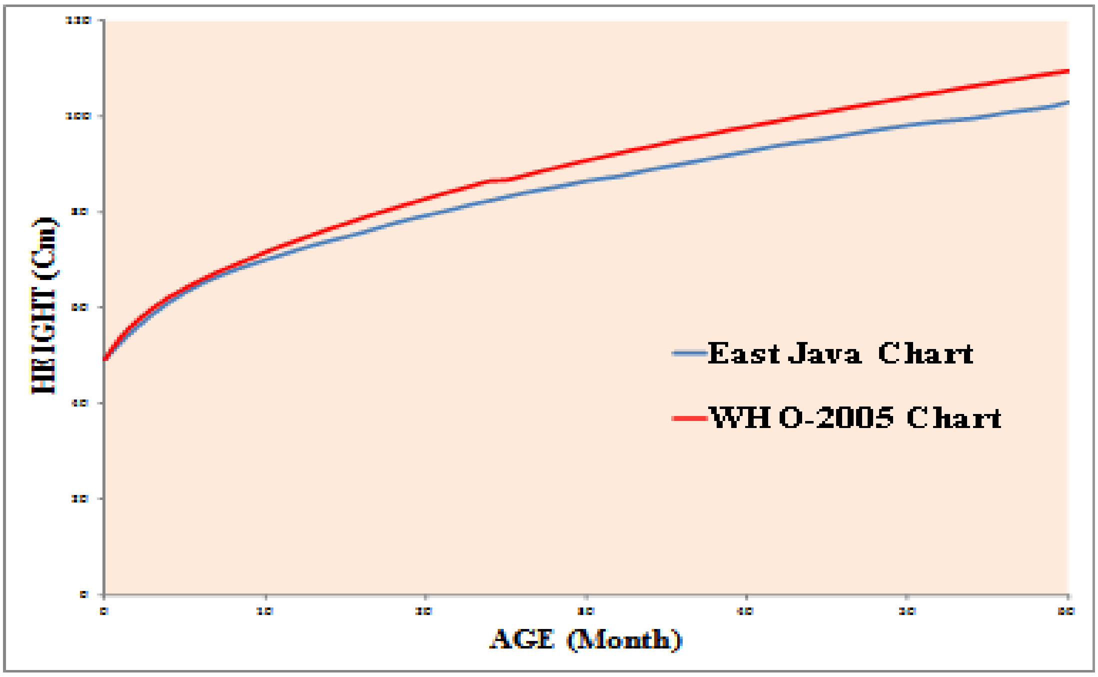


Figure 8. Comparison of median BMI for boys between East Java and WHO-2005 charts

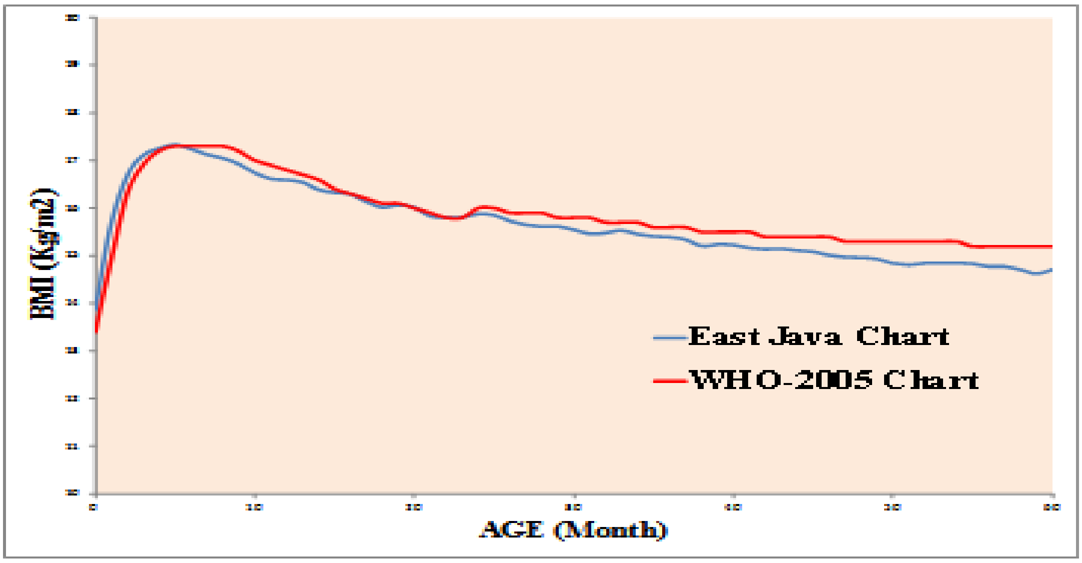


Figure 9. Comparison of median BMI for girls between East Java and WHO-2005 charts

