Estimation of Stunting and Wasting in Sumatra 2022 with Nadaraya-Watson Kernel and Penalized Spline

Cinta Rizki Oktarina[1], Sigit Nugroho[2*], Idhia Sriliana[3], Pepi Novianti[4], Etis Sunandi[5], and Reza Pahlepi[6]

[1] Department of Mathematics, Faculty of Information Technology, Institut Teknologi Batam, Indonesia
[2,3,4,5,6] Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Bengkulu, Bengkulu, Indonesia
*Corresponding author: snugroho@unib.ac.id

Received: 7 May 2025
Revised: 29 October 2025
Accepted: 31 October 2025

ABSTRACT **This study aims to estimate the prevalence of Stunting and Wasting in Sumatra in 2022 using nonparametric regression methods, specifically the Nadaraya-Watson Kernel and Penalized Spline regression models. Both models were applied to assess the relationship between these two correlated response variables and various predictor variables, such as low birth weight, sanitary facilities, poor population, and exclusive breastfeeding. The results showed that the Nadaraya-Watson Kernel regression, particularly using the Gaussian kernel, provided the best fit with minimal prediction error, as indicated by its low Generalized Cross-Validation (GCV) value of 0.024 and high R-squared values (0.9992 for Stunting and 0.9995 for Wasting). In contrast, the Epanechnikov kernel and Biweight kernel produced higher GCV values (0.110 and 0.356, respectively), indicating less optimal performance. For the Penalized Spline model, optimal parameters were determined with a smoothing parameter λ of 5 and 3 knots, which balanced model flexibility and smoothness. This research underscores the potential of nonparametric regression techniques in capturing complex relationships in health data and provides insights for improving interventions aimed at addressing child malnutrition in Indonesia.**

Keywords– **Kernel, Nonparametric Birespon Regression, Penalized Spline, Stunting, Wasting.**

I. INTRODUCTION

Regression analysis is a method used to explain how one or more response variables depend on one or more predictor variables. There are three approaches to estimating the regression curve: parametric, nonparametric, and semiparametric regression. If the pattern of the relationship between the predictor and response variables is known, parametric regression modeling can be applied [1]. However, in practice, not all data follow specific patterns. When the relationship between the predictor and response variables is unknown, nonparametric regression is the appropriate model for modeling the relationship between these variables [1]. Nonparametric regression analysis is not only for uniresponse but also for bivariate and multivariate responses. Bivariate analysis involves two correlated response variables [2]. Functions used in nonparametric regression include spline, kernel, local polynomial, Fourier series, wavelets, and MARS [3]. Kernel regression, or local averaging regression, is often applied when data points are unevenly spaced or predictor variables are random. The kernel estimator estimates the function without imposing linear or parametric assumptions [1]. This method is flexible, computationally easy, and converges quickly [4]. Nadaraya-Watson Kernel estimation is one approach with high flexibility in modeling variable relationships [5]. Compared to spline regression, kernel offers advantages in flexibility and adaptability.

Spline is a model that offers both statistical and visual interpretations that are highly specific and effective[1]. Spline regression involves polynomial functions that are segmented and continuous [6], [7]. It uses connecting points called knots, providing flexibility in capturing complex data patterns [8]. In spline regression, besides the location and number of knots, another key consideration is finding the optimal value of λ, with its use in nonparametric regression known as Penalized Spline regression. Research on kernel and spline has been widely applied. [9]conducted a study on forecasting regional $PM_{2.5}$ concentration using a new model based on empirical orthogonal function analysis and the Nadaraya-Watson Kernel regression estimator. The results showed that the average prediction accuracy of the model was 74.38%, with more than 92% of cumulative variance and varying bandwidth values for each season. A subsequent study by [10]focused on outlier identification using Penalized Spline regression to model the poverty depth index as a response variable. The study achieved an R-square value of 69.10% with optimal knots for each predictor variable being 1, 2, 4, 1, 5, 3, and 1, respectively.

The development of nonparametric regression methods has become increasingly popular in statistics due to their ability to capture complex relationships between variables without requiring a specific pattern of relationship. Among these methods, the bivariate approach with kernel and spline has advantages in providing more flexible estimation, especially in the analysis of data with two correlated response variables. For example, in children's health studies, Stunting and Wasting often occur simultaneously and reflect poor nutritional conditions, making simultaneous analysis of these two indicators crucial. In general, malnutrition in toddlers is classified into Wasting (low weight-for-height), Stunting (low height-for-age), and underweight (low weight-for-age) [11]. According to the Ministry of Health in 2022, Stunting is a growth disorder caused by chronic malnutrition and long-term infections, resulting in toddlers appearing shorter than their age peers. Meanwhile, Wasting is a condition where a toddler's weight continues to decline significantly over time, causing their weight to fall far below the growth curve standards based on height.

II. LITERATURE REVIEW

A. Kernel Nonparametric Regression

Kernel nonparametric regression, also known as local averaging regression, is an approach commonly used in cases where data points are unevenly spaced or when predictor variables are random. This method utilizes a kernel estimator to estimate the regression function without imposing assumptions of linearity or a specific parametric form on the data [1]. Nonparametric regression relies on the weighted average of the response variable, involving weights that represent the distance between the observed predictor variables, measured by bandwidth (h). Kernel nonparametric regression originates from local polynomial regression, which is considered a specific form of polynomial regression of degree 0, known as the local constant approach. In this approach, the regression function is locally approximated by a constant, with the kernel acting as a weight on the data points closest to the estimation point. One of the nonparametric regression estimation techniques is the Nadaraya-Watson Kernel estimator, which is more flexible than other nonparametric techniques [5]. With the following functions:

$$\widehat{\boldsymbol{m}}(\boldsymbol{t_i}) = \frac{\boldsymbol{K}\left(\sum_{g=1}^{G}\frac{\boldsymbol{t_g}-\boldsymbol{t_{gi}}}{\boldsymbol{h_g}}\right)\boldsymbol{y_i}}{\frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{i=1}^{n}\boldsymbol{K}\left(\sum_{g=1}^{G}\frac{\boldsymbol{t_g}-\boldsymbol{t_{gi}}}{\boldsymbol{h_g}}\right)} \tag{1}$$

$$= \boldsymbol{W_h(t_i)y_i} \tag{2}$$

with

$$W_h(t_i) = \frac{K\left(\sum_{g=1}^{G}\frac{t_g-t_{gi}}{h_g}\right)}{\frac{1}{n}\sum_{i=1}^{n}K\left(\sum_{g=1}^{G}\frac{t_g-t_{gi}}{h_g}\right)}$$

The weight $W_h(t_i)$ can be defined in the following matrix form:

$$\boldsymbol{W_h} = \begin{bmatrix} \boldsymbol{W_h(t_1)} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{W_h(t_2)} & \vdots & \vdots \\ \boldsymbol{0} & \vdots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{W_h(t_n)} \end{bmatrix} \tag{3}$$

B. Spline Nonparametric Regression

Spline regression involves polynomial functions that are segmented and continuous[12]. The Penalized Spline nonparametric regression model has high flexibility in estimating the function y, assuming that the function is smooth and defined in the Sobolev space $V_2''(a,b)$ [4]. The use of connecting points or knots in spline regression allows the model to capture significant changes in data patterns. Determining the number and location of knots is crucial in spline regression, with the optimal value of λ playing a role in controlling the smoothness of the function estimate [12]. Penalized Spline regression uses the smoothing parameter λ to avoid overfitting and provide more accurate estimates by capturing smoother data patterns. In general, the spline function with order m and the jth knot for each response can be expressed as follows [13]

$$\boldsymbol{g(t_i)} = \boldsymbol{\delta_0} + \sum_{\boldsymbol{g=1}}^{\boldsymbol{G}}\sum_{\boldsymbol{m=1}}^{\boldsymbol{M}}\left(\boldsymbol{\delta_{mg}t_{gi}^{m}} + \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}}\boldsymbol{\phi_{gk}}\left(\boldsymbol{t_{gi}}-\boldsymbol{\xi_{gk}}\right)_{+}^{\boldsymbol{m}}\right) \tag{4}$$

