JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 10, No. 3, July 2026, pp. 797-809

Nonparametric Biresponse Penalized Spline Regression for Modeling Stunting and Wasting in Kalimantan

Samsul Arifin[1*], Ardiansyah Abubakar[1], A. Fahmi Indrayani[1]
[1]Department of Statistics, Universitas Lambung Mangkurat, Indonesia
samsularrr@ulm.ac.id

Article History:
Received : 19-02-2026
Revised : 18-03-2026
Accepted : 25-03-2026
Online : 01-07-2026

Keywords:
Stunting;
Wasting;
Penalized Spline;
Birespon.

ABSTRACT

Stunting and wasting remain major nutritional problems in Indonesia, including in Kalimantan, with considerable interregional variation. This condition suggests that the relationships between determinants and indicators of child nutritional status may be nonlinear and interdependent. This study aims to develop and implement a biresponse nonparametric penalized spline regression model to simultaneously model the prevalence of stunting (Y_1) and wasting (Y_2) in Kalimantan. The data used secondary data from the 2024 Indonesian Nutritional Status Survey (SSGI) of the Ministry of Health and official publications from Statistics Indonesia (BPS), with districts/cities in Kalimantan as the unit of analysis. The predictor variables included the percentage of households with access to improved sanitation (X_1), low birth weight (X_2), and the percentage of the population covered by health insurance (X_3). The Pearson correlation test indicated a significant association between stunting and wasting (p-value = 0.012), supporting the application of a biresponse modeling approach. Model selection was conducted simultaneously for the knot points and the smoothing parameter (λ) using the minimum Generalized Cross-Validation (GCV) criterion. The optimal configuration was obtained with one knot for each predictor, namely $X_1 = 66$, $X_2 = 107$, and $X_3 = 53$, with $\lambda = 63.09$ and GCV = 20.83. Model performance evaluation yielded MSE = 28.012 and $R^2 = 0.241$ for stunting, and MSE = 4.810 and $R^2 = 0.106$ for wasting. These results indicate that the birespose penalized spline model can serve as a flexible approach for simultaneously analyzing stunting and wasting and for capturing heterogeneous, nonlinear relationships between predictors and response variables.

Crossref
https://doi.org/10.31764/jtam.v10i3.38418

———————— ◆ ————————

A. INTRODUCTION

Stunting and wasting remain major nutritional problems in Indonesia (Wulandari et al., 2025). According to the 2023 Indonesian Nutritional Status Survey (SSGI), the national prevalence of stunting remains above 20%, although it shows a declining trend compared to previous years (Rahmadiani et al., 2025). This figure remains considerably higher than the national target of 14% set out in the stunting reduction acceleration policy. Meanwhile, the national prevalence of wasting ranges between 8–10% and continues to be classified as a public health problem according to WHO criteria (Jokhu & Syauqy, 2024). These conditions indicate that chronic and acute malnutrition problems persist simultaneously among children under five in Indonesia.

In Kalimantan, the prevalence of stunting and wasting remains substantial and varies considerably across districts (Kustanto et al., 2025). Some areas report stunting rates above the national average, and wasting is also present at significant levels. Disparities in access to health services, sanitation conditions, poverty levels, and heterogeneous regional characteristics are suspected to contribute to these variations (Eryando et al., 2022). This uneven distribution pattern suggests that the relationships between determinant factors and child nutritional status are likely complex and potentially nonlinear, requiring more flexible modeling approaches capable of capturing simultaneous dynamics (Vatsa et al., 2023).

Previous studies have frequently modeled stunting and wasting separately using parametric approaches such as linear regression, logistic regression, and quantile regression (Asebe et al., 2024; Huq et al., 2023). While these approaches are effective in identifying statistically significant determinants, they generally rely on specific functional form assumptions, such as linear, quadratic, or cubic relationships (Arifin et al., 2023). In practice, however, the relationship between predictor variables and response variables is not always linear and may involve interactions or nonlinear patterns that are difficult to capture optimally using conventional parametric models (Huang & Su, 2021). Therefore, nonparametric approaches provide a relevant alternative, as they do not impose a predetermined functional form on the relationship between predictors and responses.

Nonparametric methods offer greater flexibility in capturing complex relationship patterns. Commonly used approaches include spline methods (Islamiyati et al., 2022), kernel methods (Sukran et al., 2025), and Fourier series approaches (Dani & Adrianingsih, 2021). Among these, spline regression is one of the most popular techniques due to its ability to model nonlinear relationships flexibly through piecewise polynomial functions connected at specific points known as knots (Arifin et al., 2025). The primary advantage of spline methods lies in their ability to produce smooth yet adaptive curves that respond to changes in data patterns. Additionally, spline regression remains relatively interpretable, as it retains a polynomial regression structure within each interval segment.

