RESEARCH ARTICLE • OPEN ACCESS

Nonparametric Hybrid Spline-Uniform Kernel Regression Modeling in Indonesian Malaria Cases

Nur Salsabila Wero, Rahmat Hidayat, Rika Saniarti, BQ. Nurkhalishah K, Ruliana Ruliana, Aswi Aswi

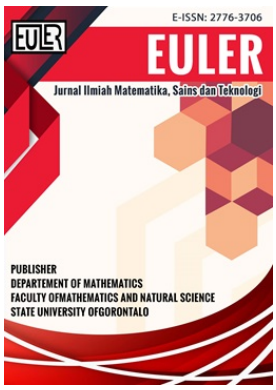
Volume 14, Issue 3, Pages 646–661, December 2026

Received 6 August 2026 , Revised 12 November 2026, Accepted 18 November 2026, Published 20 November 2026

To Cite this Article : N. S. Wero, et al., "Nonparametric Hybrid Spline-Uniform Kernel Regression Modeling in Indonesian Malaria Cases", *Euler J. Ilm. Mat. Sains dan Teknol.*, vol. 14, no. 3, pp. 646–661, 2026, doi: *https://doi.org/10.37905/euler.v14i3.40773*

© 2026 by author(s)

JOURNAL INFO • EULER : JURNAL ILMIAH MATEMATIKA, SAINS DAN TEKNOLOGI

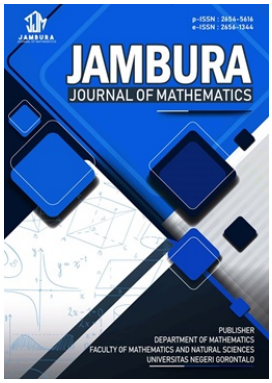


Homepage	:	http://ejurnal.ung.ac.id/index.php/euler/index
Journal Abbreviation	:	Euler J. Ilm. Mat. Sains dan Teknol.
Frequency	:	Three times a year
Publication Language	:	English
DOI	:	https://doi.org/10.37905/euler
Online ISSN	:	2776-3706
License	:	Creative Commons Attribution-NonCommercial 4.0 International License
Publisher	:	Department of Mathematics, Universitas Negeri Gorontalo
Country	:	Indonesia
OAI Address	:	http://ejurnal.ung.ac.id/index.php/euler/oai
Google Scholar ID	:	QF_r-_gAAAAJ
Email	:	euler@ung.ac.id

JAMBURA JOURNAL • FIND OUR OTHER JOURNALS



Jambura Journal of Biomathematics



Jambura Journal of Mathematics



Jambura Journal of Mathematics Education



Jambura Journal of Probability and Statistics

Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi, Volume 14, Issue 3, pp. 646–661, December 2026 *https://doi.org/10.37905/euler.v14i3.40773*

Research Article

Check for updates

Nonparametric Hybrid Spline-Uniform Kernel Regression Modeling in Indonesian Malaria Cases

Nur Salsabila Wero[1], Rahmat Hidayat[2,*], Rika Saniarti[1], BQ. Nurkhalishah K[1], Ruliana Ruliana[2], Aswi Aswi[2]

[1]*Postgraduate Student in Statistics, Makassar State University, Makassar, 90224, Indonesia*
[2]*Department of Statistics, Makassar State University, Makassar, 90224, Indonesia*

ARTICLE HISTORY
Received 6 August 2026
Revised 12 November 2026
Accepted 18 November 2026
Published 20 November 2026

KEYWORDS
Generalized Cross Validation
Hybrid Spline–Uniform Kernel
Malaria
Nonparametric Regression
Truncated Spline

ABSTRACT. *Malaria remains a significant public health concern in Indonesia, with the number of reported cases varying across provinces. This variation indicates that the relationship between malaria cases and their associated risk factors is complex, making conventional parametric regression models less suitable. This study aims to develop a Hybrid Spline–Uniform Kernel nonparametric regression model, determine the optimal knot points and bandwidths using the Generalized Cross Validation (GCV) criterion, and identify the best model for analyzing malaria cases in Indonesia. Secondary data from 2024 covering all 38 provinces were analyzed. The response variable was the number of malaria cases, while the predictor variables included population density, the percentage of the population living in poverty, the percentage of integrated health service coverage, the Gini ratio, and the percentage of households with access to improved sanitation. Population density, poverty, and integrated health service coverage were modeled using truncated spline functions, whereas the Gini ratio and sanitation access were modeled using the Uniform kernel approach. The results showed that the best model was the Hybrid Spline–Uniform Kernel model with three knot points, producing the lowest GCV value of 0.0801371. The optimal bandwidths were 2.3105260 for the Gini ratio and 1.3368420 for sanitation access. These findings demonstrate that the proposed model provides a good fit for modeling malaria cases in Indonesia and offers an effective alternative for data with complex relationships.*