C. Weighting Matrix

Based on the initial concept of bivariate regression that must have a significant relationship between response variables, that relationship can be measured using correlation analysis. One method that can be used is Pearson correlation analysis. Pearson correlation is denoted by $\hat{\rho}$, which will always be within the interval $-1 \leq \hat{\rho} \leq 1$and can be calculated using the equation (5) as follows:

$$\widehat{\boldsymbol{\rho}} = \frac{\boldsymbol{s_{y^{(1)}y^{(2)}}}}{\boldsymbol{s_{y^{(1)}}s_{y^{(2)}}}} \tag{5}$$

The hypothesis testing stages for Pearson correlation are as follows:

a) Hypothesis:
   $H_0: \rho = 0$
   $H_1: \rho \neq 0$
b) Required quantities
   Significance level, number of observations, degrees of freedom, table statistic
c) Test statistic [14]:

$$\mathbf{t-test} = \left|\frac{\widehat{\boldsymbol{\rho}}\sqrt{\boldsymbol{n-2}}}{\sqrt{\boldsymbol{1-\widehat{\rho}^2}}}\right| \tag{6}$$

d) Decision criteria
   Reject H_0 if $t_{hitung} \geq t_{\frac{\alpha}{2};n-2}$ or P-value $\leq \alpha$
e) Conclusion

Based on the test, a conclusion is drawn regarding the relationship between response variables, which is a key condition for performing bivariate regression. Weight matrices play an important role in determining parameter estimates in bivariate nonparametric regression models. The advantage of involving weight matrices is their ability to address correlation between responses within the same observation. In nonparametric regression with two responses, there is correlation between errors in the first response and errors in the second response. The covariance matrix for each observation can be represented as follows [15]:

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{s}^2_{y^{(1)}}\boldsymbol{I} & (\boldsymbol{s}_{y^{(1)}y^{(2)}})\boldsymbol{I} \\ (\boldsymbol{s}_{y^{(2)}y^{(1)}})\boldsymbol{I} & \boldsymbol{s}^2_{y^{(2)}}\boldsymbol{I} \end{bmatrix}^{-1} \tag{7}$$

$$\boldsymbol{W^{-1}} = \begin{bmatrix} \boldsymbol{z}^2_{y^{(1)}} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{z}_{y^{(1)}y^{(2)}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{z}^2_{y^{(1)}} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{z}_{y^{(1)}y^{(2)}} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{z}^2_{y^{(1)}} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{z}_{y^{(1)}y^{(2)}} \\ \boldsymbol{z}_{y^{(2)}y^{(1)}} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{z}^2_{y^{(2)}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{z}_{y^{(2)}y^{(1)}} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{z}^2_{y^{(2)}} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{z}_{y^{(2)}y^{(1)}} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{z}^2_{y^{(2)}} \end{bmatrix} \tag{8}$$

D. Nadaraya-Watson Kernel Birespon Nonparametric Regression

Bivariate nonparametric regression Kernel Nadaraya-Watson is a statistical approach used to model the relationship between two response variables and one or more predictor variables without requiring specific distributional assumptions for the data. Kernel not only functions to smooth the relationship between predictor and response variables but also involves weight functions designed to provide different contributions to each observation. This weighting reflects the level of impact that each observation has on the regression estimator, based on its proximity to the prediction point. In this approach, two response variables are analyzed simultaneously, particularly when the correlation between the two needs to be considered, such as in health, economic, or social analyses [16]. Therefore, an additional weight function is used to optimally capture the relationship between the response variables and provide the best contribution to the regression estimator. By involving these two weight functions, the bivariate nonparametric regression estimator Kernel Nadaraya-Watson is capable of capturing random relationships between the predictor variables and the two response variables. So that the Nadaraya-Watson Kernel Birespon Nonparametric Regression estimation function can be written as follows [17].

$$\widehat{\boldsymbol{\beta}}\,(\boldsymbol{t_0}) = \sum_{i=1}^{n} \frac{\left[\left(\boldsymbol{z}^2_{y^{(1)}} + 2\boldsymbol{z}_{y^{(2)}y^{(1)}} + \boldsymbol{z}^2_{y^{(2)}}\right)\left(\boldsymbol{K}\left(\sum_{g=1}^{G}\left(\frac{\boldsymbol{t_g} - \boldsymbol{t_{gi}}}{\boldsymbol{h_g}}\right)\right)\boldsymbol{y}_i^{(r)}\right)\right]}{\frac{1}{n}\sum_{i=1}^{n}\left[\left(\boldsymbol{z}^2_{y^{(1)}} + 2\boldsymbol{z}_{y^{(2)}y^{(1)}} + \boldsymbol{z}^2_{y^{(2)}}\right)\left(\boldsymbol{K}\left(\sum_{g=1}^{G}\left(\frac{\boldsymbol{t_g} - \boldsymbol{t_{gi}}}{\boldsymbol{h_g}}\right)\right)\right)\right]} \tag{9}$$

To determine the optimal kernel function and bandwidth, the common approach is minimizing the Generalized Cross-Validation (GCV). The advantage of GCV lies in its asymptotic optimality, making it effective in various data conditions [18]. The bandwidth parameter plays a key role in adjusting the smoothness of the kernel estimate. As the bandwidth increases, the estimate becomes smoother but may increase bias and cause underfitting. Conversely, reducing the bandwidth increases fluctuation in estimates but reduces bias and may lead to overfitting. The optimal bandwidth can be defined as follows [18]:

$$\boldsymbol{GCV}\left(\boldsymbol{h_{opt}}\right) = \frac{\boldsymbol{MSE}(\boldsymbol{h_{opt}})}{\left(\boldsymbol{1} - \boldsymbol{2n^{-1}tr}(\boldsymbol{B})\right)^2} \tag{10}$$

E. Spline-penalized Birespon Nonparametric Regression

Nonparametric Regression with Penalized Spline explains the relationship between one or more response variables and one or more predictor variables using the Penalized Spline estimator. This model uses paired data $(t_1, t_2, \ldots, t_G)$. The PWLS estimator, which employs smoothing parameters to control the roughness of the function, can be applied in nonparametric regression models to estimate parameters by incorporating weights in the form of the inverse covariance matrix of the response variable. This model can be expressed by Equation (11).

$$\boldsymbol{y}_i^{(r)} = \boldsymbol{\delta}_0^{(r)} + \sum_{g=1}^{G}\left(\boldsymbol{\delta}_{mg}^{(r)}\boldsymbol{t}_{gi}^{m\,(r)} + \sum_{k=1}^{K_g}\boldsymbol{\phi}_{gk}^{(r)}\left(\boldsymbol{t_{gi}} - \boldsymbol{\xi_{gk}}\right)_+^m\right) + \boldsymbol{\varepsilon}_i^{(r)};\ \boldsymbol{r = 1, 2} \tag{11}$$

After obtaining the estimator $\delta_{mg}^{(r)}$, the next step is to explain the role and location of the knots as well as the smoothing parameter λ in the Penalized Spline model. The knot (ξ_k) is the point where the behavior of a function changes over different intervals. Penalized Spline regression applies knots located at quantile points, which represent unique values of predictor variables once the data is sorted. The location of the knots can be determined using Penalized Spline regression and is expressed as follows [19]:

$$\xi_k = \frac{j}{K+1,}, k = 1,2,3,..,K \tag{12}$$

In determining the optimal smoothing parameter λ as well as the number and location of knots, the commonly used method is Generalized Cross Validation (GCV) which minimizes [20]. The advantage of the GCV method is its asymptotic optimality [18]. The smoothing parameter λ controls the roughness penalty. When the value of λ increases, the function estimate becomes smoother, while a decrease in λ results in a rougher estimate. The GCV method can be defined as follows [13]:

$$GCV(\xi_{opt}, \lambda_{opt}) = \frac{MSE(\xi_{opt}, \lambda_{opt})}{\left(1 - 2n^{-1}tr(\boldsymbol{A})\right)^2} \tag{13}$$

F. Goodness of Evaluation Model

Estimated models provide many benefits for researchers and society in decision-making. To assess how well the model meets its objectives, R-Squared and Root Mean Squared Error (RMSE) are used. R-Squared indicates how well the model explains the variance in the data, ranging from -∞ (worst) to +1 (best). A value closer to 1 shows a strong model. It is calculated as [21]:

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{14}$$

RMSE is used to measure the level of prediction error. It calculates the square root of the average squared differences between observed and predicted values, optimal when errors follow a normal distribution. The RMSE formula is [22]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{15}$$

G. Indicator of Nutritional Status of Toddlers

Nutritional status according to the Indonesian Ministry of Health and WHO is a condition caused by the balance between the intake of nutrients from food and the necessary nutritional needs. Meanwhile, nutritional status indicators are signs that can be recognized to describe a person's nutritional level. Indicators of nutritional status in toddlers are measures or parameters used to assess the nutritional condition of children under the age of five. According to the Regulation of the Minister of Health of the Republic of Indonesia (2020), indicators that can be used for children under 5 years of age are body weight for age (BB/U), body height for age (TB/U), and body weight for height (BB/TB). These three indicators can show whether a person has a nutritional status that is deficient, stunting, wasting, and obese.

