

NONPARAMETRIC SPLINE REGRESSION APPROACH IN CASE STUDY OF MATERNAL MORTALITY RATES IN SOUTH SULAWESI

Nur Salsabila Wero¹, Rahmat Hidayat^{2*}, Ruliana²

¹Postgraduate Student in Department Statistics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Makassar, South Sulawesi, Indonesia

²Department Statistics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Makassar, South Sulawesi, Indonesia

*e-mail: shalsabilaweros@gmail.com

DOI: 10.14710/j.gauss.15.2.379-390

Article Info:

Received: 2026-05-12

Accepted: 2026-10-12

Available Online: 2026-10-20

Keywords:

Maternal Mortality Rate;
Nonparametric Regression; Spline
Regression; Knot Point;
Generalized Cross Validation
(GCV)

Abstract: Maternal Mortality Rate (MMR) is an important indicator of maternal health and the quality of healthcare services. Its relationship with maternal health-related conditions may be nonlinear, requiring a flexible modeling approach. This study aims to model MMR using nonparametric spline regression and identify the optimal knot configuration for the predictor variables. The response variable was MMR, while the predictor variables consisted of Bleeding (X_1), Hypertension Disorders (X_2), Post-Miscarriage Complications (X_3), and Heart and Blood Vessel Disorders (X_4). Secondary data were obtained from the South Sulawesi Provincial Health Office for 24 districts/cities in 2023. The optimal knot configuration was determined using the Generalized Cross-Validation (GCV) criterion. The selected model used one knot for X_1 , three knots for X_2 , two knots for X_3 , and one knot for X_4 . The parameter significance test showed that Bleeding (X_1) was the only predictor with a statistically significant linear component at the 5% significance level, with a p-value of 0.0006, while the other linear and spline components were not statistically significant. The model achieved an MSE of 0.9388, RMSE of 0.9689, MAPE of 15.6494%, and R^2 of 0.9653. These results indicate that nonparametric spline regression flexibly models relationships between MMR and selected maternal health-related conditions effectively.

1. INTRODUCTION

The Maternal Mortality Rate (MMR) is an important indicator of maternal health and the quality of healthcare services in a region (Aswi *et al.*, 2025). A high MMR reflects not only medical and obstetric challenges but is also closely associated with socioeconomic conditions and access to adequate maternal healthcare services (Badan Pusat Statistik, 2023). Globally, approximately 800 women die each day from preventable causes related to pregnancy and childbirth, with more than 80% of maternal deaths considered preventable through timely and appropriate interventions (Hidayat *et al.*, 2025). In Indonesia, the MMR decreased from 346 per 100,000 live births in 2010 to 189 per 100,000 live births in 2020. However, this decline has not been evenly distributed across regions, including South Sulawesi Province, which has diverse geographical and socioeconomic conditions, as well as differences in access to maternal healthcare services (Badan Pusat Statistik, 2024).

Several medical conditions are associated with maternal mortality. Hemorrhage is one of the major conditions that may lead to maternal death, particularly when timely treatment is not provided. Hypertensive disorders during pregnancy also require particular attention because they may progress to more severe complications (Sampurna *et al.*, 2023). In addition, complications following miscarriage may result in serious health problems when appropriate treatment is not provided. Heart and blood vessel disorders may also increase the risk of

complications during pregnancy and childbirth, as physiological changes during pregnancy place additional demands on the cardiovascular system (Aswi *et al.*, 2025).

The occurrence of these conditions varies across districts and municipalities. Such variation may result in different patterns of association with maternal mortality. The relationship between a particular condition and maternal mortality may also not change at a constant rate across different levels of occurrence (Raina *et al.*, 2023). At certain levels, a change in the number of cases may be associated with a different change in maternal mortality than that observed at other levels. This suggests that the relationship between maternal mortality and its associated factors should be examined without imposing a specific functional form a priori (Qomari and Notobroto, 2022).

Several studies in recent years have applied spline-based approaches to model maternal mortality. (Fadhilah and Islamiyati, 2025) analyzed maternal mortality in South Sulawesi while accounting for spatial variation across regions. (Fadhilah, Islamiyati and Nirwan, 2025) applied Multivariate Adaptive Generalized Poisson Regression Spline to maternal mortality data in South Sulawesi and demonstrated the applicability of spline-based methods in capturing nonlinear relationships. More recently, (Annas *et al.*, 2026) applied nonparametric B-spline regression to maternal mortality rate data in Indonesia and identified nonlinear relationships between maternal mortality and the factors considered. These studies demonstrate the potential of spline-based approaches to flexibly characterize the relationship between maternal mortality and its associated factors.

Thus, there remains a need to examine nonlinear relationships between maternal mortality and conditions associated with maternal deaths at the district and municipal level in South Sulawesi using nonparametric spline regression, without incorporating spatial or Poisson components as the primary structure of the model. Furthermore, previous studies in South Sulawesi have not specifically examined the relationship between maternal mortality and the combination of hemorrhage, hypertensive disorders, complications following miscarriage, and heart and blood vessel disorders using a nonparametric spline regression framework.