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonComercial 4.0 International License. ***Editorial of EULER:*** *Department of Mathematics, Universitas Negeri Gorontalo, Jln. Prof. Dr. Ing. B. J. Habibie, Bone Bolango 96554, Indonesia.*

1. Introduction

Regression analysis is one of the most widely used statistical methods for modeling the relationship between a response variable and one or more predictor variables [1, 2]. Regression models play an important role in describing the relationships among variables, identifying the effects of predictor variables on the response variable, and providing a basis for estimation and prediction. Based on the functional form of the regression model, regression analysis can be classified into parametric, nonparametric, and semiparametric regression [3, 4]. However, in many real-world applications, the relationships among variables often do not follow a specific functional form, making parametric regression approaches less appropriate. Consequently, nonparametric regression has emerged as a flexible alternative because it does not require a predetermined regression function [5, 6].

Nonparametric regression is capable of constructing regression curves that adapt to the characteristics of the observed data, making it suitable for modeling complex and nonlinear relationships. Several estimators have been developed in nonparametric regression, including truncated spline, kernel, Fourier, and local polynomial estimators, each possessing distinct

*Corresponding Author.

Email : *rahmat.hidayat@unm.ac.id.* (R. Hidayat)
Homepage : http://ejurnal.ung.ac.id/index.php/euler/index / E-ISSN : 2776-3706
© 2026 by the Author(s).

characteristics that should be selected according to the underlying relationship between the response and predictor variables [7]. Among these, the truncated spline estimator has been widely applied because it effectively models data exhibiting structural changes at specific knot points [8, 9]. Its primary advantage lies in its ability to capture local variations, thereby producing a more flexible regression model. Nevertheless, the performance of the truncated spline estimator depends heavily on the number and locations of the selected knot points, requiring an appropriate optimization procedure [10].

Another widely used estimator in nonparametric regression is the kernel estimator, which is particularly suitable for modeling smooth relationships among variables. Kernel estimators construct regression curves by assigning weights to observations within the neighborhood of a target point through a kernel function [11]. One commonly used kernel function is the Uniform kernel, which assigns equal weights to all observations within the bandwidth interval and zero weight to observations outside the interval [12]. This characteristic makes the Uniform kernel computationally simple while providing stable estimates for data with relatively smooth patterns [13, 14]. Similar to spline estimation, the performance of kernel estimation is strongly influenced by the choice of bandwidth, which determines the degree of smoothing applied to the data [15].

The distinct characteristics of truncated spline and kernel estimators indicate that each is better suited to different types of data patterns. Truncated spline estimators are more appropriate for data exhibiting abrupt changes at particular points, whereas kernel estimators perform better when relationships vary continuously and smoothly [14]. These complementary characteristics have encouraged the development of hybrid nonparametric regression models that combine both estimators within a single framework, allowing each predictor variable to be modeled using the estimator that best reflects its underlying relationship with the response variable [11, 16].

Hybrid nonparametric regression models have been increasingly developed to improve modeling flexibility. Previous studies have shown that Hybrid Kernel–Truncated Spline models incorporating the Gaussian kernel produce lower Generalized Cross Validation (GCV) values than models based on a single estimator [17]. Likewise, the Mixed Smoothing Spline–Kernel approach has demonstrated improved capability in capturing nonlinear relationships in multivariate data [11]. Furthermore, comparative studies have demonstrated the effectiveness of truncated spline regression over Multivariate Adaptive Regression Splines (MARS) under specific data conditions. In particular, truncated spline regression achieved lower Mean Squared Error (MSE) values than MARS in simulation studies [18]. Similar findings were reported in an application involving the Environmental Quality Index in Indonesia, where truncated spline regression demonstrated better model performance than MARS [19].

Although hybrid spline–kernel regression models have been widely investigated [3, 11, 14, 17], further investigation is needed into the use of alternative kernel functions within this framework. In particular, the Uniform kernel offers a compact-support weighting scheme in which observations within the bandwidth receive equal weights, while observations outside the bandwidth receive zero weights. This weighting characteristic provides a different local estimation mechanism from the Gaussian kernel, which assigns continuously decreasing weights according to the distance between observations. Therefore, examining the Uniform kernel in a hybrid spline–kernel framework provides an opportunity to evaluate its suitability for capturing local patterns in the data.