Stunting, according to the [23], is a condition of impaired growth and development in children due to malnutrition, repeated infections, and inadequate psychosocial stimulation. Children are considered stunted if their height for age is below two standard deviations of the WHO Child Growth Standards. This condition can lead to cognitive delays, reduced productivity in adulthood, and an increased risk of chronic diseases. In Indonesia, stunting is caused by various factors including family and household conditions, infectious diseases, and poor sanitation [24]. Efforts to prevent stunting include improving nutrition, access to clean water, and healthcare services to ensure better growth and development of children.

Wasting, as defined by UNICEF, is a severe form of malnutrition characterized by low body weight relative to height. This condition is often caused by inadequate nutrition or repeated infections. Risk factors for wasting include lack of exclusive breastfeeding, improper complementary feeding, and poor access to healthcare and sanitation services. Children suffering from wasting are at a higher risk of stunting and cognitive impairment. In Indonesia, malnutrition remains a serious issue, affected by various factors, and efforts must focus on early detection and proper intervention to prevent and manage wasting in children [25].

III. METHODOLOGY

This study covers 10 provinces with 154 districts/cities as observation locations. The variables used include the prevalence of Stunting (Y_1) and Wasting (Y_2), as well as predictor variables such as the percentage of Low Birth Weight (T_1), Sanitary Facilities (T_2), Poor Population (T_3), and Infants Receiving Exclusive Breastfeeding (T_4). All variables are measured in percentage, with data sources from Health Profiles and the Central Bureau of Statistics. The steps in this study using the Kernel Nadaraya-Watson and Penalized Spline estimator are as follows:

1. Collect and determine data on Stunting and Wasting, along with suspected factors based on previous theories and research.
2. Measure the correlation between response variables using Pearson correlation.
3. Visualize data using scatterplots between response and predictor variables to determine the relationship pattern.
4. Estimate the Kernel Nadaraya-Watson function.
   a. Define the kernel function to be used.
   b. Determine the upper and lower bounds, and the increment value of the bandwidth.
   c. Define the local data points.

d. Construct the weight matrix involving the inverse of the variance and covariance values, along with the kernel function.
e. Estimate the function using the kernel function and the optimal bandwidth value obtained in step (a).

5. Estimate the model using the PWLS estimator.
   a. Determine the maximum order to be used.
   b. Set the upper and lower bounds, and the increment value for lambda.
   c. Determine the maximum number of knots to be used.
   d. Determine the order, location of knots, and the number of knots, as well as vary the value of lambda to obtain the optimal lambda value based on the minimum GCV.
   e. Estimate the PWLS model using the order, location of knots, and optimal number of knots and lambda value obtained in step (a).
6. Evaluate model performance using R-squared and RMSE.
7. Segment the model and interpret segmentation results.

IV. RESULTS AND DISCUSSIONS

A. Correlation between Response Variables

The relationship between two response variables can be measured using Pearson correlation, the results of testing the two response variables are as follows:

Table 1 Correlation output of $\boldsymbol{Y}^{(1)}$ and $\boldsymbol{Y}^{(2)}$

$H_0: \rho = 0$	
Statistics	**Value**
$\hat{\rho}$	0.216
t-test	2.735
t-crit	2.264
p-value	0.007

Based on Table 1, the value of t-test = 2.735 > t-crit = 2.264. Therefore, at the 0.05 significance level, we can reject H_0 and conclude that there is a strong and significant correlation of 0.216 between the response variables, namely the prevalence of Stunting and Wasting in Sumatra in 2022. Thus, the assumption of correlation is fulfilled, confirming the significant relationship between Stunting and Wasting prevalence.

B. Visualization of Relationship Pattern of Response Variable and Each Predictor Variable

Scatter plot is used to see and identify the pattern of relationship between response variables and predictor variables. The purpose of this analysis is to provide an initial picture of the pattern of the relationship between the predictor variables and the response variable seen before proceeding to further modeling stages.

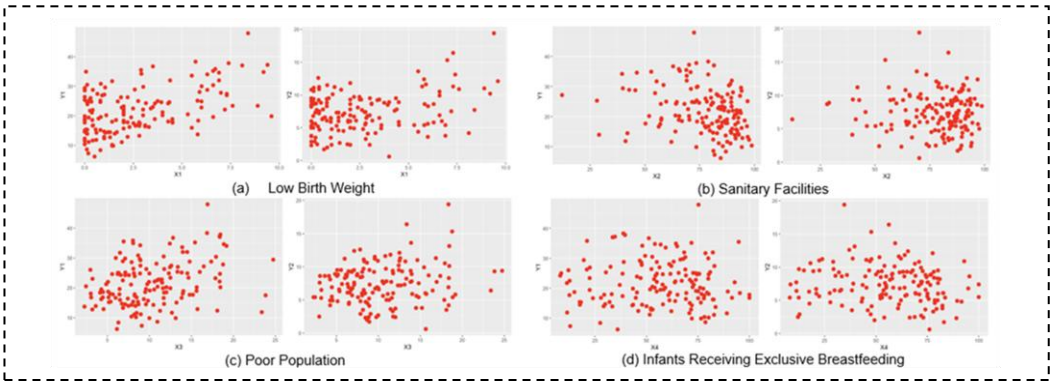

Figure 1 Scatter pattern of the relationship between stunting and wasting prevalence and their predictor variables in Sumatra, 2022.

Based on Figure 1, the two scatter plots show the relationship between each predictor and Y_1 (Prevalence of Stunting) and Y_2 (Prevalence of Wasting). The data points are scattered randomly without a clear pattern, indicating that each predictor does not have a strong or consistent relationship with the two response variables. Although there are some points with high values of each predictor associated with high Y_1 or Y_2 values, the data distribution remains random.

C. Nadaraya-Watson Kernel Birespon Nonparametric Regression

The bivariate nonparametric regression using the Nadaraya-Watson Kernel is employed to model two response variables without assuming any specific distribution, utilizing distance-based weighting of predictors. This study uses the Epanechnikov, Gaussian, and Biweight kernels, and determines the optimal bandwidth within the range of 5 to 100 with an increment of 5 based on the minimum GCV criterion. The local estimation points are taken from the latest observation data to calculate the weighted average of the response values based on the proximity of the predictors. The next step is to estimate the regression function using the kernel functions determined in the previous stage. The estimation is performed by applying the selected kernel functions, namely Epanechnikov, Gaussian, or Biweight, to the predictor variables. The following is the formula of the kernel function that will be used by memorizing $\boldsymbol{u_i} = \sum_{g=1}^{4} \left(\frac{t_g - t_{gi}}{h_g} \right)$:

1. Epanechnikov

$$\hat{m}_E(u) = \sum_{i=1}^{154} \frac{\left[\left(z^2_{y^{(1)}} + 2z_{y^{(2)}y^{(1)}} + z^2_{y^{(2)}}\right)\left(\left(\frac{3}{4}(1-(u_i)^2)\right)\left(y_i^{(r)}\right)\right)\right]}{\frac{1}{154}\sum_{i=1}^{154}\left[\left(z^2_{y^{(1)}} + 2z_{y^{(2)}y^{(1)}} + z^2_{y^{(2)}}\right)\left(\frac{3}{4}(1-(u_i)^2)\right)\right]} \quad (16)$$