Nonparametric spline regression provides an appropriate framework for this purpose because it does not require the functional relationship between the response and explanatory variables to be specified in advance. Spline functions are constructed over several intervals and connected at selected knot points, allowing changes in the relationship across different ranges of the explanatory variables to be captured. This flexibility enables the association between maternal mortality and its related conditions to be characterized according to the underlying pattern observed in the data.

Based on these considerations, this study applies nonparametric spline regression to model maternal mortality in South Sulawesi. The approach is used to examine the patterns of association between maternal mortality and hemorrhage, hypertensive disorders, complications following miscarriage, and heart and blood vessel disorders. The findings are expected to provide a clearer description of these relationships across districts and municipalities in South Sulawesi and contribute to a better understanding of factors associated with maternal mortality.

2. LITERATURE REVIEW

Nonparametric regression is a regression method in which the functional form of the regression curve is not specified in advance (Muthahharah and Hidayat, 2025). In nonparametric regression, the regression curve is simply assumed to operate smoothly within a certain function space, providing a high degree of flexibility (Abdy, 2019). The nonparametric regression model can be expressed in equation (Puspitasari *et al.*, 2023).

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, 3, \dots, n \quad (1)$$

Where y_i is the response variable for the i -th observation, $f(x_i)$ is an unknown smooth function, X_i is the predictor variable for the i -th observation and ε_i is a random error term. The error term is assumed to be independently and identically distributed with mean zero and variance σ^2 , such that $\varepsilon_i \sim N(0, \sigma^2)$. Spline regression is one of the methods used in nonparametric regression to model the relationship between response and predictor variables when the relationship is nonlinear (Hesikumalasari *et al.*, 2022). This method is useful when the relationship between variables cannot be adequately explained by simple parametric models, such as linear regression, because it may have a complex form and change across different intervals (Safitri *et al.*, 2025). The spline regression approach uses polynomial functions that are defined over several intervals and connected at selected knot points (Hidayat and Ilyas, 2023).

These intervals are determined by specific points known as knot points (Safitri *et al.*, 2025). The presence of knot points allows the regression function to adapt to changes in the data pattern across different intervals, producing a more flexible curve than conventional regression models (Syamsualam and Hidayat, 2022). Although a spline consists of several polynomial pieces, the resulting function remains smooth because continuity conditions are imposed at the knot points (Hidayat *et al.*, 2021). Because existence condition continuity at knot points (Suriaslan, Budiantara and Ratnasari, 2025). In general, spline regression is able to balance model flexibility and estimation stability (Arifanti and Hidayat, 2021). This makes spline regression an effective method for modeling data with unknown patterns, particularly in the fields of health, economics, and the environment. The spline regression model can be expressed as follows (Nurhuda, Wasono and Nohe, 2022):

$$f(x_i) = \sum_{j=0}^p \beta_j x_i^j + \sum_{l=1}^K \beta_{(p+l)} (x_i - k_l)^p \quad (2)$$

with

$$(x_i - k_l)^p = \begin{cases} (x_i - k_l)^p, & x_i \geq k_l \\ 0, & x_i < k_l \end{cases} \quad (3)$$

where $f(x_i)$ is spline regression function, $k_1, k_2, \dots, k_K$ is knot point, x_i is predictor variable, β is regression parameters and p is degree of polynomial. The parameters of the spline regression model are estimated using the Ordinary Least Squares (OLS) method. In matrix form, the spline regression model can be expressed as (Ramadhan, Nohe and Wahyuningsih, 2023):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (4)$$

where $\mathbf{y}$ is the response vector, $\mathbf{X}$ is the spline design matrix, $\boldsymbol{\beta}$ is the vector of regression parameters, and $\boldsymbol{\varepsilon}$ is the error vector. The OLS estimator is obtained by minimizing the sum of squared residuals, resulting in (Handajani, Pratiwi and Qona, 2023):

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

The estimated parameters are then substituted into the spline regression function to obtain the fitted model. The model shows that the spline function is a combination of a global polynomial and additional components that are active after passing certain knot points. This structure allows the model to adapt to local changes in the data pattern while maintaining overall stability (Sriliana, Budiantara and Ratnasari, 2022). The selection of knot points is an important stage in spline regression because it directly affects the shape and performance of the resulting model. The knot point functions as barrier between intervals that allow the model to adjust change data patterns in each part (Hidayat, Muthahharah and Resmawan, 2025). The inclusion of knot points allows spline models to flexibly capture the relationship between variables (Sanusi, Syam and Adawiyah, 2019).