However, determining the number and placement of knots in classical spline models presents a methodological challenge (Yang et al., 2021). Too few knots may lead to underfitting, while too many knots can result in overfitting and high estimation variance (Arnes et al., 2023). To address this issue, penalized spline approaches have been developed, combining the flexibility of spline functions with a penalty mechanism that controls curve smoothness (Berry & Helwig, 2021). In penalized spline regression, parameters are estimated by incorporating a penalty component into the objective function, typically penalizing either the magnitude of coefficients or the degree of curvature (Gascoigne & Smith, 2023). This approach allows the model to automatically control its complexity through a smoothing parameter, resulting in curves that are more stable, robust, and generalizable.

This study develops a biresponse nonparametric penalized spline regression model that enables flexible estimation of regression functions for each response variable while accounting for the correlation structure between responses. Based on this framework, the study aims to develop and implement a biresponse nonparametric penalized spline regression model to analyze stunting and wasting in Kalimantan simultaneously. This approach is expected to provide a more flexible, efficient, and informative modelling framework for identifying nonlinear relationships between socioeconomic determinant factors and the two indicators of child nutritional status.

B. LITERATURE REVIEW

1. Nonparametric Spline Regression

Nonparametric regression is a modeling approach used when the functional form of the relationship between the response variable and predictor variables is unknown (Arifin et al., 2020). Unlike parametric regression, which requires a predetermined functional form, such as a linear or fixed-degree polynomial, nonparametric regression does not impose a rigid functional structure from the outset. This flexibility enables it to adapt more effectively to complex, nonlinear relationship patterns (Hidayat et al., 2025). Therefore, nonparametric regression is particularly suitable when the data exhibit structures that simple parametric models cannot adequately represent.

Suppose there are n observations with one response variable y_i and p predictor variables $x_{1i}, x_{2i}, \dots, x_{pi}$. The nonparametric regression model is given in Equation (1) below.

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{pi}) + \varepsilon_i, i = 1,2, \dots, n \tag{1}$$

where $f(\cdot)$ is an unknown regression function and $\varepsilon_i$represents random error terms assumed to be independent and normally distributed with mean zero and variance σ^2 (Nisa & Budiantara, 2020). To enhance interpretability and avoid high-dimensional complexity, the regression function is commonly assumed to take an additive form as shown in Equation (2) below.

$$y_i = \sum_{j=1}^{p} f_j\,(x_{ji}) + \varepsilon_i \tag{2}$$

where $f_j(\cdot)$ is a smooth univariate function. One of the most popular approaches in nonparametric regression is spline regression. A spline is a piecewise polynomial function connected at specific points known as knots (Husain et al., 2024). A spline function of order $q$with $r$knots $k_1, k_2, \dots, k_r$ can be written as follows (Equation 3) below.

$$(x_{ji}) = \sum_{l=0}^{q} \beta_l\, x_{ji}^l + \sum_{h=1}^{r} \beta_{q+h}\,(x_{ji} - k_h)_+^q \tag{3}$$

where

$$(x_{ji} - k_h)_+^q = \begin{cases} (x_{ji} - k_h)^q, & x_{ji} \geq k_h \\ 0, & x_{ji} < k_h \end{cases}$$

Equation (3) can be expressed in matrix form as shown in Equation (4) below.

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{4}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{p1}^q & (x_{11}-k_{11})_+^q & \cdots & (x_{p1}-k_{pr})_+^q \\ 1 & x_{12} & \cdots & x_{p2}^q & (x_{12}-k_{11})_+^q & \cdots & (x_{p2}-k_{pr})_+^q \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn}^q & (x_{1n}-k_{11})_+^q & \cdots & (x_{pn}-k_{pr})_+^q \end{bmatrix} \begin{bmatrix} \beta_{0j} \\ \beta_{p1} \\ \vdots \\ \beta_{pq} \\ \beta_{1(q+1)} \\ \vdots \\ \beta_{p(q+r)} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

The parameter estimates in Equation (4) can be obtained using the Least Squares method, yielding the estimator in Equation (5) below.

$$\widehat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \mathbf{y} \tag{5}$$

However, when the number of knots is relatively large, classical spline models may produce curves that excessively follow data fluctuations, leading to overfitting. To address this limitation, the penalized spline approach was developed (Chamidah et al., 2022), which incorporates a penalty term to control model smoothness and complexity.