2. Gaussian

$$\hat{m}_G(u) = \sum_{i=1}^{154} \frac{\left[\left(z^2_{y^{(1)}} + 2z_{y^{(2)}y^{(1)}} + z^2_{y^{(2)}}\right)\left(\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{u_i^2}{2}}\right)\left(y_i^{(r)}\right)\right)\right]}{\frac{1}{154}\sum_{i=1}^{154}\left[\left(z^2_{y^{(1)}} + 2z_{y^{(2)}y^{(1)}} + z^2_{y^{(2)}}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{u_i^2}{2}}\right)\right]} \quad (17)$$

3. Biweight

$$\hat{m}_B(u) = \sum_{i=1}^{154} \frac{\left[\left(z^2_{y^{(1)}} + 2z_{y^{(2)}y^{(1)}} + z^2_{y^{(2)}}\right)\left(\left(\frac{15}{16}(1-u_i{}^2)^2\right)\left(y_i^{(r)}\right)\right)\right]}{\frac{1}{154}\sum_{i=1}^{154}\left[\left(z^2_{y^{(1)}} + 2z_{y^{(2)}y^{(1)}} + z^2_{y^{(2)}}\right)\left(\frac{15}{16}(1-u_i{}^2)^2\right)\right]} \quad (18)$$

The estimation results based on the three kernel functions are obtained as follows:

Table 2 Comparison of GCV Values of Kernel Functions

Kernel Function	GCV	Combination Number	Optimal Bandwidth t_1, t_2, t_3, t_4
Epanechnikov	0.110	159987	35,100,100,100
Gaussian	0.024	159987	35,100,100,100
Biweight	0.356	159988	40,100,100,100

Based on Table 2, the GCV values for the three kernel functions used in Nonparametric Bivariate Regression estimation (Epanechnikov, Gaussian, and Biweight) show that the smaller the GCV value, the better the kernel function is at providing optimal estimation. The table shows that the Gaussian kernel has the smallest GCV value (0.024), compared to Epanechnikov (0.110) and Biweight (0.356). This indicates that the Gaussian kernel produces the best estimation, as it optimally balances bias and variance in the Nonparametric Bivariate Regression model. Therefore, the Gaussian kernel is chosen as the best kernel function for regression estimation in this study. Based on the estimated functions $y^{(1)}$ and $y^{(2)}$, the prediction results can be plotted against the actual data, as shown in Figure 2.

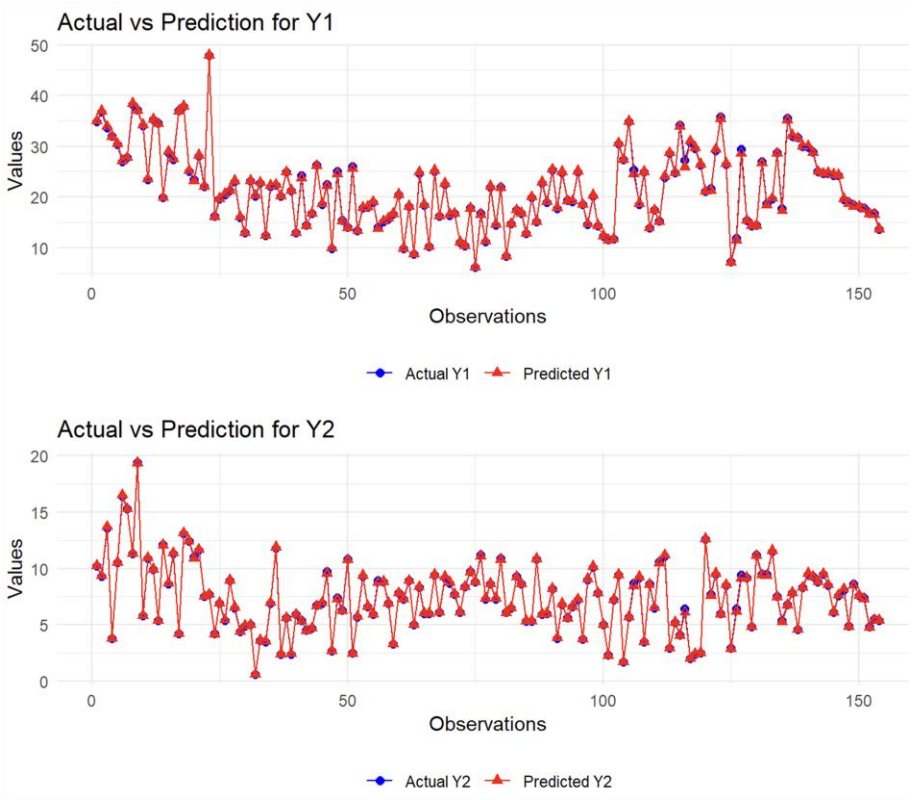

Figure 2 Actual and predicted values of Y_1 and Y_2 obtained using the Nadaraya–Watson kernel estimator.

D. Spline-penalized Birespon Nonparametric Regression

After applying the Kernel Nadaraya-Watson estimator, the analysis is continued using the Penalized Spline estimator to model the relationship between the response variables and predictors without assuming a specific functional form. Parameter estimation is performed using Penalized Weighted Least Squares to improve model accuracy. In bivariate nonparametric regression using Penalized Spline, the selection of polynomial order, lambda value, and number of knots greatly determines the model's flexibility. The maximum order is limited to 3, with combinations of order 1 (linear), order 2 (quadratic), and order 3 (cubic), where the optimal order is selected based on the minimum GCV value. Lambda values are tested within the range of 5 to 100 with increments of 5 to achieve the best smoothing level based on GCV. Meanwhile, the number of knots is limited to a maximum of 3 to maintain a balance between model flexibility and the risk of overfitting, with the optimal number also determined based on the minimum GCV value.The next step is to determine the optimal combination of order, knot locations, number of knots, and lambda values which is summarized in the following table:

Table 3 Optimal Knot Location

ξ_1	ξ_2	ξ_3	ξ_4
0.810	70.087	6.816	39.280
2.600	81.240	9.415	57.800
4.505	89.103	13.350	72.220
$Orde\ (3,3)$			
$\lambda = 5$			
$GCV = 6.960 \times 10^{-5}$			

Figure 3 shows the relationship between λ and GCV. As λ increases, it indicates that larger values of λ cause the model to become too smooth (underfitting) and lose data patterns. The optimal selection of λ is generally done by finding the minimum GCV value. However, in this graph, GCV continues to increase as λ becomes smaller. It is advised to avoid over-smoothing.

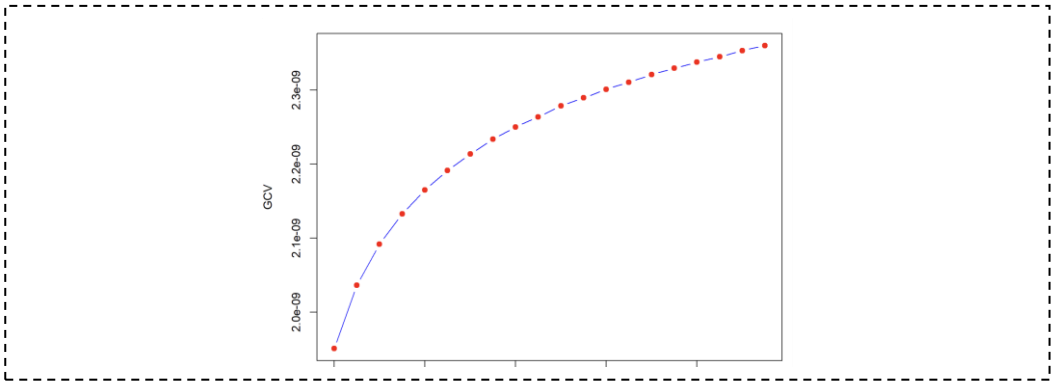

Figure 3 Relationship between the smoothing parameter (λ) and the Generalized Cross-Validation (GCV) value in the Penalized Spline model.