However, the number and locations of knot points should be selected carefully. If too few knot points are used, the model may not be sufficiently flexible to capture the underlying data pattern, resulting in underfitting (Kaseside and Loklomin, 2021). Conversely, using too many knot points may make the model overly complex and cause it to capture random fluctuations in the data, resulting in overfitting. Therefore that, is necessary something methods that can help determine the most optimal knot point (Ariesta, Gusriani and Parmikanti, 2021). One of frequent methods used in election knot point is Generalized Cross Validation (GCV This method is used to evaluate model performance by balancing prediction error and model complexity. In general, good model is a model that has the smallest GCV value. Mathematically, the GCV value can be written as follows (Puspitasari *et al.*, 2023):

$$\text{GCV} = \frac{\text{MSE}}{\left(1 - \frac{\text{tr}(\mathbf{H})}{n}\right)^2} \quad (6)$$

with:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

Where n is number of observations, y_i is observed value of the response variable, $\hat{y}_i$ is predicted value of the response variable, $\mathbf{H}$ is a hat matrix, and $\text{tr}(\mathbf{H})$ is trace of the hat matrix. Obtained from the sum of its diagonal elements. The hat matrix is defined as (Ariesta, Gusriani and Parmikanti, 2021):

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (8)$$

Where $\mathbf{X}$ is the design matrix containing the spline basis functions corresponding to a particular knot configuration. For each candidate knot configuration, a different spline design matrix $\mathbf{X}$ is constructed according to the number and position of the selected knot points. Consequently, the corresponding hat matrix $\mathbf{H}$ and its trace $\text{tr}(\mathbf{H})$ also change for each candidate model. The GCV value is then calculated for each candidate knot configuration using Equation (6). The knot configuration that produces the smallest GCV value is selected as the optimal configuration for the spline regression model (Handajani, Pratiwi and Qona, 2023).

Once the spline regression model has been constructed, parameter testing is conducted to determine whether the variables included in the model significantly contribute to the response variable. Parameter testing generally consists of simultaneous and partial tests. The simultaneous test is used to determine whether the model parameters jointly influence the response variable, whereas the partial test is used to evaluate the significance of individual parameters (Suriaslan, Budiantara and Ratnasari, 2025).

Parameter significance testing was conducted to evaluate whether the estimated model parameters significantly contributed to the response variable. The significance of the individual parameters was evaluated using the t-test. This procedure is appropriate because, after the spline basis functions and knot locations are fixed, the spline regression model is linear in its parameters, and the parameters are estimated using the Ordinary Least Squares (OLS) method. Under the assumptions required for OLS-based inference, including independently and identically distributed errors with mean zero and constant variance, the estimated coefficients can be evaluated using t-statistics. This inference concerns the coefficients of the spline basis functions and does not imply that the underlying regression function itself is parametric. The t-test statistic is expressed as follows (Safitri *et al.*, 2025):

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (9)$$

where $\hat{\beta}_j$ denotes the estimated j -th regression parameter and $SE(\hat{\beta}_j)$ denotes its standard error. A parameter is considered statistically significant when its corresponding p-value is

smaller than the predetermined significance level α . Then, coefficient determination (R^2) used for measure how much big model capabilities in explain variations in variables response. The R^2 value shows proportion diversity of data that can be explained by the variables existing predictors in the model. The R^2 value is in the range of 0 to 1. A higher R^2 value indicates that the model explains a greater proportion of the variability in the data. Conversely, an R^2 value close to zero indicates that the model has limited ability to explain the variation in the response variable. In terms of mathematical, coefficient determination can written as following (Kaseside and Loklomin, 2021):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted value of the model, $\bar{y}$ is the average of the response variable and n is the number of observations.

3. METHODOLOGY

This study employed a quantitative approach to model the relationship between the response and predictor variables using nonparametric spline regression. The data used in this study were secondary data obtained from the South Sulawesi Provincial Health Office. The observational units consisted of 24 districts/cities in South Sulawesi Province in 2023, resulting in 24 observations. The response variable was the Maternal Mortality Rate (MMR), measured as the number of maternal deaths per 100,000 live births. The predictor variables consisted of Bleeding (X_1), Hypertension Disorders (X_2), Post-Miscarriage Complications (X_3), and Heart and Blood Vessel Disorders (X_4), measured as the number of maternal deaths attributed to each respective condition in each district/city. These predictor variables represent maternal health-related causes of maternal deaths reported in the same observation units. They were not used as components to calculate or construct the MMR response variable.

The data analysis was conducted through several sequential stages based on the nonparametric spline regression model presented in the previous section. The first stage involved descriptive analysis to examine the characteristics of the response and predictor variables. Data exploration was then performed using scatter plots to identify the initial patterns of the relationships between each predictor variable and MMR and to assess the suitability of a nonlinear modeling approach.

The next stage involved constructing the nonparametric spline regression model based on the formulation presented in Equation (1). The truncated spline basis functions were constructed according to Equation (2). For each predictor variable, one-, two-, and three-knot configurations were considered as candidate models. The optimal knot configuration was determined using the Generalized Cross-Validation (GCV) criterion formulated in Equation (3). The knot configuration with the smallest GCV value was selected as the optimal configuration for each predictor variable.

After the optimal knot points were determined, the corresponding spline basis matrix was constructed, and the model parameters were estimated using the least squares method as formulated in Equation (4). The fitted model was then evaluated to determine its statistical and predictive performance. The final stage involved evaluating the fitted model using the coefficient of determination (R^2) and the prediction error measures defined in the previous section. The selected model was subsequently interpreted to describe the relationships between Bleeding (X_1), Hypertension Disorders (X_2), Post-Miscarriage Complications (X_3), and Heart and Blood Vessel Disorders (X_4) and the Maternal Mortality Rate (MMR).