2. Bivariate Nonparametric Penalized Spline Regression

Suppose there are two response variables analyzed simultaneously, namely Y_{1i} representing the prevalence of stunting and Y_{2i} representing the prevalence of wasting for the i-th observation, where $i = 1,2,\dots,n$. The bivariate nonparametric regression model is expressed in Equation (6) below.

$$Y_{1i} = \sum_{j=1}^{p} f_{1j}\left(X_{ji}\right) + \varepsilon_{1i};\ Y_{2i} = \sum_{j=1}^{p} f_{2j}\left(X_{ji}\right) + \varepsilon_{2i} \tag{6}$$

with $E(\varepsilon_{ki}) = 0$, for $k = 1,2$, and the error covariance structure given in Equation (7) below.

$$\text{Cov}\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \tag{7}$$

This structure allows for correlation between the two response variables. The penalized spline model is applied to each response variable, resulting in Equation (8) below.

$$\begin{aligned} \mathbf{Y}_1 &= X\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1 \\ \mathbf{Y}_2 &= X\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2 \end{aligned} \tag{8}$$

Both equations can be combined into the matrix form shown in Equation (9) below.

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} = \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} + \boldsymbol{\varepsilon} \tag{9}$$

Alternatively, more compactly written as in Equation (10) below.

$$\mathbf{Y} = Z\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{10}$$

To control model complexity in nonparametric regression, a penalty function is imposed on the spline parameters (Witte et al., 2024). The Penalized Least Squares objective function is defined in Equation (11) below.

$$Q(\boldsymbol{\beta}) = (\mathbf{Y} - Z\boldsymbol{\beta})^T(\mathbf{Y} - Z\boldsymbol{\beta}) + \boldsymbol{\beta}^T\Lambda\boldsymbol{\beta} \tag{11}$$

where:

$$\Lambda = \begin{pmatrix} \lambda_1 P & 0 \\ 0 & \lambda_2 P \end{pmatrix}$$

Here, P is a second-order penalty matrix, while λ_1 and λ_2 are smoothing parameters corresponding to each response variable. To obtain the parameter estimates, Equation (11) is differentiated with respect to β as shown in Equation (12) below.

$$\frac{\partial Q}{\partial \boldsymbol{\beta}} = -2Z^T(\mathbf{Y} - Z\boldsymbol{\beta}) + 2\Lambda\boldsymbol{\beta} \tag{12}$$

Setting Equation (12) equal to zero yields the estimation equation in Equation (13) below.

$$Z^T Z\boldsymbol{\beta} + \Lambda\boldsymbol{\beta} = Z^T\mathbf{Y} \tag{13}$$

Thus, the estimator for the bivariate nonparametric penalized spline regression model is given in Equation (14) below.

$$\hat{\boldsymbol{\beta}} = (Z^T Z + \Lambda)^{-1} Z^T \mathbf{Y} \tag{14}$$

The smoothing parameters λ_1 and λ_2 are determined using the Generalized Cross Validation (GCV) method by selecting the values that minimize the GCV function.

3. Best Model Selection

In the bivariate nonparametric penalized spline approach, model flexibility is determined not only by the number of knot points but also by the smoothing parameters λ. Therefore, model selection is conducted through the simultaneous optimization of the smoothing parameters and the determination of knot locations. The selection of knot points for the bivariate penalized spline model using the Generalized Cross Validation (GCV) function is expressed in Equation (15) below.

$$GCV(\lambda_1, \lambda_2) = \frac{\frac{1}{2n} \parallel \mathbf{Y} - \hat{\mathbf{Y}} \parallel^2}{\left[1 - \frac{1}{2n}\text{trace}(A_\lambda)\right]^2} \tag{15}$$

with

$$A_\lambda = Z(Z^T Z + \Lambda)^{-1} Z^T$$

where Y is the combined response vector of dimension $2n \times 1$, Z is the block-diagonal design matrix, Λ is the penalty matrix, A_λ is the penalized spline smoother matrix, and λ_1 and λ_2 are the smoothing parameters corresponding to each response variable the optimal values of λ_1 and λ_2 are obtained by minimizing the GCV function.

C. METHODS

This study is a quantitative study based secondary data obtained from the Indonesian Nutritional Status Survey (SSGI) conducted by the Ministry of Health and official publications from Statistics Indonesia (BPS). ..The unit of analysis in this study is districts/cities in Kalimantan Island, with a total of 56 observations. The variables used in this study consist of two response variables: Stunting Prevalence (Y_1) and Wasting Prevalence (Y_2). The predictor variables include the percentage of households with access to proper sanitation (X_1), low-birth-weight (X_2), and the percentage of the population covered by health insurance (X_3). The model used in this study is a bivariate nonparametric penalized spline regression that allows simultaneous modeling of two response variables. The analysis was conducted using RStudio version 25.0.0. The analytical steps are as follows:

1. Correlation Test Between Response Variables
   A correlation test between the two response variables, namely stunting prevalence (Y_1) and wasting prevalence (Y_2), was performed to ensure that they are statistically related. This step is important because a bivariate model is generally more appropriate when the responses are not independent.
2. Exploratory Data Analysis
   Descriptive analysis was conducted for all research variables, including minimum value, maximum value, mean, and standard deviation. In addition, scatter plots between each response variable and predictor variable were created to detect potential nonlinear relationship patterns.
3. Determination of Knot Points and Smoothing Parameters
   The selection of knot points for each predictor variable was carried out by testing several candidate knot locations. For each candidate, the smoothing parameter $\lambda$was chosen based on the smallest GCV value. The combination of knot points and λthat produced the minimum GCV value was selected as the best model.
4. Construction of the Penalized Spline Model
   Based on the selected knot points, the relationships between the predictors and each response variable (stunting and wasting) were modeled using a bivariate penalized spline approach. Model parameters were estimated via penalized least squares, yielding spline functions that are flexible yet smooth for modeling both responses simultaneously.
5. Model Goodness-of-Fit Evaluation
   Model evaluation was conducted using the coefficient of determination (R^2) and Mean Squared Error (MSE) for each response variable. The R^2value indicates the proportion of variation in stunting and wasting explained by the predictor variables, while MSE measures the average squared prediction error.

D. RESULT AND DISCUSSION

1. Correlation Test Between Response Variables

To assess the correlation among the response variables, the Pearson correlation coefficient was computed. This test was conducted as a preliminary step to ensure that the assumption of interdependence between responses is satisfied, thereby justifying the use of a bivariate model. The hypotheses tested were $H_0: \rho = 0$ (No correlation between response variables) against $H_1: \rho \neq 0$ (There is a correlation between response variables). The test of the response variables yielded a test statistic of t = 2.65 with df = 54 and a p-value of 0.012. At the 5% significance level, the null hypothesis is rejected. Therefore, there is a statistically significant linear correlation between stunting prevalence and wasting prevalence. Thus, the two response variables are not independent and are appropriately modeled simultaneously using a bivariate regression model.

2. Exploratory Data Analysis

An exploratory analysis was conducted to describe the characteristics of the data and the variability of each research variable. The descriptive statistics calculated include the number of observations (N), minimum, maximum, mean, and standard deviation (SD) for all response and predictor variables. The summary of descriptive statistics is presented in Table 1.

Table 1. Exploratory Data Analysis

Variabel	N	Min	Max	Mean	SD
Stunting (y1)	56	7.40	36.40	22.48	6.13
Wasting (y2)	56	1.10	13.30	7.68	2.34
Sanitation (x1)	56	57.71	97.70	82.83	10.02
Low-birth-weight (x2)	56	20.00	1404.00	396.89	373.69
Health Insurance (x3)	56	41.17	100.00	72.23	16.16

Based on Table 1, the total number of observations for all variables is 56. Stunting prevalence (y_1) ranges from 7.40 to 36.40, with a mean of 22.48 and a standard deviation of 6.13, indicating substantial variation in stunting across observations. Wasting prevalence (y_2) ranges from 1.10 to 13.30, with a mean of 7.68 and a standard deviation of 2.34, suggesting relatively lower variability compared to stunting. For the predictor variables, sanitation (x_1) has a mean of 82.83, a standard deviation of 10.02, and ranges from 57.71 to 97.70, indicating considerable differences in sanitation coverage across observations. The low-birth-weight variable (x_2) exhibits the highest variability, with a range of 20.00 to 1404.00, a mean of 396.89, and a standard deviation of 373.69, reflecting strong heterogeneity in this indicator. Meanwhile, health insurance coverage (x_3) ranges from 41.17 to 100.00, with a mean of 72.23 and a standard deviation of 16.16, indicating substantial variation in health insurance coverage across observations. Subsequently, scatter plots were generated between each response variable and each predictor variable to examine relationship patterns and detect potential nonlinearity. The scatter plots are presented in Figure 1 and serve as the basis for applying the bivariate nonparametric penalized spline regression model.