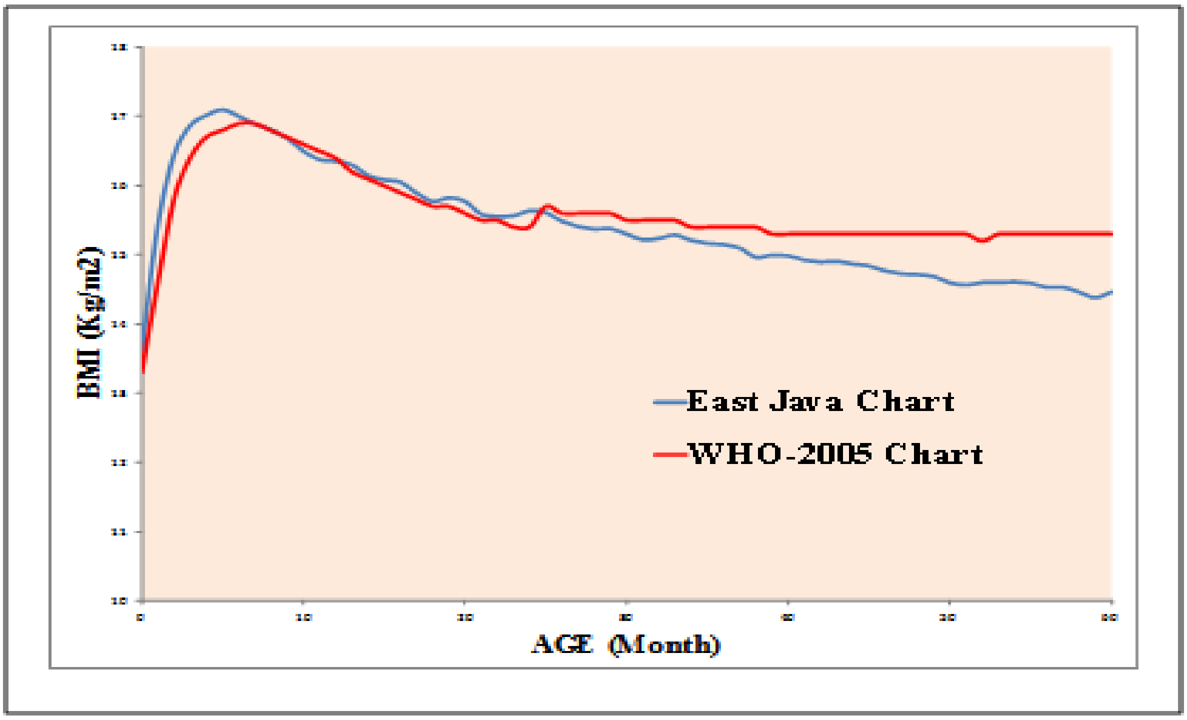


Table 2: Optimal h , Minimum GCV, MSE and R2 values in each Percentile

Percentile	Optimal $h_1 h_2 h_3$	Minimum GCV	MSE	R^2
3rd	1; 2.1; 1.2	0.7276	0.46	99.58%
15th	1; 1.4; 0.7	0.4832	0.10	99.99%
50th	1.2; 1.1; 0.7	0.5030	0.21	99.7%
85th	1.5; 1; 0.9	0.4390	0.14	99.99%
97th	2.8; 1.4; 1	1.2775	0.72	99.98%