4. RESULT

The first stage of the analysis involved conducting descriptive statistics to provide an initial overview of the characteristics of the study variables. Descriptive statistics of the Maternal Mortality Rate (MMR) and the independent variables, namely Bleeding (X_1), Hypertension Disorders (X_2), Post-Miscarriage Complications (X_3), and Heart and Blood Vessel Disorders (X_4), are presented in Table 1.

Table 1. Statistics Descriptive Variables Study

Variable	Mean	Minimum	Maximum	Standard Deviation
Y	7.375	2	20	5.315
X_1	2.167	0	5	1.341
X_2	2.250	0	8	2.090
X_3	1.750	0	7	2.027
X_4	2.167	1	5	1.239

Table 1 presents the descriptive statistics of Maternal Mortality (MMR) and the four predictor variables. The mean Maternal Mortality was 7.375, with values ranging from 2 to 20 and a standard deviation of 5.315, indicating considerable variation in maternal mortality across the observed regions. Among the predictor variables, Hypertension Disorders (X_2) had the highest mean value of 2.250, followed by Bleeding (X_1) and Heart and Blood Vessel Disorders (X_4), both with a mean of 2.167. Post-Miscarriage Complications (X_3) had a mean value of 1.750. The variability of the predictor variables, as indicated by their standard deviations, ranged from 1.239 for Heart and Blood Vessel Disorders to 2.090 for Hypertension Disorders.

The next stage of the analysis was to examine the relationship between each independent variable and the Maternal Mortality Rate (MMR) using scatter plots. This step was conducted to identify the initial pattern of the relationship between the predictor variables and the response variable and to determine whether a linear or non-linear relationship should be considered in the subsequent modeling stage.

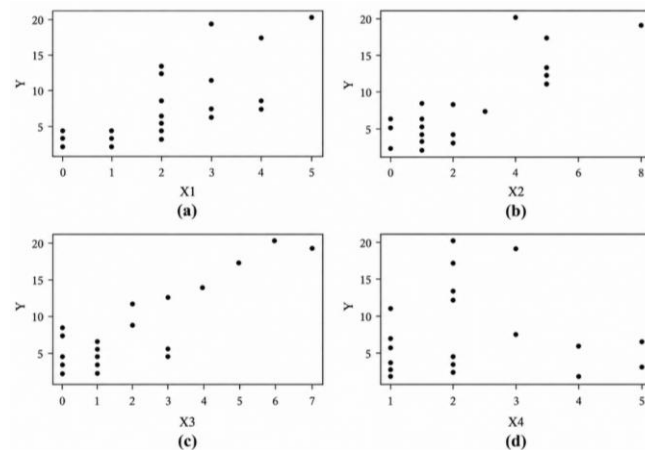


Figure 1. (a) Bleeding and MMR, (b) Hypertension Disorder and MMR, (c) Post-Miscarriage Complications and MMR, (d) Heart and Blood Vessel Disorders and MMR

Based on Figure 1, the scatter plot results between independent variables and the response variable, the relationships between X_1 , X_2 , X_3 , X_4 , and the Maternal Mortality Rate (MMR) show varying patterns across the observed values. Therefore, a nonparametric spline approach was applied to all four predictor variables to allow the model to accommodate potential changes in the relationships across different ranges of each predictor. The number and locations of the knot points for each predictor were subsequently determined using the Generalized Cross-Validation (GCV) criterion.

The next stage of the analysis was to determine the optimal knot points for the spline components. One-, two-, and three-knot models were considered as candidate models for each predictor variable, and the optimal knot configuration was selected based on the Generalized Cross-Validation (GCV) criterion. The knot configuration with the lowest GCV value was selected as the optimal configuration for each predictor variable. The first step in constructing the spline regression model was to determine the optimal knot point for each predictor variable. At this stage, the predictor variables, namely X_1 , X_2 , X_3 , and X_4 , were each fitted using a single knot point. The optimal knot point was determined by evaluating candidate knot locations using the Generalized Cross-Validation (GCV) criterion. The knot configuration yielding the minimum GCV value was selected as the optimal configuration for each predictor variable.

The use of the Generalized Cross-Validation (GCV) criterion for selecting optimal knot points is consistent with previous international studies on truncated spline regression. (Putra, Fadhlurrahman and Gunardi, 2023) used GCV to determine optimal knot points in geographically weighted truncated spline nonparametric regression and evaluated different knot configurations to obtain the model with the smallest GCV value. This approach supports the use of GCV as a criterion for selecting knot locations while considering model fit and complexity. The results of the optimization for the one-knot model are presented in Table 2.

Table 2. GCV Values for the One-Knot Model

Predictor Variable	Optimal Knot	Minimum GCV
X_1	4.0	16.4673
X_2	1.5	9.5369
X_3	3.0	7.7226
X_4	3.0	28.5077

Table 2 shows that the minimum GCV values for X_1 , X_2 , X_3 , and X_4 were 16.4673, 9.5369, 7.7226, and 28.5077, respectively. The corresponding optimal knot points were 4.0 for X_1 , 1.5 for X_2 , 3.0 for X_3 , and 3.0 for X_4 . These results represent the best one-knot configurations for each predictor variable based on the minimum GCV criterion. The next step was to determine the optimal knot points for the spline regression model with two knot points. At this stage, all candidate combinations of two knot points were evaluated for each predictor variable. The optimal combination was selected based on the Generalized Cross-Validation (GCV) criterion, with the best configuration corresponding to the minimum GCV value. The results of the optimization for the two-knot model are presented in Table 3.