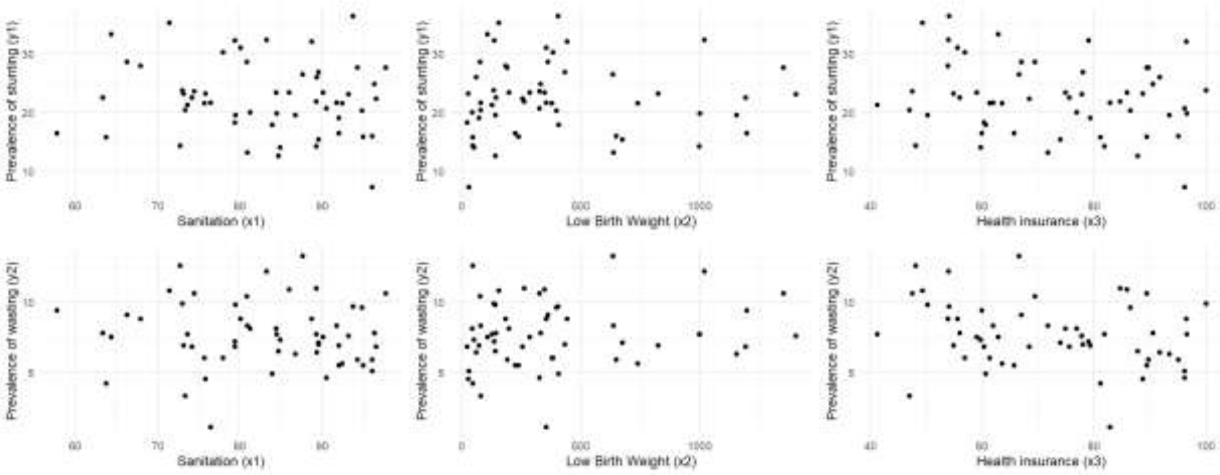

Figure 1. Scatter Plots of Stunting and Wasting Prevalence Against the Predictor Variables

3. Selection of Knot Points and Optimal Lambda

In the bivariate nonparametric penalized spline regression model, determining the number and locations of knots, as well as selecting the smoothing parameter lambda, are crucial components. These selections were performed simultaneously through a model selection procedure based on the minimum GCV criterion. The results are presented in Table 2.

Table 2. Selection of Knot Points and Optimal Lambda

Number of Knots	Knot Points			Lambda	GCV
	X1	X2	X3		
1	66	107	53	63.09	20.83
	79	143	76	15848.93	21.66
	87	431	61	251188.60	22.65
2	67; 75	107; 560	61; 88	6309.57	21.91
	71; 84	135; 823	53; 77	10000.00	21.95
	74; 92	285; 441	46; 81	398107.20	22.57
3	64; 79; 88	190; 403; 674	59; 77; 96	39810.72	21.66
	67; 84; 96	141; 396; 738	60; 76; 91	39810.72	21.63
	71; 83; 94	155; 354; 998	61; 85; 93	63095.73	22.08

Based on Table 2, the minimum GCV value is obtained for the model with one knot, where the knot locations for each predictor are $X_1 = 66$, $X_2 = 107$, and $X_3 = 53$, with a smoothing parameter $\lambda = 63.09$and a GCV value of 20.83. These results indicate that this model provides the best performance among all knot configurations considered. Therefore, nonparametric bivariate penalized spline regression is subsequently implemented using a one-knot model.

4. Bivariate Penalized Spline Model

The bivariate nonparametric penalized spline model was estimated using one knot and the optimal smoothing parameter λ selected in the previous stage. Parameter estimation was performed for each response variable, namely stunting (y_1) and wasting (y_2), resulting in regression coefficients that represent the contribution of the spline components for each predictor. The estimated model parameters are presented in Table 3.

Table 3. Estimation of the Bivariate Nonparametric Penalized Spline Model

Parameter	Stunting (y1)	Wasting (y2)
β_{0i}	-43.032	9.804
β_{11}	0.578	0.009
β_{12}	-0.677	0.004
β_{21}	0.100	0.010
β_{22}	-0.103	-0.009
β_{31}	0.416	-0.063
β_{32}	-0.490	0.022

Based on Table 3, the estimated bivariate nonparametric penalized spline regression models are given by the following equations:

$$\begin{aligned}\hat{y}_{1i} &= -43.032 + 0.578\,x_{1i} - 0.677\,(x_{1i} - 66)_+ + 0.100\,x_{2i} - 0.103\,(x_{2i} - 107)_+ \\ &\quad + 0.416\,x_{3i} - 0.490\,(x_{3i} - 53)_+ \\ \hat{y}_{2i} &= 9.804 + 0.009\,x_{1i} + 0.004\,(x_{1i} - 66)_+ + 0.010\,x_{2i} - 0.009\,(x_{2i} - 107)_+ \\ &\quad - 0.063\,x_{3i} + 0.022\,(x_{3i} - 53)_+\end{aligned}$$

After constructing the bivariate nonparametric penalized spline regression model, the interpretation of each predictor's effect is carried out through segmentation based on the selected knot points, namely, sanitation ($x_1 = 66$), low-birth-weight ($x_2 = 107$), and health insurance coverage ($x_3 = 53$). This segmentation confirms that the relationship between the response variables and the predictors is not uniform across the entire data range; rather, it varies across segments. Consequently, this approach provides a richer understanding of the dynamics of the factors associated with stunting (y1) and wasting (y2). The segmentation results for the sanitation variable (x_1) are presented in the following equations.