Figure 10. Standard growth chart of weight for boys in East Java Province

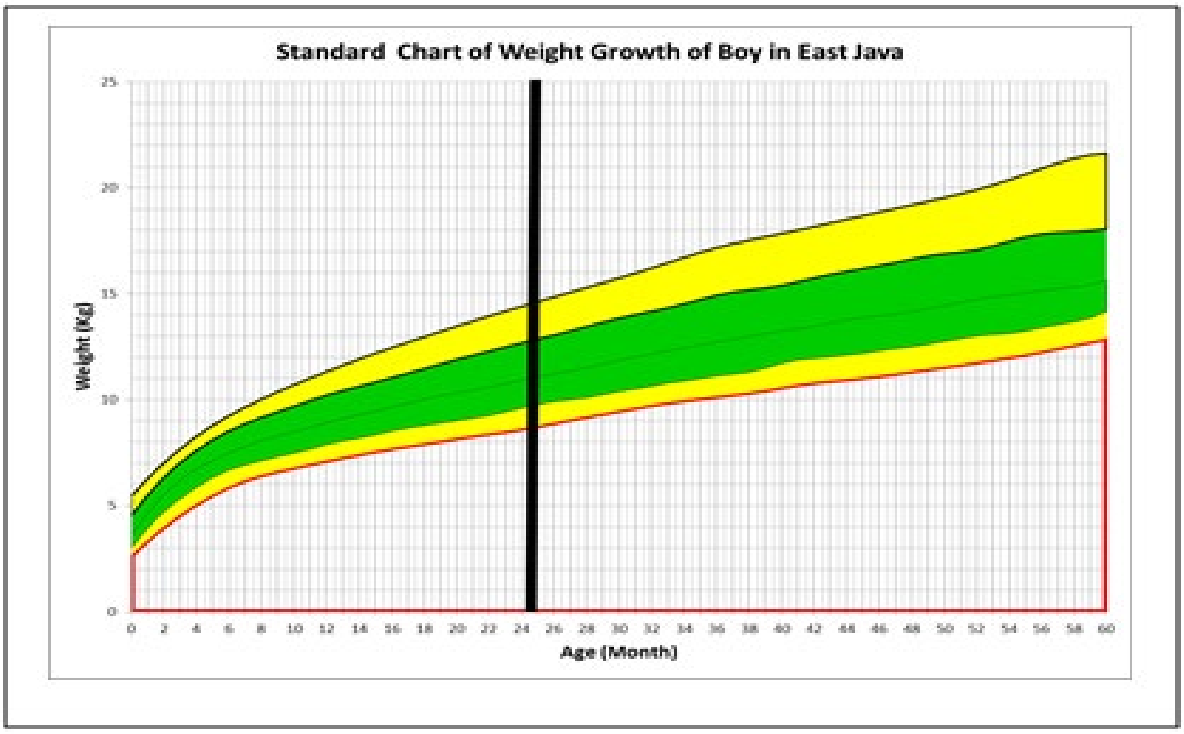


Figure 11. Standard growth chart of weight for girls in East Java Province

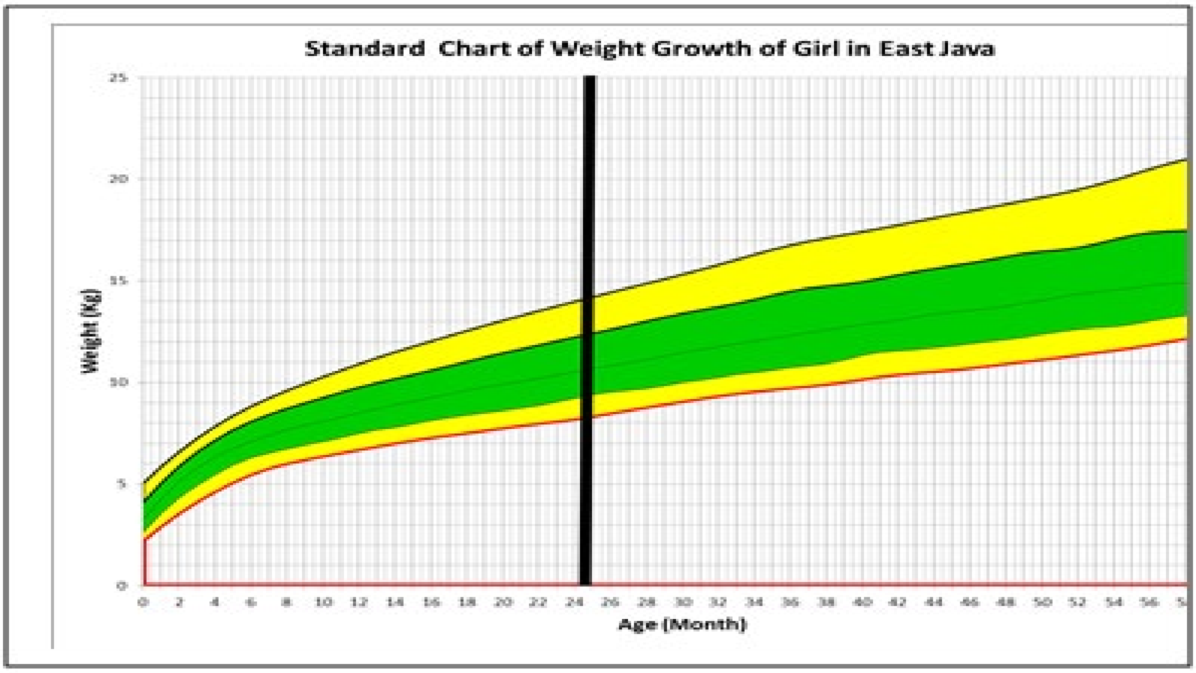


Figure 12. Standard growth chart of height for boys in East Java Province

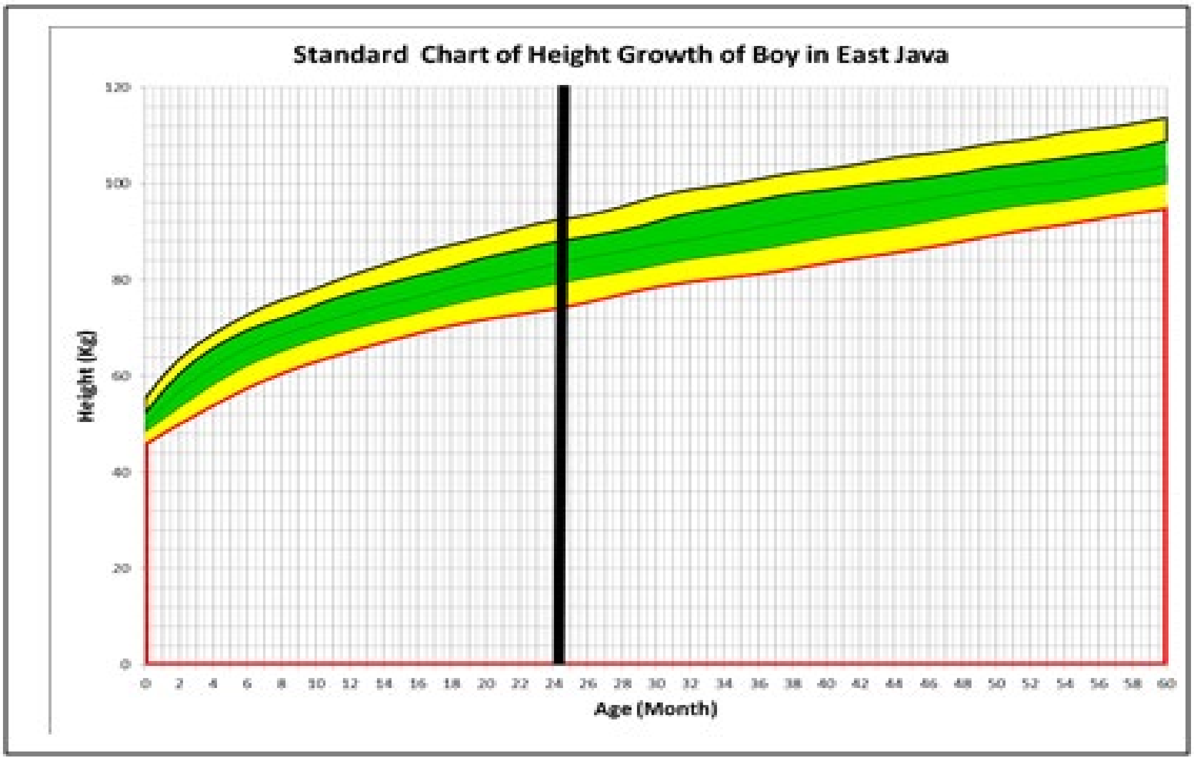


Figure 13. Standard growth chart of height for girls in East Java Province

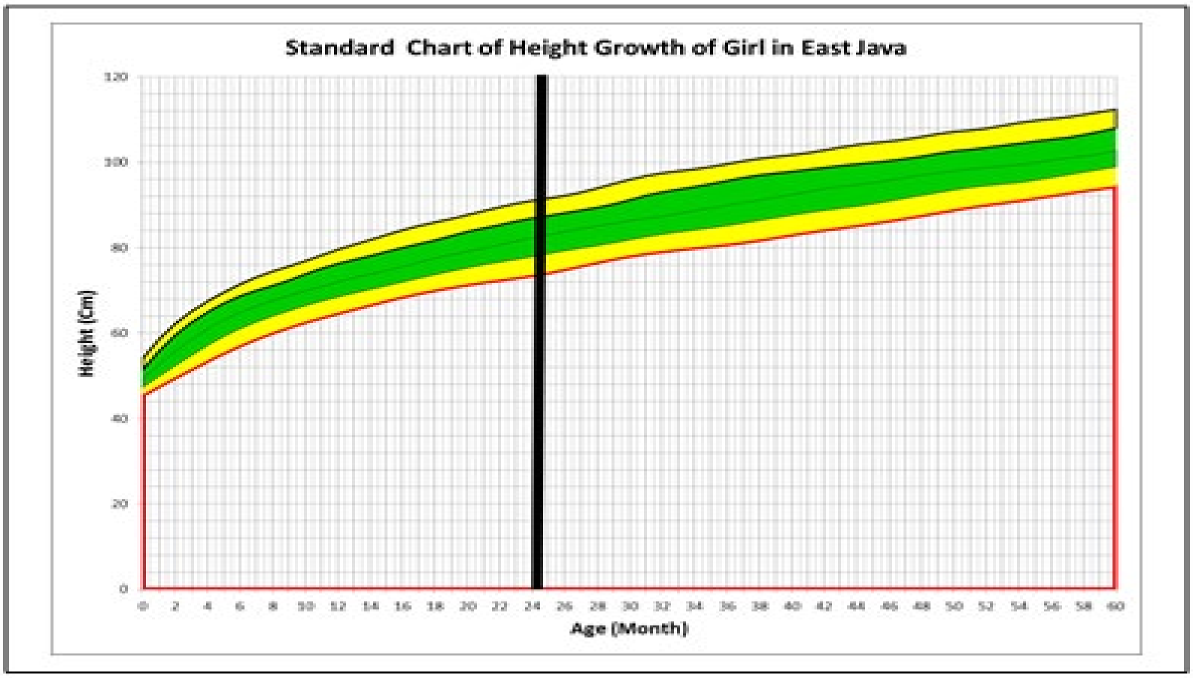


Figure 14. Standard growth chart of BMI for boys in East Java Province