Based on the optimal combination of order, knot locations, number of knots, and lambda value, the bivariate nonparametric Penalized Spline regression models for both response variables are obtained as follows:

$$\begin{aligned}\hat{y}_i^{(1)} = {} & 32.410 - 1.660 \times 10^{0} t_{1i} + 8.115 \times 10^{-1} t_{1i}^2 - 5.621 \times 10^{-2} t_{1i}^3 + 2.384 \times 10^{-5} (t_{1i} - 0.810)_+^3 \\ & + 5.195 \times 10^{-4} (t_{1i} - 2.600)_+^3 - 2.221 \times 10^{-3} (t_{1i} - 4.505)_+^3 - 1.276 \times 10^{0} t_{2i} + 3.013 \times 10^{-2} t_{2i}^2 \\ & - 2.146 \times 10^{-4} t_{2i}^3 + 7.906 \times 10^{-4} (t_{2i} - 70.087)_+^3 + 1.604 \times 10^{-3} (t_{2i} - 81.240)_+^3 \\ & - 1.627 \times 10^{-2} (t_{2i} - 89.103)_+^3 + 2.740 \times 10^{0} t_{3i} - 2.496 \times 10^{-1} t_{3i}^2 + 6.942 \times 10^{-3} t_{3i}^3 \\ & + 3.268 \times 10^{-3} (t_{3i} - 6.816)_+^3 - 8.398 \times 10^{-4} (t_{3i} - 9.415)_+^3 - 1.808 \times 10^{-2} (t_{3i} - 13.350)_+^3 \\ & + 6.837 \times 10^{-1} t_{4i} + 2.773 \times 10^{-2} t_{4i}^2 - 3.058 \times 10^{-4} t_{4i}^3 + 4.846 \times 10^{-4} (t_{4i} - 39.280)_+^3 \\ & - 4.155 \times 10^{-4} (t_{4i} - 57.800)_+^3 \\ & + 8.274 \times 10^{-4} (t_{4i} - 72.220)_+^3 \end{aligned}$$

$$\begin{aligned}\hat{y}_i^{(2)} = {} & -3.301 \times 10^{0} - 2.849 \times 10^{-1} t_{1i} + 8.057 \times 10^{-2} t_{1i}^2 - 1.105 \times 10^{-3} t_{1i}^3 - 1.573 \times 10^{-4} (t_{1i} - 0.810)_+^3 \\ & + 2.438 \times 10^{-3} (t_{1i} - 2.600)_+^3 + 8.997 \times 10^{-3} (t_{1i} - 4.505)_+^3 - 2.190 \times 10^{-1} t_{2i} + 6.073 \times 10^{-3} t_{2i}^2 \\ & - 4.643 \times 10^{-5} t_{2i}^3 + 5.255 \times 10^{-4} (t_{2i} - 70.087)_+^3 - 1.912 \times 10^{-3} (t_{2i} - 81.240)_+^3 \\ & + 5.205 \times 10^{-3} (t_{2i} - 89.103)_+^3 + 4.091 \times 10^{0} t_{3i} - 6.091 \times 10^{-1} t_{3i}^2 + 2.867 \times 10^{-2} t_{3i}^3 \\ & - 1.585 \times 10^{-2} (t_{3i} - 6.816)_+^3 - 2.349 \times 10^{-2} (t_{3i} - 9.415)_+^3 + 1.653 \times 10^{-2} (t_{3i} - 13.350)_+^3 \\ & + 3.147 \times 10^{-1} t_{4i} - 7.297 \times 10^{-3} t_{4i}^2 + 5.469 \times 10^{-5} t_{4i}^3 - 5.131 \times 10^{-5} (t_{4i} - 39.280)_+^3 \\ & + 8.787 \times 10^{-5} (t_{4i} - 57.800)_+^3 \\ & + 4.101 \times 10^{-4} (t_{4i} - 72.220)_+^3 \end{aligned}$$

After establishing the PWLS model for the Stunting and Wasting case study in the Sumatra region in 2022, the next step is to segment the population based on predictor variables and interpret the regression coefficients to understand the relative influence of each variable on children's nutritional status. Through this analysis, it is expected to gain a deeper understanding of the factors affecting nutritional status, serving as a basis for more effective decision-making and policy planning. The segmentation results for each variable are presented as follows.

$$\begin{aligned} f^{(1)}(t_{1i}) = {} & -1.660 t_{1i} + 8.115 \times 10^{-1} t_{1i}^2 - 5.621 \times 10^{-2} t_{1i}^3 + 2.384 \times 10^{-5} (t_{1i} - 0.810)_+^3 + \\ & 5.195 \times 10^{-4} (t_{1i} - 2.600)_+^3 - 2.221 \times 10^{-3} (t_{1i} - 4.505)_+^3 \end{aligned}$$

$$f^{(1)}(t_{1i}) = \begin{cases} -1.660 t_{1i} + 0.8115 t_{1i}^2 - 0.05621 t_{1i}^3. & 0 < t_{1i} \le 0.810 \\ -1.659 t_{1i} + 0.8114 t_{1i}^2 - 0.05618 t_{1i}^3 - 0.00002. & 0.810 < t_{1i} \le 2.600 \\ -1.649 t_{1i} + 0.8074 t_{1i}^2 - 0.05566 t_{1i}^3 - 0.009. & 2.600 < t_{1i} \le 4.505 \\ -1.784 t_{1i} + 0.8374 t_{1i}^2 - 0.05788 t_{1i}^3 - 0.194. & t_{1i} > 4.505 \end{cases}$$

Based on the segmentation results of the Low Birth Weight Percentage variable, the effect of one-unit $\boldsymbol{t_{1i}}$ on LBW shows a varied pattern in each segment. In Segment I, a one-unit increase in $\boldsymbol{t_{1i}}$causes a decrease in $\boldsymbol{y^{(1)}}$, which means that more babies born with low birth weight will reduce the prevalence of stunting. Entering Segment II, a one-unit increase in $\boldsymbol{t_{1i}}$ causes a decrease in $\boldsymbol{y^{(1)}}$, which means that the more babies born with low body weight, the lower the prevalence of stunting. In Segment III, a one-unit increase in $\boldsymbol{t_{1i}}$ causes a decrease in $\boldsymbol{y^{(1)}}$, which means that the more babies born with low body weight, the lower the prevalence of stunting. However, in Segment IV, a one-unit increase in $\boldsymbol{t_{1i}}$ causes a decrease in $\boldsymbol{y^{(1)}}$, which means that the more babies born with low birth weight, the lower the prevalence of stunting.

$$\begin{aligned} f^{(1)}(t_{2i}) = {} & -1.276 t_{2i} + 3.013 \times 10^{-2} t_{2i}^2 - 2.146 \times 10^{-4} t_{2i}^3 + 7.906 \times 10^{-4} (t_{2i} - 70.087)_+^3 + \\ & 1.604 \times 10^{-3} (t_{2i} - 81.240)_+^3 - 1.627 \times 10^{-2} (t_{2i} - 89.103)_+^3 \end{aligned}$$

$$f^{(1)}(t_{2i}) = \begin{cases} -1.276 t_{2i} + 0.003 t_{2i}^2 - 0.00002 t_{2i}^3. & 0 < t_{2i} \le 70.087 \\ 10.374 t_{2i} + 0.163 t_{2i}^2 - 0.0007 t_{2i}^3 - 272.188. & 70.087 < t_{2i} \le 81.240 \\ 42.133 t_{2i} - 0.228 t_{2i}^2 - 0.0009 t_{2i}^3 - 1132.219. & 81.240 < t_{2i} \le 89.103 \\ -345.486 t_{2i} + 4.121 t_{2i}^2 - 0.017 t_{2i}^3 - 10377.495. & t_{2i} > 89.103 \end{cases}$$