Table 3. GCV Values for the Two-Knot Model

Predictor Variable	Optimal Knot Points	Minimum GCV
X_1	1.0, 4.0	17.8169
X_2	2.0, 4.4	8.9472
X_3	3.0, 4.7	7.6293
X_4	3.0, 4.0	30.5100

Table 3 shows that the minimum GCV values for X_1 , X_2 , X_3 , and X_4 were 17.8169, 8.9472, 7.6293, and 30.5100, respectively. The corresponding optimal knot points were 1.0 and 4.0 for X_1 , 2.0 and 4.4 for X_2 , 3.0 and 4.7 for X_3 , and 3.0 and 4.0 for X_4 . Compared with the one-knot model, the two-knot configuration produced a lower GCV value only for X_3 , whereas the GCV values for X_1 , X_2 , and X_4 increased. Therefore, based on the GCV criterion, the two-knot model provides an improvement only for X_3 . The final candidate model considered was the spline regression model with three knot points. As in the previous stages, all candidate combinations of three knot points were evaluated for each predictor

variable, and the optimal combination was determined based on the minimum GCV value. The optimization results for the three-knot model are presented in Table 4.

Predictor Variable	Optimal Knot Points	Minimum GCV
X_1	1.0, 3.0, 4.0	19.0424
X_2	3.0, 4.4, 5.0	5.7179
X_3	1.0, 2.1, 3.0	8.0144
X_4	2.0, 3.0, 4.0	33.5335

Table 4 shows that the minimum GCV values for X_1 , X_2 , X_3 , and X_4 were 19.0424, 5.7179, 8.0144, and 33.5335, respectively. The corresponding optimal knot points were 1.0, 3.0, and 4.0 for X_1 ; 3.0, 4.4, and 5.0 for X_2 ; 1.0, 2.1, and 3.0 for X_3 ; and 2.0, 3.0, and 4.0 for X_4 . Among the candidate models, the three-knot configuration produced the lowest GCV value for X_2 , indicating that the additional knot points improved the model's ability to capture the relationship between X_2 and the response variable. After evaluating the one-knot, two-knot, and three-knot models, the optimal knot configuration for each predictor variable was selected based on the minimum GCV value. The comparison of GCV values for the three candidate knot configurations is presented in Table 5.

Predictor Variable	One Knot	Two Knots	Three Knots	Optimal Configuration
X_1	16.4673	17.8169	19.0424	One knot
X_2	9.5369	8.9472	5.7179	Three knots
X_3	7.7226	7.6293	8.0144	Two knots
X_4	28.5077	30.5100	33.5335	One knot

Based on Table 5, the optimal knot configuration differs across the predictor variables. Bleeding (X_1) has the lowest GCV value under the one-knot model, with a GCV of 16.4673. Hypertension Disorders (X_2) achieves its lowest GCV value under the three-knot model, with a GCV of 5.7179. Meanwhile, Post-Miscarriage Complications (X_3) has the lowest GCV value under the two-knot model, with a GCV of 7.6293. Heart and Blood Vessel Disorders (X_4) has the lowest GCV value under the one-knot model, with a GCV of 28.5077. Therefore, the selected knot configuration consists of one knot for X_1 , three knots for X_2 , two knots for X_3 , and one knot for X_4 . The corresponding optimal knot points are 4 for X_1 ; 3.0, 4.4, and 5.0 for X_2 ; 3.0 and 4.7 for X_3 ; and 3 for X_4 . These selected knot points were subsequently used to construct the final spline regression model.

Based on the optimal knot configuration obtained in the previous subsection, the spline regression model with one knot for X_1 , three knots for X_2 , two knots for X_3 , and one knot for X_4 was selected as the final model. This model was then used to examine the contribution of each independent variable to the Maternal Mortality Rate (MMR). The results of the model parameter estimation are presented in Table 6. The use of parameter significance testing in truncated spline regression is also supported by previous research. (Ayu, Srinadi and Sukarsa, 2024) applied truncated spline nonparametric regression and conducted partial parameter significance testing after estimating the spline model. Their study reported coefficient estimates together with t-statistics and p-values for individual spline parameters, demonstrating the use of partial inference to evaluate the statistical significance of estimated spline coefficients.