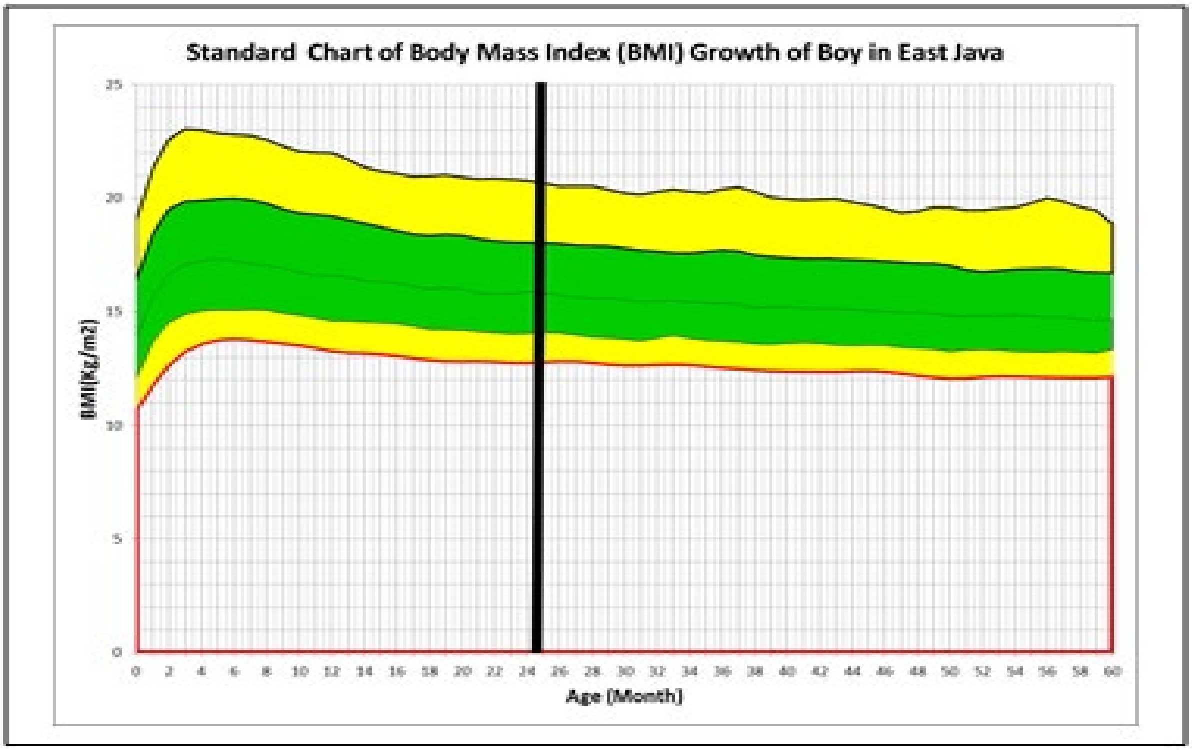
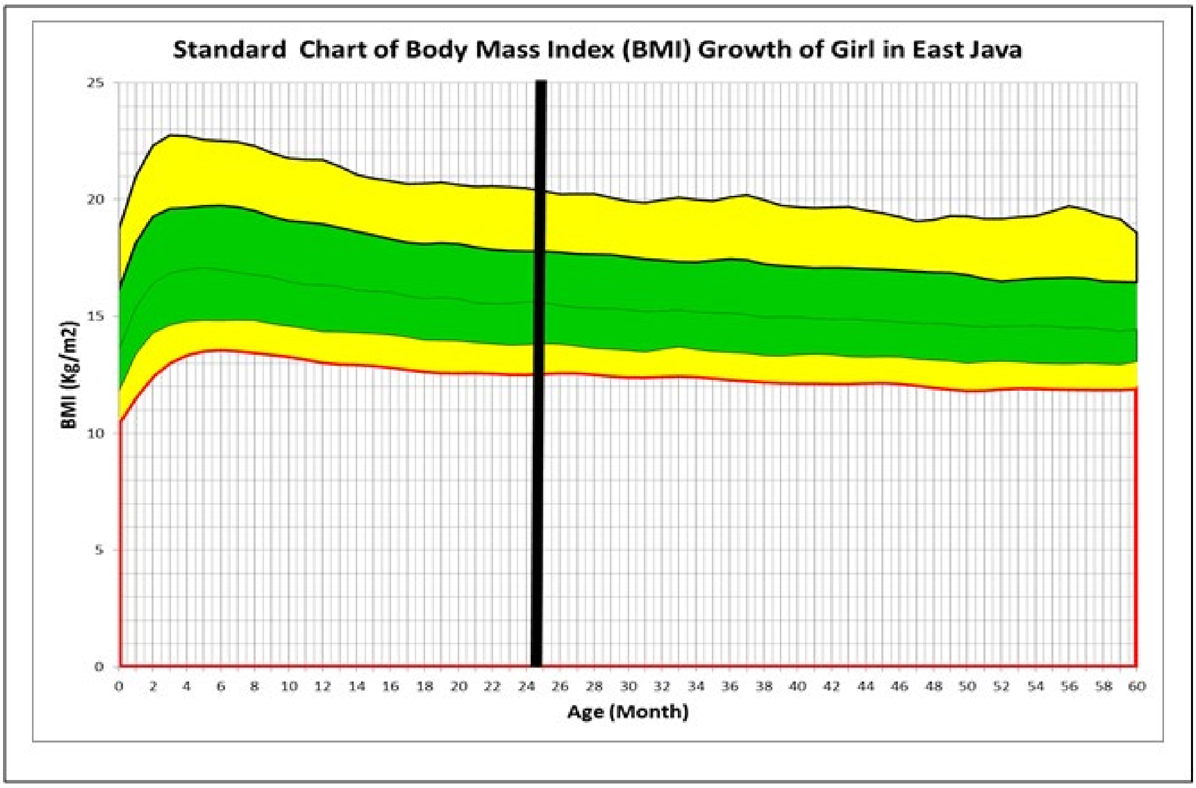


Figure 15. Standard growth chart of BMI for girls in East Java Province



Assessing Status of Nutrition for Children in East Java Province

We assessed the nutritional status of 30,490 boys and 28,680 girls recorded from 20 cities in East Java province. The percentage of nutritional status for boys and girls is indicated in Tables 3-5. Based on Table 3-5 and according to LSGC the percentage of severely underweight nutritional status, severely stunted nutritional status, and severely wasted nutritional status for boys and girls regarding weight (for age), height (for age), and BMI (for age) are smaller than those according to WSGC. This means that East Javanese standard growth chart for weight/age, height/age, and BMI/age by using local linear estimator of multi-response semiparametric regression model approach is lower than the WHO-2005 standard growth chart. Furthermore, according to LSGC the percentage of normal status of nutrition based on weight/age and normal status of nutrition status based on height/age for boys and girls are greater than those according to WSGC. The differences of percentage of normal nutritional status for boys based on weight/age, height/age, and BMI/age between LSGC and WSGC are 19.76%, 32.36%, and 19.7%, respectively, while the differences of percentages of normal nutritional status for girls based on weight/age, height/age, and BMI/age between LSGC and WSGC are 16.6%, 29.79%, and 17.27%, respectively. Consequently, nutritional status based on weight/age, height/age, and BMI/age for girls in East Java is closer to the WHO-2005 standard growth chart than those for boys.