Based on the segmentation results of the percentage of proper sanitation variable, the effect of $\boldsymbol{t_{2i}}$ on the percentage of proper sanitation shows a varied pattern in each segment. In Segment I, a one-unit increase in $\boldsymbol{t_{2i}}$ causes a decrease in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of proper sanitation, the lower the prevalence of stunting. Entering Segment II, an increase in $\boldsymbol{t_{2i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of proper sanitation, the higher the prevalence of stunting. In Segment III, an increase in $\boldsymbol{t_{2i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of proper sanitation, the higher the prevalence of stunting. However, in Segment IV, a one-unit increase in $\boldsymbol{t_{2i}}$ causes a decrease in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of proper sanitation, the lower the prevalence of stunting.

$$f^{(1)}(t_{3i}) = 2.740t_{3i} - 2.496 \times 10^{-1}t_{3i}^2 + 6.942 \times 10^{-3}t_{3i}^3 + 3.268 \times 10^{-3}(t_{3i} - 6.816)_+^3$$
$$-8.398 \times 10^{-4}\,(t_{3i} - 9.415)_+^3 - 1.808 \times 10^{-2}(t_{3i} - 13.350)_+^3$$

$$f^{(1)}(t_{3i}) = \begin{cases} 2.740t_{3i} - 0.249t_{3i}^2 + 0.007t_{3i}^3. & 0 < t_{3i} \leq 6.816 \\ 3.195t_{3i} - 0.316t_{3i}^2 - 0.004t_{3i}^3 - 1.035. & 6.816 < t_{3i} \leq 9.415 \\ 2.972t_{3i} - 0.292t_{3i}^2 - 0.005t_{3i}^3 - 0.334. & 9.415 < t_{3i} \leq 13.350 \\ -6.695t_{3i} - 0.432t_{3i}^2 - 0.023t_{3i}^3 + 42.683. & t_{3i} > 13.350 \end{cases}$$

Based on the segmentation results of the percentage of poor people variable, the effect of t_3i on the percentage of poor people shows a varied pattern in each segment. In Segment I, a one-unit increase in $\boldsymbol{t_{3i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of poor people, the higher the prevalence of stunting. Entering Segment II, an increase in $\boldsymbol{t_{3i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of poor people, the higher the prevalence of stunting. In Segment III, an increase in $\boldsymbol{t_{3i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of poor people, the higher the prevalence of stunting. However, in Segment IV, a one-unit increase in $\boldsymbol{t_{3i}}$ causes a decrease in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of poor people, the lower the prevalence of stunting.

$$f^{(1)}(t_{4i}) = 6.837 \times 10^{-1}t_{4i} + 2.773 \times 10^{-2}t_{4i}^2 - 3.058 \times 10^{-4}t_{4i}^3 + 4.846 \times 10^{-4}(t_{4i} - 39.280)_+^3$$
$$-4.155 \times 10^{-4}(t_{4i} - 57.800)_+^3 + 8.274 \times 10^{-4}(t_{4i} - 72.220)_+^3$$

$$f^{(1)}(t_{4i}) = \begin{cases} 0.684t_{4i} + 0.027t_{4i}^2 - 0.0003t_{4i}^3. & 0 < t_{4i} \leq 39.280 \\ 2.927t_{4i} - 0.085t_{4i}^2 + 0.0002t_{4i}^3 - 29.369. & 39.280 < t_{4i} \leq 57.800 \\ -1.237t_{4i} - 0.013t_{4i}^2 - 0.0006t_{4i}^3 + 50.864. & 57.800 < t_{4i} \leq 72.220 \\ 11.709t_{4i} - 0.192t_{4i}^2 + 0.00023t_{4i}^3 - 260.801. & t_{4i} > 72.220 \end{cases}$$

Based on the segmentation results of the percentage of ASI-E variable, the effect of $\boldsymbol{t_{4i}}$ on the percentage of ASI-E shows a varied pattern in each segment. In Segment I, a one-unit increase in $\boldsymbol{t_{4i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of ASI-E, the higher the prevalence of stunting. Entering Segment II, an increase in $\boldsymbol{t_{4i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of breastfeeding-E, the higher the prevalence of stunting. In Segment III, an increase in $\boldsymbol{t_{4i}}$ causes an increase in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of breastfeeding-E, the higher the prevalence of stunting. However, in Segment IV, a one-unit increase in $\boldsymbol{t_{4i}}$ causes a decrease in $\boldsymbol{y^{(1)}}$, which means that the higher the percentage of E-breastfeeding, the lower the prevalence of stunting. Response function estimation results 1 as followed:

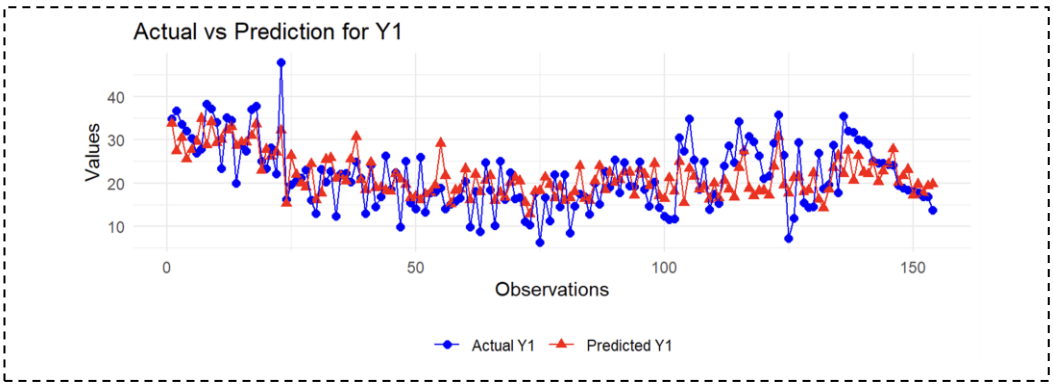

Figure 4 Estimated $\boldsymbol{y^{(1)}}$spline function showing the relationship between the actual and predicted values for $\boldsymbol{Y_1}$.

$$f^{(2)}(t_{1i}) = -2.849 \times 10^{-1}t_{1i} + 8.057 \times 10^{-2}t_{1i}^2 - 1.105 \times 10^{-3}t_{1i}^3 - 1.573 \times 10^{-4}(t_{1i} - 0.810)_+^3$$
$$+ 2.438 \times 10^{-3}(t_{1i} - 2.600)_+^3 + 8.997 \times 10^{-3}(t_{1i} - 4.505)_+^3$$

$$f^{(2)}(t_{1i}) = \begin{cases} -0.285t_{1i} + 0.0806t_{1i}^2 - 0.0011t_{1i}^3. & 0 < t_{1i} \leq 0.810 \\ -0.285t_{1i} + 0.0809t_{1i}^2 - 0.0013t_{1i}^3 + 0.000084. & 0.810 < t_{1i} \leq 2.600 \\ -0.2355t_{1i} + 0.0618t_{1i}^2 + 0.0011t_{1i}^3 - 0.0428. & 2.600 < t_{1i} \leq 4.505 \\ 0.3123t_{1i} - 0.0598t_{1i}^2 + 0.0101t_{1i}^3 - 0.8654. & t_{1i} > 4.505 \end{cases}$$

Based on the segmentation results of the Low Birth Weight Percentage variable, the effect of one-unit t_{1i} on LBW shows a varied pattern in each segment. In Segment I, a one-unit increase in t_{1i}causes a decrease in $y^{(2)}$, which means that more babies born with low birth weight will reduce the prevalence of wasting. Entering Segment II, a one-unit increase in t_{1i} causes a decrease in $y^{(2)}$, which means that the more babies born with low body weight, the lower the prevalence of wasting. In Segment III, a one-unit increase in t_{1i} causes a decrease in $y^{(2)}$, which means that the more babies born with low body weight, the lower the prevalence of wasting. However, in Segment IV, an increase in t_{1i} causes an increase in $y^{(2)}$, which means that the more babies born with low weight will increase the prevalence of wasting.