In this study, a Hybrid Spline–Uniform Kernel regression model is developed and applied to public health data. The estimation procedure uses the Generalized Cross Validation (GCV)

criterion to determine the spline knot configuration and kernel bandwidth through a sequential optimization process. The contribution of this study lies in investigating the applicability and estimation performance of the Uniform kernel within the hybrid spline framework, particularly in modeling public health outcomes. The resulting model is evaluated based on its predictive and estimation accuracy to assess its effectiveness for the data under consideration.

Malaria remains one of the major public health problems in Indonesia [20]. The disease is caused by *Plasmodium* parasites transmitted through the bites of infected female *Anopheles* mosquitoes and continues to pose a challenge to the national malaria elimination program [21]. Malaria morbidity is commonly measured using the Annual Parasite Incidence (API) per 1,000 population, which is calculated as the proportion of confirmed malaria cases among the population at risk. Indonesia successfully reduced the API to below one case per 1,000 population during 2015–2020. However, since 2021 the API has increased, reaching 1.9 per 1,000 population in 2024. During the same year, malaria surveillance recorded 4,353,097 examinations and 543,965 confirmed malaria cases, indicating that malaria remains a serious public health concern requiring effective prevention and control measures [22].

The distribution of malaria cases across provinces in Indonesia reflects the interaction of socioeconomic, demographic, environmental, and healthcare-related factors. Previous ecological research in Indonesia has identified poverty and access to improved sanitation as factors associated with malaria incidence at the provincial level [23]. Furthermore, research conducted in endemic areas has highlighted the importance of healthcare accessibility, socioeconomic conditions, and housing characteristics in understanding malaria prevalence [24]. Population density may also represent an important demographic characteristic in examining malaria distribution, although its relationship with transmission can vary depending on local environmental and epidemiological conditions. In this study, the percentage of integrated health service coverage, the Gini ratio, and the percentage of households with access to improved sanitation are considered indicators of healthcare accessibility and socioeconomic and environmental conditions that may be associated with variations in malaria cases [23, 24].

The nonlinear relationship between malaria cases and their associated factors has also been investigated using nonparametric regression approaches. Previous studies have applied Least Square Spline estimation to model malaria prevalence in Indonesia, identifying a significant relationship between access to improved sanitation and malaria morbidity [20]. Similarly, kernel smoothing has been employed in quantile regression to investigate factors associated with malaria distribution in Indonesia, including sanitation conditions, health-related behavior, and healthcare resources [25]. These studies demonstrate the applicability of spline-based and kernel-based nonparametric methods in analyzing malaria-related outcomes. Therefore, the use of a flexible hybrid regression framework provides an alternative approach to investigating potentially nonlinear relationships between malaria cases and their associated factors.

Therefore, the Hybrid Spline–Uniform Kernel nonparametric regression approach is considered appropriate because it accommodates different predictor characteristics by combining truncated spline estimation for variables with local structural changes and Uniform kernel estimation for variables exhibiting smooth relationships. In addition, the simultaneous optimization of knot points and bandwidths using the Generalized Cross Validation criterion is expected to improve model accuracy. Accordingly, this study aims to develop a Hybrid Spline–Uniform Kernel nonparametric regression model for malaria cases in Indonesia, determine the optimal knot points and bandwidths using the GCV criterion, and identify the best model based on the minimum GCV value.

2. Methods

2.1. Data and Data Sources

This study utilized secondary cross-sectional data collected from all 38 provinces in Indonesia. The dataset comprised one response variable and five predictor variables derived from official publications issued by Statistics Indonesia (BPS) and the Ministry of Health of the Republic of Indonesia for the year 2024 [22, 26]. The definitions and descriptions of the variables used in this study are presented in Table Table 1.

Table 1. Description of the research variables.

Variable	Description	Unit
Number of malaria cases (Y)	Total reported malaria cases in each province	Cases
Population density (X_1)	Number of inhabitants per square kilometer	Persons/km^2
Percentage of the population living in poverty (X_2)	Proportion of the population living below the poverty line	Percent (%)
Percentage of active Posyandu (X_3)	Proportion of active Posyandu relative to the total registered Posyandu in each province	Percent (%)
Gini ratio (X_4)	Measure of income inequality	Index
Percentage of households with access to improved sanitation (X_5)	Proportion of households with access to improved sanitation facilities	Percent (%)