$$f_1^{(1)}(x_{1i}) = \begin{cases} 0.578\,x_{1i}, & x_{1i} \le 66 \\ -0{,}099\,x_{1i} + 44.682, & x_{1i} > 66 \end{cases}$$

$$f_1^{(2)}(x_{1i}) = \begin{cases} 0.009\,x_{1i}, & x_{1i} \le 66 \\ 0.013\,x_{1i} - 0.264, & x_{1i} > 66 \end{cases}$$

For the sanitation variable (x_1), the stunting model shows a clear change in the direction of the effect. In the segment $x_1 \le 66$. The coefficient is positive, indicating that higher sanitation coverage is associated with higher stunting rates within this range. However, after exceeding the knot point ($x_1 > 66$), the slope becomes negative, meaning that higher levels of sanitation are associated with a decrease in stunting. This pattern suggests a strong nonlinear relationship and the possibility of a threshold effect of sanitation on stunting. In contrast, for wasting, the effect of sanitation is relatively small and consistently positive, increasing from 0.009 when $x_1 \le 66$ to 0.013 when $x_1 > 66$. Next, the segmentation results for the low-birth-weight variable (x_2) are presented in the following equations.

$$f_2^{(1)}(x_{2i}) = \begin{cases} 0.100\,x_{2i}, & x_{2i} \le 107 \\ -0.003\,x_{2i} + 11.021, & x_{2i} > 107 \end{cases}$$

$$f_2^{(2)}(x_{2i}) = \begin{cases} 0.010\, x_{2i}, & x_{2i} \leq 107 \\ 0.001\, x_{2i} + 0.963, & x_{2i} > 107 \end{cases}$$

For the low-birth-weight variable (x_2), the stunting model also exhibits a clear change in the direction of the effect. When $x_2 \leq 107$. The positive slope indicates that an increase in low-birth-weight cases is associated with an increase in stunting. However, in the segment $x_2 > 107$, the slope approaches zero and is slightly negative, suggesting that beyond this threshold, further increases in low birth weight are no longer strongly associated with changes in stunting. For wasting, the direction of the low-birth-weight effect remains positive, but the magnitude is much smaller and weakens after the knot point. Finally, the segmentation results for the health insurance coverage variable (x_3) are presented in the following equations.

$$f_3^{(1)}(x_{3i}) = \begin{cases} 0.416\, x_{3i}, & x_{3i} \leq 53 \\ -0.074\, x_{3i} + 25.970, & x_{3i} > 53 \end{cases}$$

$$f_3^{(2)}(x_{3i}) = \begin{cases} -0.063\, x_{3i}, & x_{3i} \leq 53 \\ -0.041\, x_{3i} - 1.166, & x_{3i} > 53 \end{cases}$$

For the health insurance coverage variable (x3), the pattern differs between stunting and wasting. In the stunting model, before the knot point ($x_3 \leq 53$), the relatively large positive slope indicates that increased health insurance coverage is associated with higher stunting prevalence within this range. However, after exceeding the knot ($x_3 > 53$), the direction of the effect reverses and becomes negative, suggesting that higher levels of health insurance coverage are associated with a reduction in stunting. In contrast, for wasting, the relationship with health insurance coverage tends to be negative overall. Based on these results, the three predictors exhibit different relationships with stunting and wasting. Stunting tends to show stronger nonlinearity, particularly for sanitation and health insurance coverage, both of which demonstrate changes in the direction of effect after passing the knot points. Conversely, waste is generally associated with smaller effect magnitudes and more stable patterns, except for health insurance coverage, which remains negative but shows a flattening slope beyond the threshold.

5. Model Goodness-of-Fit Evaluation

The performance of the bivariate nonparametric penalized spline regression model was evaluated separately for each response variable, namely stunting (y_1) and wasting (y_2). This evaluation was conducted to assess the model's ability to explain data variability and its predictive accuracy. Two evaluation measures were used: Mean Squared Error (MSE) and the coefficient of determination (R^2). MSE represents the average squared prediction error, while R^2indicates the proportion of response variability explained by the predictors included in the model. The evaluation results of the bivariate nonparametric penalized spline regression model are presented in Table 3.

Table 3. Evaluation Results of the Bivariate Nonparametric Penalized Spline Model

Evaluasi Model	Stunting (y1)	Wasting (y2)
MSE	28.012	4.810
R^2	0.241	0.106

Based on Table 3, for the stunting response (y_1), the model yields an MSE of 28.012 and an R^2 value of 0.241. This R^2 indicates that the model explains approximately 24.1% of the variation in stunting, while the remaining variation is influenced by other factors not included among the predictors or by random variability. For the wasting response (y_2), the model produces an MSE of 4.810 and an R^2 value of 0.106, indicating that approximately 10.6% of the variation in wasting is explained by the model. Compared with wasting, the model demonstrates better explanatory performance for stunting. However, both R^2 values suggest that stunting and wasting are likely influenced by additional determinants that are not yet incorporated into the model