$$f^{(2)}(t_{2i}) = -2.190 \times 10^{-1} t_{2i} + 6.073 \times 10^{-3} t_{2i}^2 - 4.643 \times 10^{-5} t_{2i}^3 + 5.255 \times 10^{-4} (t_{2i} - 70.087)_+^3$$
$$-1.912 \times 10^{-3} (t_{2i} - 81.240)_+^3 + 5.205 \times 10^{-3} (t_{2i} - 89.103)_+^3$$

$$f^{(2)}(t_{2i}) = \begin{cases} -0.219 t_{2i} + 0.006 t_{2i}^2 - 0.000052 t_{2i}^3. & 0 < t_{2i} \leq 70.087 \\ 7.525 t_{2i} - 0.1044 t_{2i}^2 + 0.00048 t_{2i}^3 - 180.919. & 70.087 < t_{2i} \leq 81.240 \\ -30.332 t_{2i} + 0.362 t_{2i}^2 - 0.0114 t_{2i}^3 + 844.255. & 81.240 < t_{2i} \leq 89.103 \\ 93.641 t_{2i} - 1.029 t_{2i}^2 - 0.0062 t_{2i}^3 - 2837.863. & t_{2i} > 89.103 \end{cases}$$

Based on the segmentation results of the percentage of proper sanitation variable, the effect of t_{2i} on the percentage of proper sanitation shows a varied pattern in each segment. In Segment I, a one-unit increase in t_{2i} causes a decrease in $y^{(2)}$, which means that the higher the percentage of proper sanitation, the lower the prevalence of wasting. Entering Segment II, an increase in t_{2i} causes an increase in $y^{(2)}$, which means that the higher the percentage of proper sanitation, the higher the prevalence of wasting. In Segment III, an increase in t_{2i} causes a decrease in $y^{(2)}$, which means that the higher the percentage of proper sanitation, the lower the prevalence of wasting. However, in Segment IV, a one-unit increase in t_{2i} causes an increase in $y^{(2)}$, which means that the higher the percentage of proper sanitation, the higher the prevalence of wasting.

$$f^{(2)}(t_{3i}) = 4.091 t_{3i} - 6.091 \times 10^{-1} t_{3i}^2 + 2.867 \times 10^{-2} t_{3i}^3 - 1.585 \times 10^{-2} (t_{3i} - 6.816)_+^3$$
$$-2.349 \times 10^{-2} (t_{3i} - 9.415)_+^3 + 1.653 \times 10^{-2} (t_{3i} - 13.350)_+^3$$

$$f^{(2)}(t_{3i}) = \begin{cases} 4.091 t_{3i} - 0.609 t_{3i}^2 + 0.0087 t_{3i}^3. & 0 < t_{3i} \leq 6.816 \\ 1.882 t_{3i} - 0.285 t_{3i}^2 - 0.0071 t_{3i}^3 + 5.109. & 6.816 < t_{3i} \leq 9.415 \\ 1.257 t_{3i} - 0.218 t_{3i}^2 - 0.0095 t_{3i}^3 + 6.979. & 9.415 < t_{3i} \leq 13.350 \\ 10.095 t_{3i} - 0.880 t_{3i}^2 + 0.0070 t_{3i}^3 - 32.350. & t_{3i} > 13.350 \end{cases}$$

Based on the segmentation results of the percentage of poor people variable, the effect of t_{3i} on the percentage of poor people shows a varied pattern in each segment. In Segment I, a one-unit increase in t_{3i} causes an increase in $y^{(2)}$, which means that the higher the percentage of poor people, the higher the prevalence of wasting. Entering Segment II, an increase in t_{3i} causes an increase in $y^{(2)}$, which means that the higher the percentage of poor people, the higher the prevalence of wasting. In Segment III, an increase in t_{3i} causes an increase in $y^{(2)}$, which means that the higher the percentage of poor people, the higher the prevalence of wasting. However, in Segment IV, a one-unit increase in t_{3i} causes an increase in $y^{(2)}$, which means that the higher the percentage of poor people, the higher the prevalence of wasting.

$$f^{(2)}(t_{4i}) = 3.147 \times 10^{-1} t_{4i} - 7.297 \times 10^{-3} t_{4i}^2 + 5.469 \times 10^{-5} t_{4i}^3 - 5.131 \times 10^{-5} (t_{4i} - 39.280)_+^3$$
$$+8.787 \times 10^{-5} (t_{4i} - 57.800)_+^3 + 4.101 \times 10^{-4} (t_{4i} - 72.220)_+^3$$

$$f^{(2)}(t_{4i}) = \begin{cases} 0.315 t_{4i} + 0.0073 t_{4i}^2 - 0.00005 t_{4i}^3. & 0 < t_{4i} \leq 39.280 \\ 0.0772 t_{4i} - 0.001 t_{4i}^2 + 0.000003 t_{4i}^3 - 3.109. & 39.280 < t_{4i} \leq 57.800 \\ 0.958 t_{4i} - 0.016 t_{4i}^2 - 0.00009 t_{4i}^3 - 20.077. & 57.800 < t_{4i} \leq 72.220 \\ 0.958 t_{4i} - 0.016 t_{4i}^2 - 0.00009 t_{4i}^3 - 20.077. & t_{4i} > 72.220 \end{cases}$$

Based on the segmentation results of the percentage of ASI-E variable, the effect of t_{4i} on the percentage of ASI-E shows a varied pattern in each segment. In Segment I, a one-unit increase in t_{4i} causes an increase in $y^{(2)}$, which means that the higher the percentage of ASI-E, the higher the prevalence of wasting. Entering Segment II, an increase in t_{4i} causes an increase in $y^{(2)}$, which means that the higher the percentage of breastfeeding-E, the higher the prevalence of wasting. In Segment III, an increase in t_{4i} causes an increase in $y^{(2)}$, which means that the higher the percentage of breastfeeding-E, the higher the prevalence of wasting. However, in Segment IV, a one-unit increase in t_{4i} causes a decrease in $y^{(2)}$, which means that the higher the percentage of breastfeeding-E, the lower the prevalence of wasting. Response function estimation results 2 as followed:

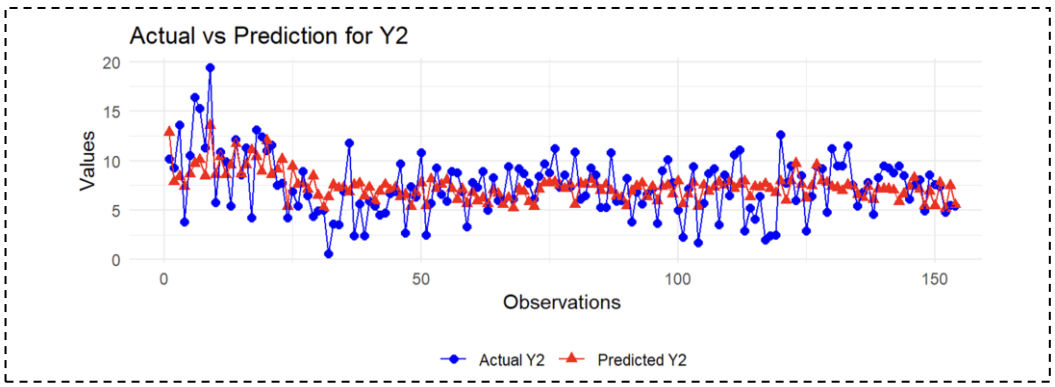

Figure 5 Estimated $\boldsymbol{y}^{(2)}$spline function showing the relationship between the actual and predicted values for $\boldsymbol{Y_2}$.

E. Model Selection and Accuracy

The best model between Kernel Nadaraya-Watson and Penalized Spline is selected based on the minimum Mean Squared Error (MSE). The Kernel Nadaraya-Watson model shows a significantly lower MSE (0.024) compared to the Penalized Spline model (19.917), indicating that it provides better predictions. Based on the evaluation in section 5.6, the Kernel model is chosen as the optimal model for regression estimation, as it effectively captures data patterns with minimal error.

The accuracy of the model is further validated using R-Squared and Root Mean Squared Error (RMSE). The Kernel Nadaraya-Watson model demonstrated superior performance with high R-Squared values (0.9992 for response 1 and 0.9995 for response 2) and lower RMSE, indicating excellent predictive capability. These results confirm that the Kernel model is highly effective in explaining data variance and making accurate predictions, making it the preferred model for further analysis and decision-making.