Table 6. Spline Regression Model Parameter Estimation Results

Variable	Coefficient	t-count	p-value
----------	-------------	---------	---------

Intercept	1.1305	0.9858	0.3437
X_1 (Bleeding)	1.2660	4.5829	0.0006
X_2 (Hypertension Disorders)	0.2153	0.5021	0.6247
X_3 (Post-Miscarriage Complications)	0.3008	0.6443	0.5315
X_4 (Heart and Blood Vessel Disorders)	0.5015	0.7334	0.4774
B_{11} (Spline $X_1 - 4$)	3.0764	0.2927	0.7748
B_{21} (Spline $X_2 - 3.0$)	2.6441	2.0422	0.0637
B_{22} (Spline $X_2 - 4.4$)	-1.1082	-1.0060	0.3343
B_{23} (Spline $X_2 - 5.0$)	-1.3979	-0.2303	0.8217
B_{31} (Spline $X_3 - 3.0$)	1.4310	0.8065	0.4357
B_{32} (Spline $X_3 - 4.7$)	-0.8797	-0.0802	0.9374
B_{41} (Spline $X_4 - 3$)	-1.0731	-0.7326	0.4779

From Table 6, the spline regression model formed is:

$$\hat{Y} = 1.1305 + 1.2660X_1 + 0.2153X_2 + 0.3008X_3 + 0.5015X_4 + 3.0764(X_1 - 4)_+ + 2.6441(X_2 - 3.0)_+ - 1.1082(X_2 - 4.4)_+ - 1.3979(X_2 - 5.0)_+ + 1.4310(X_3 - 3.0)_+ - 0.8797(X_3 - 4.7)_+ - 1.0731(X_4 - 3)_+$$

The partial test results indicate that Bleeding (X_1) is the only predictor variable that is statistically significant at the 5% significance level, with a p-value of 0.0006. The positive coefficient of X_1 (1.2660) indicates a positive association between the linear component of Bleeding and the estimated MMR. Meanwhile, Hypertension Disorders (X_2), Post-Miscarriage Complications (X_3), and Heart and Blood Vessel Disorders (X_4) are not statistically significant at the 5% level, with p-values of 0.6247, 0.5315, and 0.4774, respectively. None of the spline components is statistically significant at the 5% significance level. Among the spline components, the spline component of X_2 at the knot point 3.0 has the smallest p-value, namely 0.0637, but this value remains above the 0.05 significance level. Therefore, there is insufficient statistical evidence to conclude that the individual spline components are statistically significant at the 5% level.

The results of the partial significance test are also supported by recent research on statistical inference in nonparametric regression. (Ramli, Budiantara and Ratnasari, 2024) developed a statistical inference framework for nonparametric regression and specifically addressed partial hypothesis testing for individual model parameters. These findings support the use of partial significance testing for evaluating individual coefficients in the present truncated spline regression model, provided that the underlying assumptions of the inferential procedure are satisfied. After obtaining the final spline regression model, the next stage was to evaluate its performance in explaining and predicting the Maternal Mortality Rate (MMR). Model evaluation was conducted using the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2). The spline regression model obtained an MSE value of 0.9388 and an RMSE value of 0.9689. These values indicate that the difference between the observed and estimated MMR values is relatively small. The model also produced a MAPE value of 15.6494%, indicating that the average absolute percentage error of the model predictions is approximately 15.65%. Furthermore, the coefficient of determination (R^2) was 0.9653, indicating that approximately 96.53% of the variation in MMR can be explained by the predictor variables and spline components included in the model, while the remaining 3.47% is associated with factors not included in the model. These results are consistent with previous studies applying truncated spline nonparametric regression. Saidah et al. (2022) evaluated a multivariable truncated spline model using the coefficient of determination (R^2) and Mean Absolute Percentage Error (MAPE) after selecting the optimal knot configuration using GCV.

5. CONCLUSION

Based on the analysis, the nonparametric spline regression model provided a flexible approach for modeling the Maternal Mortality Rate (MMR) across 24 districts/cities in South Sulawesi in 2023. Based on the Generalized Cross-Validation (GCV) criterion, the optimal knot configuration consisted of one knot for Bleeding (X_1), three knots for Hypertension Disorders (X_2), two knots for Post-Miscarriage Complications (X_3), and one knot for Heart and Blood Vessel Disorders (X_4). The corresponding knot points were 4 for X_1 ; 3.0, 4.4, and 5.0 for X_2 ; 3.0 and 4.7 for X_3 ; and 3 for X_4 . The parameter significance test showed that Bleeding (X_1) was the only predictor with a statistically significant linear component at the 5% significance level, with a p-value of 0.0006. The other linear components and all spline components were not statistically significant at the 5% significance level.

The selected model achieved an MSE of 0.9388, an RMSE of 0.9689, a MAPE of 15.6494%, and an R^2 of 0.9653. The R^2 value indicates that 96.53% of the variation in MMR is explained by the predictor variables and spline components included in the model. Among the spline components, the component of X_2 at the knot point 3.0 had the smallest p-value (0.0637), but it remained above the 0.05 significance level. Overall, the results indicate that nonparametric spline regression can provide a flexible framework for modeling the relationship between MMR and the selected maternal health-related conditions.