E. CONCLUSION AND SUGGESTIONS

This study shows that the bivariate nonparametric penalized spline regression model is capable of capturing the nonlinear relationship between the predictors and the two nutritional indicators, namely stunting and wasting prevalence. The findings suggest that proper sanitation, low birth weight, and health insurance coverage are important factors associated with variations in stunting and wasting across districts/cities in Kalimantan. These results have practical implications for policy and intervention design. Efforts to reduce stunting and wasting should not rely on a single strategy, but rather on integrated policies that improve household sanitation, strengthen maternal and child health services to prevent low birth weight, and expand effective health insurance coverage to ensure better access to nutrition and health care services. Therefore, local governments and public health authorities in Kalimantan may use these findings as evidence to support more targeted and context-specific nutrition intervention programs

For future research, it is recommended that broader explanatory variables, such as socioeconomic conditions, access to health services, maternal education, poverty levels, and environmental or regional factors, be incorporated to explain variations in stunting and wasting more comprehensively. In addition, incorporating spatial effects across Kalimantan may provide deeper insights, given that nutritional indicators often exhibit geographic patterns. Finally, applying alternative nonparametric approaches, such as kernel methods or different spline bases, is suggested to assess the robustness and consistency of the findings.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Kementerian Kesehatan Republik Indonesia and the Badan Pusat Statistik for providing access to the Indonesian Nutritional Status Survey (SSGI) data and official statistical publications, which served as the primary data sources for this study. They also extend their appreciation to Lambung Mangkurat University for the academic support and research facilities provided throughout the preparation and analysis of this research. Their support and contributions were invaluable to the successful completion of this study.

REFERENCES

Arifin, S., Anggraini, D., Salam, N., Arifin, S., Anggraini, D., & Salam, N. (2025). A Comparative Study of Linear and Quadratic Spline Regression Models for Predicting HbA1c Levels in Patients with Diabetes Mellitus. *Jambura: Journal of Mathematics*, *7*(2), 183–188. https://doi.org/10.37905/jjom.v7i2.33292

Arifin, S., Islamiyati, A., & Herdiani, E. T. (2023). Ability of Ordinal Spline Logistic Regression Model in the Classification of Nutritional Status Data. *Communications in Mathematical Biology and Neuroscience*, *2023*(1), 1–11. https://doi.org/10.28919/cmbn/8072

Arifin, S., Islamiyati, A., & Raupong, R. (2020). Kemampuan Estimator Spline Linear dalam Analisis Komponen Utama. *ESTIMASI: Journal of Statistics and Its Application*, *1*(1), 40. https://doi.org/10.20956/ejsa.v1i1.9262

Arnes, J. I., Hapfelmeier, A., Horsch, A., & Braaten, T. (2023). Greedy knot selection algorithm for restricted cubic spline regression. *Frontiers in Epidemiology*, *3*(12), 1–14. https://doi.org/10.3389/fepid.2023.1283705

Asebe, H. A., Asmare, Z. A., Mare, K. U., Kase, B. F., Tebeje, T. M., Asgedom, Y. S., Shibeshi, A. H., Lombebo, A. A., Sabo, K. G., Fente, B. M., Bezie, M. M., & Seifu, B. L. (2024). The level of wasting and associated factors among children aged 6–59 months in sub-Saharan African countries: Multilevel ordinal logistic regression analysis. *Frontiers in Nutrition*, *11*(June), 1–14. https://doi.org/10.3389/fnut.2024.1336864

Berry, L. N., & Helwig, N. E. (2021). Cross-Validation, Information Theory, or Maximum Likelihood? A Comparison of Tuning Methods for Penalized Splines. *Stats*, *2021*(4), 701–724. https://doi.org/10.3390/stats4030042

Chamidah, N., Rifada, M., & Amelia, D. (2022). A Theoretical Discussion on Modeling the Number of Covid-19 Death Cases Using Penalized Spline Negative Binomial Regression Approach. *Communications in Mathematical Biology and Neuroscience*, *2022*(82), 1–14. https://doi.org/10.28919/cmbn/7518

Dani, A. T. R., & Adrianingsih, N. Y. (2021). Pemodelan Regresi Nonparametrik dengan Estimator Spline Truncated dan Deret Fourier. *Jambura: Journal of Mathematics*, *3*(1), 26–36. https://doi.org/10.34312/jjom.v3i1.7713

Eryando, T., Sipahutar, T., Budhiharsana, M. P., Siregar, K. N., Aidi, M. N., Utari, D. M., & Rahmaniati, M. (2022). Spatial analysis of stunting determinants in 514 Indonesian districts/cities: Implications for intervention and setting of priority. *Geospatial Health*, *17*(1055), 1–7. https://doi.org/10.4081/gh.2022.1055