V. CONCLUSIONS AND SUGGESTIONS

This study highlights the importance of nonparametric regression techniques, specifically Nadaraya-Watson Kernel and Penalized Spline regression, in modeling the relationship between two correlated response variables: Stunting and Wasting. These models offer flexibility and adaptability in capturing complex patterns in data, particularly in health studies where such relationships are prevalent. The analysis indicates that these models, especially the Kernel method, provide more accurate predictions and better explain the variance in the data, with the Gaussian kernel showing the best performance. The findings emphasize the significance of addressing the factors influencing children's nutritional status, such as birth weight, sanitation, and breastfeeding, which are crucial for formulating effective interventions to reduce malnutrition in Indonesia. Future studies could explore further refinements in these models and extend their applications to other health and social determinants.

ACKNOWLEDGEMET

I would like to express my deepest gratitude to my supervisor, Sigit Nugroho, for her invaluable guidance, continuous support, and encouragement throughout this research. I am equally grateful to my co-supervisor, Idhia Sriliana, for her insightful feedback and constructive suggestions, which greatly enhanced the quality of this work. Additionally, I extend my heartfelt appreciation to my fellow "spline" companions, who have shared this challenging yet rewarding journey with me. Your unwavering support, collaboration, and encouragement have been an immense source of motivation. Thank you all for making this journey memorable and enriching.

REFERENCES

[1] R. L. Eubank, *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker, Inc, 1999.

[2] P. P. Gabrela, J. D. T. Purnomo, and I. N. Budiantara, "The Estimation of Mixed Truncated Spline and Fourier Series Estimator in Bi-Response Nonparametric Regression," *AIP Conf. Proc.*, vol. 2903, no. 1, 2023, doi: 10.1063/5.0177224.

[3] I. Sriliana, I. N. Budiantara, and V. Ratnasari, "The Performance of Mixed Truncated Spline-Local Linear Nonparametric Regression Model for Longitudinal Data," *MethodsX*, vol. 12, no. July 2023, 2024, doi: 10.1016/j.mex.2024.102652.

[4] R. Hidayat, I. N. Budiantara, B. W. Otok, and V. Ratnasari, "The regression curve estimation by using mixed smoothing spline and kernel (MsS-K) model," *Commun. Stat. - Theory Methods*, vol. 50, no. 17, pp. 3942–3953, 2021, doi: 10.1080/03610926.2019.1710201.

[5] T. H. Ali, H. A. A. M. Hayawi, and D. S. I. Botani, "Estimation of the bandwidth parameter in Nadaraya-Watson kernel Non-Parametric Regression Based on Universal Threshold Level," *Commun. Stat. Simul. Comput.*, vol. 52, no. 4, pp. 1476–1489, 2021, doi: 10.1080/03610918.2021.1884719.

[6] G. Yang, B. Zhang, and M. Zhang, "Estimation of Knots in Linear Spline Models," *J. Am. Stat. Assoc.*, vol. 118, no. 541, pp. 639–

650, 2023, doi: 10.1080/01621459.2021.1947307.

[7] D. Barry and W. Hardle, *Applied Nonparametric Regression.*, 1st ed. Cambridge University Press., 1994. doi: 10.2307/2982873.

[8] C. R. Oktarina, I. Sriliana, and S. Nugroho, "Penalized Spline Semiparametric Regression for Bivariate Response in Modeling Macro Poverty Indicators," *Indones. J. Appl. Stat.*, vol. 8, no. 1, pp. 13–23, 2025, doi: https://doi.org/10.13057/ijas.v8i1.94370.

[9] K. Xie, F. Meng, and D. Zhang, "Regional Forecasting of PM2.5 Concentrations: A Novel Model Based on The Empirical Orthogonal Function Analysis and Nadaraya–Watson Kernel Regression Estimator," *Environ. Model. Softw.*, vol. 170, no. January, p. 105857, 2023, doi: 10.1016/j.envsoft.2023.105857.

[10] A. R. Fadilah, A. Fitrianto, and I. M. Sumertajaya, "Outlier Identification on Penalized Spline Regression Modeling for Poverty Gap Index in Java," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 4, pp. 1231–1240, 2022, doi: 10.30598/barekengvol16iss4pp1231-1240.

[11] M. Blössner and M. de Onis, "Malnutrition Quantifying the Health Impact at National and Local Levels," 12th ed., no. 12, Geneva: World Health Organization, 2005.

[12] I. N. Budiantara, *Regresi Nonparametrik Spline Truncated*. Surabaya: ITS Press, 2019.

[13] D. Ruppert, M. P. Wand, and R. . Carroll, *Semiparametric Regression*. Cambridge University Press, 2003. doi: 10.1201/9781420091984-c17.

[14] Q. Yang, Q. Kang, Q. Huang, Z. Cui, Y. Bai, and H. Wei, "Linear correlation analysis of ammunition storage environment based on Pearson correlation analysis," *J. Phys. Conf. Ser.*, vol. 1948, no. 1, 2021, doi: 10.1088/1742-6596/1948/1/012064.

[15] Sifriyani, A. R. M. Sari, A. T. R. Dani, and S. Jalaluddin, "Bi-Response Truncated Spline Nonparametric Regression With Optimal Knot Point Selection Using Generalized Cross-Validation in Diabetes Mellitus Patient'S Blood Sugar Levels," *Commun. Math. Biol. Neurosci.*, vol. 2023, pp. 1–18, 2023, doi: 10.28919/cmbn/7903.

[16] S. Tosatto, R. Akrour, and J. Peters, "An Upper Bound of the Bias of Nadaraya-Watson Kernel Regression Under Lipschitz Assumptions," *Stats*, vol. 4, no. 1, pp. 1–17, 2021, doi: 10.3390/stats4010001.

[17] L. Selk and J. Gertheiss, "Nonparametric Regression and Classification with Functional, Categorical, and Mixed Covariates," *Adv. Data Anal. Classif.*, vol. 17, no. 2, pp. 519–543, 2022, doi: 10.1007/s11634-022-00513-7.

[18] T. Misiakiewicz and B. Saeed, "A non-Asymptotic Theory of Kernel Ridge Regression: Deterministic Equivalents, Test Error, and GCV Estimator," pp. 1–131, 2024, [Online]. Available: http://arxiv.org/abs/2403.08938

[19] L. Yang and Y. Hong, "Adaptive Penalized Splines for Data Smoothing," *Comput. Stat. Data Anal.*, vol. 108, pp. 70–83, 2017, doi: 10.1016/j.csda.2016.10.022.

[20] X. Li *et al.*, "A Practical and Effective Regularized Polynomial Smoothing (RPS) Method for High-Gradient Strain Field Measurement In Digital Image Correlation," *Opt. Lasers Eng.*, vol. 121, no. April, pp. 215–226, 2019, doi: 10.1016/j.optlaseng.2019.04.017.

[21] R. Pahlepi, I. Sriliana, W. Agwil, and C. R. Oktarina, "Biresponse Spline Truncated Nonparametric Regression Modeling for Longitudinal Data on Monthly Stock Prices of Three Private Banks in Indonesia," *Barekeng*, vol. 19, no. 4, pp. 2467–2480, 2025, doi: 10.30598/barekengvol19iss4pp2467-2480.

[22] I. Sriliana, I. N. Budiantara, and V. Ratnasari, "Poverty gap index modeling in Bengkulu Province using truncated spline regression for longitudinal data," *AIP Conf. Proc.*, vol. 2975, no. 1, 2023, doi: https://doi.org/10.1063/5.0181178.

[23] World Health Organization, "Stunting in a Nutshell." [Online]. Available: https://www.who.int/news/item/19-11-2015-stunting-in-a-nutshell

[24] S. S. Anton, N. M. U. K. Dewi, and I. G. Adiba, "Kajian Determinan Stunting Pada Anak di Indonesia," *J. Yoga dan Kesehat.*, vol. 6, no. 2, pp. 201–217, 2023, doi: 10.25078/jyk.v6i2.2907.

[25] D. A. Ningsih, "Kajian Determinan yang Berhubungan dengan Status Gizi Kurang pada Balita," *J. Ilmu Gizi Indones.*, vol. 3, no. 1, pp. 28–34, 2022, doi: 10.57084/jigzi.v3i1.885.