DAFTAR PUSTAKA

- Abdy, M. (2019) "Tinjauan Singkat Tentang Regresi Parametrik dan Regresi non Parametrik," *Saintifik*, 5(1), pp. 58–62. Available at: <https://doi.org/10.31605/saintifik.v5i1.199>.
- Annas, S., Aswi, A. and Hidayat, R. (2026) "Determinants of Maternal Mortality in Indonesia: A B-Spline Nonparametric Regression Approach to Identify Nonlinear Relationship Patterns," *MATHLINE: Jurnal Matematika dan Pendidikan Matematika*, 11(1), pp. 31–44. <https://doi.org/10.31943/mathline.v11i1.1015>.
- Ariesta, D., Gusriani, N. and Parmikanti, K. (2021) "Estimasi Parameter Model Regresi Nonparametrik B-Spline E Pada Angka Kematian Maternal," *Jurnal Matematika UNAND*, 10(3), pp. 342–354. Available at: <https://doi.org/10.25077/jmu.10.3.342-354.2021>.
- Arifanti, D.R. and Hidayat, R. (2021) "Statistical Inference on Nonparametric Spline Models and its Applications," *Journal of Physics: Conference Series*, 2123(1–9). Available at: <https://doi.org/10.1088/1742-6596/2123/1/012023>.
- Aswi, A., Sanusi, W., Tiro, M.A., Sukarna, S., Haekal, M.F., Palarungi, A.G., Putri, S.C.A. and Oktaviana, O. (2025) "Evaluasi Performa Model Regresi Poisson Tweedie dan Conway Maxwell Poisson dalam Menangani Masalah Dispersi: Studi Angka Kematian Ibu di Provinsi Sulawesi Selatan," *Indonesian Journal of Fundamental Sciences (IJFS)*, 11(2), pp. 122–134. Available at: <https://doi.org/10.26858/ijfs.v11i2.77506>.
- Ayu, I.G., Srinadi, M. and Sukarsa, I.K.G. (2024) "Pemodelan Regresi Nonparametrik Spline Truncated Kasus Kejadian Diare di Provinsi Bali," *E-Jurnal Matematika*, 13(1), pp. 28–37.
- Badan Pusat Statistik (2023) *Analisis Tematik Kependudukan Indonesia: Fertilitas Remaja, Kematian Maternal, Kematian Bayi, dan Penyandang Disabilitas*. Badan Pusat Statistik.
- Badan Pusat Statistik (2024) *Profil Kesehatan Ibu dan Anak*. Badan Pusat Statistik.
- Fadhilah, R.N., Islamiyati, A. and Nirwan (2025) "Analysis of maternal mortality using geographically weighted Poisson regression with spline truncated estimator,"

- Communications in Mathematical Biology and Neuroscience, 2025(2025 AR–95). Available at: <https://doi.org/10.28919/cmbn/9195>.
- Handajani, S.S., Pratiwi, H. and Qona, N. (2023) “Comparison of B-Spline and Truncated Spline Regression Models for Temperature Forecast,” *Barekeng: Journal of Mathematics and Its Applications*, 17(4), pp. 1969–1984.
- Hesikumalasari, H., Budiantara, I.N., Ratnasari, V. and Nisa', K. (2022) “Estimation of Semiparametric Regression Curve with Mixed Estimator of Multivariable Linear Truncated Spline and Multivariable Kernel,” *Media Statistika*, 15(1), pp. 12–23 Available at: <https://doi.org/10.14710/medstat.15.1.12-23>.
- Hidayat, R., Budiantara, I.N., Otok, B.W. and Ratnasari, V. (2021) “The Regression Curve Estimation by Using Mixed Smoothing Spline and Kernel (MsS-K) Model,” *Communications in Statistics - Theory and Methods*, 50(17), pp. 3942–3953. Available at: <https://doi.org/10.1080/03610926.2019.1710201>.
- Hidayat, R., Annas, S., Aswi, A., Putri, S.C.A. and Vivianti, V. (2025) “Pendekatan Regresi Nonparametrik Spline Truncated untuk Mengidentifikasi Determinan Angka Kematian Ibu di Indonesia,” *Indonesian Journal of Fundamental Sciences*, 11(2), pp. 113–121. <https://doi.org/10.26858/ijfs.v11i2.77643>.
- Hidayat, R. and Ilyas, M. (2023) “Spline Model in The Case of Cervical Cancer Patient Resilience,” *Trends in Sciences*, 20(8), pp. 1–14.
- Hidayat, R., Muthahharah, I. and Resmawan, R. (2025) “Kernel-Truncated Spline : Estimator Fleksibel untuk Regresi Nonparametrik,” *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, 13(2), pp. 271–277.
- Kaseside, M. and Loklomin, S.B. (2021) “Analisis Angka Kematian Bayi di Kabupaten Halmahera Utara dengan Metode Regresi Nonparametrik Spline Truncated,” *Variance: Journal of Statistics and Its Applications*, 3(1), pp. 1–5.
- Muthahharah, I. and Hidayat, R. (2025) “The Impact of Malnutrition on Infant Mortality Rate in Indonesia : A Spline Regression Approach,” *Journal of Mathematics, Computations, and Statistics*, 8(2), pp. 474–481.
- Nurhuda, G., Wasono and Nohe, D.A. (2022) “Pemodelan Regresi Nonparametrik Berdasarkan Estimator Spline Truncated pada Data Simulasi,” *Jurnal Matematika, Statistika dan Komputasi*, 19(1), pp. 172–182. Available at: <https://doi.org/10.20956/j.v19i1.21534>.
- Puspitasari, A.A., Fernandes, A.A.R., Efendi, A., Astutik, S. and Sumarminingsih, E. (2023) “Development of Nonparametric Truncated Spline at Various Levels of Autocorrelation of Longitudinal Generating Data,” *Journal of Statistics Applications & Probability*, 12(2), pp. 757–766. <https://doi.org/10.18576/jsap/120234>
- Putra, R., Fadhlurrahman, M.G. and Gunardi (2023) “Determination of the best knot and bandwidth in geographically weighted truncated spline nonparametric regression using generalized cross validation,” *MethodsX*, 10(January). Available at: <https://doi.org/10.1016/j.mex.2022.101994>.
- Qomari, Y.A.N. and Notobroto, H.B. (2022) “Kunjungan Ibu Hamil K1 Dan K4 Terhadap Angka Kematian Ibu Di Provinsi Jawa Timur,” *Preventif: Jurnal Kesehatan Masyarakat*, 13(4), pp. 586–595. Available at: <https://doi.org/https://doi.org/10.22487/preventif.v13i4.309>.
- Ramadhan, M.R., Nohe, D.A. and Wahyuningsih, S. (2023) “Truncated Spline Nonparametric Regression Model on Human Development Index in Indonesia,” *Jurnal EKSPONENSIAL*, 14(November), pp. 91–98.
- Ramli, M., Budiantara, I.N. and Ratnasari, V. (2024) “Statistical inference on nonparametric regression model with approximation of Fourier series function: Estimation and