Gascoigne, C., & Smith, T. (2023). Penalized smoothing splines resolve the curvature identifiability problem in age-period-cohort models with unequal intervals. *Statistics in Medicine*, *42*(February), 1888–1908. https://doi.org/10.1002/sim.9703

Hidayat, R., Muthahharah, I., Resmawan, R., Fleksibel, E., Ilm, E. J., & Sains, M. (2025). Kernel-Truncated Spline : Estimator Fleksibel untuk Regresi Nonparametrik Kernel-Truncated Spline : Estimator Fleksibel untuk Regresi Nonparametrik. *EULER: Jurnal Ilmiah Matematika, Sains dan Teknologi*, *13*(2), 271–277. https://doi.org/10.37905/euler.v13i2.33062

Huang, J. Z., & Su, Y. (2021). Asymptotic Properties of Penalized Spline Estimators in Concave Extended Linear Models: Rates of Convergence. *The Annals of Statistics*, *49*(6), 3383–3407. https://doi.org/10.1214/21-AOS2088

Huq, M., Farjana, M., Islam, A., & Khan, M. (2023). Children and Youth Services Review Factors associated with stunting status among under-5 years children in Bangladesh: Quantile regression modelling approach. *Children and Youth Services Review*, *155*(February), 107199. https://doi.org/10.1016/j.childyouth.2023.107199

Husain, H., Dewi, A. F., & Wardani, A. E. (2024). Modeling of Indonesian Stunting Prevalence Using Spline Truncated Non- parametric Regression. *Journal of Analytical Research, Statistics and Computation*, *3*(1), 1–13. https://doi.org/10.4590/jarsic.v3i1.26

Islamiyati, A., Raupong, Kalondeng, A., & Sari, U. (2022). Estimating the confidence interval of the regression coefficient of the blood sugar model through a multivariable linear spline with known variance. *STATISTICS IN TRANSITION new series*, *23*(1), 201–212. https://doi.org/10.2478/stattrans-2022-0012

Jokhu, L. A., & Syauqy, A. (2024). Determinants of concurrent wasting and stunting among children 6 to 23 mo in Indonesia. *Nutrition*, *122*. https://doi.org/10.1016/j.nut.2024.112390

Kustanto, A., Rachmat, O., & Setyadi, S. (2025). The Prevalence of Stunting in Indonesia: An Examination of the Health, Socioeconomic Status, and Environmental Determinants. *Journal of Iranian Medical Council*, *8*(1), 67–79. https://doi.org/10.18502/jimc.v8i1.17062

Nisa, K., & Budiantara, I. N. (2020). Modeling east java indonesia life expectancy using semiparametric regression mixed spline truncated and fourier series 1. *Media Statistika*, *13*(2), 149–160. https://doi.org/10.14710/medstat.13.2.149-160

Rahmadiani, I., Fibriana, A. I., Nisa, A. A., Abdul, S., Shabbir, & Azam, M. (2025). Impact of Low Birth Weight and Other Determinants on Stunting in Children Under - Five Years Old: Evidence from Indonesia’s Nutrition Status Survey. *International Journal of Preventive Medicine*, *16*(76), 1–8. https://doi.org/10.4103/ijpvm.ijpvm

Sukran, A. M., Budiantara, I. N., & Ratnasari, V. (2025). The Curve Estimation Nonparametric Regression Multiresponse Mixed with Truncated Spline, Fourier Series, and Kernel. *Mandalika Mathematics and Education Journal*, *7*(2), 766–787. https://doi.org/10.29303/jm.v7i2.9188

Vatsa, R., Id, U. G., Yasmin, K., & Hasan, F. J. (2023). Determinants of undernutrition and overnutrition among reproductive-age women in Bangladesh : Trend analysis using spatial modeling. *PLOS ONE*, 1–17. https://doi.org/10.1371/journal.pone.0282998

Witte, D. De, Abad, A. A., Neyens, T., Verbeke, G., & Molenberghs, G. (2024). A joint penalized spline smoothing model for the number of positive and negative COVID- 19 tests. *PLOS ONE*, *19*(5), 1–21. https://doi.org/10.1371/journal.pone.0303254

Wulandari, D. T., Azizah, N., Rahmawati, V. E., & Annuchasari, H. (2025). Community-Based Insights on Maternal Nutrition Knowledge and Its Role in Preventing Stunting and Wasting in Rural Indonesian Children. *Journal of Current Health Sciences*, *5*(3), 145–154. https://doi.org/10.47679/jchs.2025128

Yang, G., Zhang, B., Zhang, M., Yang, G., Zhang, B., & Zhang, M. (2021). Estimation of Knots in Linear Spline Models. *Journal of the American Statistical Association*, *0*(0), 1–12. https://doi.org/10.1080/01621459.2021.1947307