Table 3: Percentage of Nutritional Status of Boys and Girls for Weight/Age in East Java

Nutritional Status for Weight/Age	Boys		Girls	
	LSGC	WSGC	LSGC	WSGC
Severely Underweight (< 3rd percentile)	3.05 %	16.95 %	3.19 %	14.57 %
Underweight (≥3rd percentile and <15th percentile)	12.20 %	26.13 %	12.03 %	24.98 %

Table 4: Percentage of Nutritional Status of Boys and Girls for Height/Age in East Java

Nutritional Status for Height/Age	Boys		Girls	
	LSGC	WSGC	LSGC	WSGC
Severely Stunted (< 3rd percentile)	3.18 %	29.33%	2.93%	26.26%
Stunted	12.05 %	27.13 %	11.94%	26.73%

Table 5: Percentage of Nutritional Status of Boys and Girls for BMI/Age in East Java

Nutritional Status for BMI/Age	Boys		Girls	
	LSGC	WSGC	LSGC	WSGC
Severely Wasted (< 3rd percentile)	2.98%	11.46%	3.14%	9.78%
Wasted (≥3rd percentile and <15th percentile)	12.02%	16.12%	12.01%	16.26%
Normal (≥15th percentile and ≤85th percentile)	69.81%	50.07%	69.57%	52.30%
Risk Overweight (>85th percentile)	15.19%	22.35%	15.28%	21.66%

Within these anthropometric measurements, the overall standard growth chart based on BMI/age for girls of LSGC is closest to that of WSGC. Also, the differences severely stunted nutritional status of boys between LSGC and WSGC is the greatest compared with severely underweight and severely wasted nutritional status, i.e., equalling 32.36%. Generally, the LSGC of children is lower than WSGC . This is due to the fact that e samples used to design LSGC and WSGC are very different. WHO-2005 standard growth chart used samples of multi ethnic and multi states (WHO-MGRS, 2006), while, LSGC used samples of multi ethnic and mono state. The actual condition of children from these six countries is physically different from the real condition of children in East Javanese Province, Indonesia, especially their status of nutrition based on height/age.

Conclusion

The estimated model of local standard growth charts (LSGC) of children using multi-response local linear estimator satisfies goodness of fit criteria that are an average of R^2, i.e., 99.85 % and the average of MSE, i.e., 0.326 (based on Table 2). The LSGC of children are designed based on samples of children in East Java, Indonesia, and by using local linear estimator of a multi-response semiparametric regression model for assessing the nutritional status for children in East Java Province, Indonesia are significantly more suitable than WSGC, due to the fact that the actual condition of children from six multi-ethnic and multi-state countries used as sample by WSGC is physically different from the condition of children from East Java, Indonesia, especially based on height/age.

Acknowledgment

The authors thank the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for funding this research through the Fundamental Research (*Penelitian Dasar*

Unggulan Perguruan Tinggi - PDUPT) Grant during the fiscal year 2019 with Contract Number: 727/UN3.14/LT/2019.

REFERENCES

Ana, E., Chamidah, N., Andriani, P. and Lestari, B. (2019). Modeling of hypertension risk factors using local linear of additive nonparametric logistic regression. *Journal of Physics: Conference Series*, **1397** 012067.

Aydin, D. and Yilmaz, E. (2018). Modified spline regression based on randomly right-censored data: A comparative study. *Communications in Statistics-Simulation and Computation*, **47**(9), 2587-2611.

Chamidah, N., Budiantara, I.N., Sunaryo, S. and Zain, I. (2012). Designing of child growth curve based on multiresponse local polynomial modeling. *Journal of Mathematics and Statistics*, **8**(3), 342-347.

Chamidah, N., Gusti, K. H., Tjahjono, E. and Lestari, B. (2019a). Improving of classification accuracy of cyst and tumour using local polynomial estimator. *TELKOMNIKA*, **17**(3), 1492-1500.

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

Chamidah, N. and Lestari, B. (2016). Spline estimator in homoscedastic multiresponse nonparametric regression model in case of unbalanced number of observations. *Far East Journal of Mathematical Sciences*, **100**(9), 1433-1453.