- hypothesis testing,” *MethodsX*, 13, p. 102922. Available at: <https://doi.org/10.1016/j.mex.2024.102922>.
- Raina, N., Khanna, R., Gupta, S., Jayathilaka, C.A., Mehta, R. and Behera, S. (2023) “Progress in Achieving SDG Targets for Mortality Reduction among Mothers, Newborns, and Children in the WHO South-East Asia Region,” *The Lancet Regional Health – Southeast Asia*, 18, Article 100307. Available at: <https://doi.org/10.1016/j.lansea.2023.100307>.
- Safitri, U.N., Sriliana, I., Adelisa, R., Hafiz, M. and Novianti, P. (2025) “Perbandingan Regresi Nonparametrik Spline Truncated dan Kernel Gaussian dalam Menganalisis Faktor-Faktor Penentu Indeks Pembangunan Manusia (IPM) di Indonesia,” *Jurnal Gaussian*, 14(2), pp. 554–564. Available at: <https://doi.org/10.14710/j.gauss.14.2.554-564>.
- Saidah, L.M., Suparti and Sudarno (2022) “Pemodelan Faktor Ekonomi Makro terhadap Harga Saham Telkom Menggunakan Regresi Spline Truncated Multivariabel Dilengkapi GUI R,” 11, pp. 344–354. Available at: <https://doi.org/10.14710/J.GAUSS.11.3.334-354>.
- Sampurna, M.T.A., Handayani, K.D., Utomo, M.T., Angelika, D., Etika, R., Harianto, A., Mapindra, M.P., Mahindra, M.P., Efendi, F., Kaban, R.K., Rohsiswatmo, R., Visuddho, V. and Permana, P.B.D. (2023) “Determinants of Neonatal Deaths in Indonesia: A National Survey Data Analysis of 10,838 Newborns,” *Heliyon*, 9(1), Article e12980. Available at: <https://doi.org/10.1016/j.heliyon.2023.e12980>.
- Sanusi, W., Syam, R. and Adawiyah, R. (2019) “Model Regresi Nonparametrik dengan Pendekatan Spline (Studi Kasus: Berat Badan Lahir Rendah di Rumah Sakit Ibu dan Anak Siti Fatimah Makassar),” *Journal of Mathematics, Computations, and Statistics*, 2(1), p. 70. Available at: <https://doi.org/10.35580/jmathcos.v2i1.12460>.
- Sriliana, I., Budiantara, I.N. and Ratnasari, V. (2022) “A Truncated Spline and Local Linear Mixed Estimator in Nonparametric Regression for Longitudinal Data and Its Application,” *Symmetry*, 14(12), pp. 1–19.
- Suriaslan, A.S., Budiantara, I.N. and Ratnasari, V. (2025) “Nonparametric Regression Estimation Using Multivariable Truncated Splines for Binary Response Data,” *MethodsX*, 14, p. 103084. Available at: <https://doi.org/https://doi.org/10.1016/j.mex.2024.103084>.
- Syamsualam, N. and Hidayat, R. (2022) “Application Of Truncated Spline Nonparametric RegreSyamsualam, N. and Hidayat, R. (2022) “Application Of Truncated Spline Nonparametric Regression In Modeling Traffic Accident Rate In Palopo City Penerapan Regresi Nonparametrik Spline Truncated Dalam Memo,” *Mathline : Jurnal Matematika dan Pendidikan Matematika*, 7(2), pp. 185–196. Available at: <https://doi.org/10.31943/mathline.v7i2.275>.