Chamidah, N. and Lestari, B. (2019). Estimation of covariance matrix using multiresponse local polynomial estimator for designing children growth charts: A theoretically discussion. *Journal of Physics: Conference Series*, **1397** 012072.

Chamidah, N., Lestari, B. and Saifudin, T. (2019). Predicting blood pressures and heart rate associated with stress level using spline estimator: A theoretically discussion. *International Journal of Academic and Applied Research*, **3**(10), 5-12.

Chamidah, N. and Rifada, M. (2016a). Local linear estimator in biresponse semiparametric regression model for estimating median growth chart of children. *Far East Journal of Mathematical Sciences*, **99**(8), 1233-1244.

Chamidah, N. and Rifada, M. (2016b). Estimation of median growth curves for children up two years old based on biresponse local linear estimator. *AIP Conference Proceedings*, **1718** 110001.

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.

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.

Chamidah, N., Zaman, B., Muniroh, L. and Lestari, B. (2019b). Estimation of median growth charts for height of children in East Java province of Indonesia using penalized spline estimator. *GCEAS-Hokkaido International Conference Proceeding*, **201907**, 68-78.

Darnah, Utoyo, M. I. and Chamidah, N. (2019). Modeling of maternal mortality and infant mortality cases in East Kalimantan using poisson regression approach based on local linear estimator. *IOP Conf. Series: Earth and Environmental Science*, **243** 012023.

Global Nutrition Report (2014). *Actions and accountability to accelerate the world's progress on nutrition. A-Peer-Reviewed Publication*, Washington DC, USA.

Hidayati, L., Chamidah, N. and Budiantara, I. N. (2019). Spline truncated estimator in multiresponse semiparametric regression model for computer based national exam in West Nusa Tenggara. *IOP Conf. Series: Materials Science and Engineering*, **546** 052029.

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.

Lestari, B., Fatmawati, Budiantara, I., N. & Chamidah, N. (2019). Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and kernel estimators approaches. *Journal of Physics: Conference Series*, **1397** 012064.

Lestari, B., Fatmawati and Budiantara, I. N. (2019a). Smoothing spline estimator in multi-response nonparametric regression for predicting blood pressures and pulse: a theoretically discussion. *GCEAS-Hokkaido International Conference Proceeding*, **201907**, 81-93.

Lestari, B., Fatmawati and Budiantara, I. N. (2019b). Smoothing spline estimator in multi-response nonparametric regression for predicting blood pressures and heart rate. *International Journal of Academic and Applied Research*, **3**(9), 1-8.

Lestari, B., Fatmawati and Budiantara, I. N. (2019c). Estimation of multiresponse nonparametric regression model using smoothing spline estimator: A simulation study. *International Journal of Academic and Applied Research*, **3**(10), 1-4.

Massaid, A., Hanif, M., Febrianti, D. and Chamidah, N. (2019). Modelling of poverty percentage of non-food per capita expenditures in Indonesia using least square spline estimator. *IOP Conf. Series: Materials Science and Engineering*, **546** 052044.

Murbarani, N., Swastika, Y., Dwi, A., Aris, B. and Chamidah, N. (2019). Modeling of the percentage of AIDS sufferers in East Java province using nonparametric regression approach based on truncated spline estimator. *Indonesian Journal of Stat. and Its Applications*, **3**(2), 139-147.

Nidhomuddin, Chamidah, N. and Kurniawan, A. (2019). Admission test modelling of state Islamic college in Indonesia using local linear for bivariate longitudinal data. *IOP Conf. Series: Materials Science and Engineering*, **546** 052047.

Nottingham Q.J., and Cook, D. F. (2011). Local linear regression for estimating time series data. *Journal of Computational Statistics and Data Analysis*, **37**, 209–217.

Puspitawati, A. and Chamidah, N. (2019). Choroidal neovascularisation classification on fundus retinal images using local linear estimator. *IOP Conf. Series: Materials Science and Engineering*, **546** 052056.

Ramadan, W., Chamidah, N., Zaman, B., Muniroh, L. and Lestari, B. (2019). Standard growth chart of weight for height to determine wasting nutritional status in East Java based on semiparametric least square spline estimator. *IOP Conf. Series: Materials Science and Engineering*, **546** 052063.

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.

WHO-MGRS (Multicentre Growth Reference Study Group) (2006). *WHO Child Growth Standards: Length/height-for-age, Weight-forage, Weight-for-length, Weight-for-height and Body mass index-for-age: Methods and Development*, World Health Organization, Geneva.