2.2. Analysis Procedure

2.2.1. Development of the Hybrid Spline–Uniform Kernel Nonparametric Regression Model

The Hybrid Spline–Uniform Kernel nonparametric regression model combines the truncated spline estimator and the Uniform kernel estimator [27, 28]. Truncated spline regression is a flexible nonparametric approach for modeling nonlinear relationships by combining polynomial functions with truncated basis functions at selected knot points [29]. Predictor variables exhibiting structural changes at specific points are modeled using truncated spline functions, whereas predictor variables with smooth functional relationships are modeled using the Uniform kernel estimator [30, 31]. Kernel regression estimates the response locally through a weighting function, while the bandwidth controls the degree of smoothing and the number of observations contributing to each local estimate [32]. The combination of different regression components provides flexibility in modeling relationships that may not be adequately represented by a single functional form [33]. Accordingly, the hybrid regression model is expressed as follows:

$$y_i = f(x_i) + g(t_i) + \varepsilon_i, \tag{1}$$

Where:

$$f(x_i) = \sum_{j=1}^{3} \beta_j x_{ji} + \sum_{j=1}^{3} \sum_{k=1}^{r} \phi_{jk} \left(x_{ji} - K_{jk}\right)_+, \tag{2}$$

The terms in Eq. (2) are defined as follows: $f(x_i)$ represents the spline regression function, β_j denotes the regression coefficient of the j-th predictor, ϕ_{jk} represents the coefficient of the k-th truncated spline basis for the j-th predictor, K_{jk} denotes the k-th knot of the j-th predictor, and r represents the number of knots assigned to each spline predictor. The truncated function $(x_{ji} - K_{jk})_+$ is defined as $\max(0, x_{ji} - K_{jk})$. The predictor variables are standardized as follows:

$$x_{ji}^* = \frac{x_{ji} - \bar{x}_j}{s_j}, \tag{3}$$

The standardized predictor x_{jt}^* is obtained by subtracting the sample mean $\bar{x}_j$ from the original predictor value x_{ji} and dividing the result by its sample standard deviation s_j. The response variable was also standardized using the z-score transformation as follows:

$$y_i^* = \frac{y_i - \bar{y}}{s_y}, \tag{4}$$

where y_i represents the original response value, $\bar{y}$ denotes the sample mean, and s_y is the sample standard deviation of the response variable. Therefore, all model estimation procedures were performed using standardized response and predictor variables. For interpretation, the estimated knot locations were transformed back into their original measurement units using:

$$K_{jk}^{\text{original}} = \bar{x}_j + s_j K_{jk}^*, \tag{5}$$

where K_{jk}^* represents the standardized knot location, $\bar{x}_j$ is the sample mean, and s_j is the sample standard deviation of the corresponding predictor variable.

and

$$g(t_i) = \sum_{j=1}^{n} V_h(t_i, t_j) y_j. \tag{6}$$

Here, $\mathbf{t}_i = (v_{1i}, v_{2i})$ represents the two standardized kernel predictors, while $\mathbf{h} = (h_1, h_2)$ denotes their respective bandwidth parameters. The Uniform kernel function is defined as:

$$K(u) = \begin{cases} \frac{1}{2}, & |u| \le 1, \\ 0, & |u| > 1. \end{cases} \tag{7}$$

The standardized distances for the two kernel predictors are defined as:

$$u_{1ij} = \frac{v_{1i} - v_{1j}}{h_1}, \qquad u_{2ij} = \frac{v_{2i} - v_{2j}}{h_2}. \tag{8}$$

Here, u_{1ij} and u_{2ij} denote the standardized distances between the i-th and j-th observations for the two kernel predictors, respectively, while h_1 and h_2 represent their corresponding bandwidth parameters, expressed in standardized units. The two kernel predictors are combined using a Product Uniform Kernel, defined as:

$$W_{ij} = K(u_{1ij}) K(u_{2ij}). \tag{9}$$

The normalized kernel weights are calculated using the Nadaraya–Watson weighting scheme:

$$V_{ij} = \frac{W_{ij}}{\sum_{\ell=1}^{n} W_{i\ell}}. \tag{10}$$

Here, W_{ij} denotes the product kernel weight, while V_{ij} represents the normalized Nadaraya–Watson kernel weight used to construct the smoothing matrix V.

2.2.2. Matrix Representation of the Model

The model in Eq. (1) can be expressed in matrix form as

$$y = G\theta + Vy + \varepsilon, \tag{11}$$

where y denotes the response vector, G denotes the truncated spline basis matrix, $\boldsymbol{\theta}$ represents the vector of spline parameters, V is the Uniform kernel weighting matrix, and $\boldsymbol{\varepsilon}$ is the random error vector. The Uniform kernel weighting matrix is constructed from the Uniform kernel function and subsequently normalized to obtain the kernel smoothing matrix.

2.2.3. Estimation of Spline Parameters

The spline parameters were estimated using the Ordinary Least Squares (OLS) method by minimizing the following objective function [34]:

$$\min_{\boldsymbol{\theta}} Q(\boldsymbol{\theta})^T Q(\boldsymbol{\theta}), \tag{12}$$

Where:

$$Q(\boldsymbol{\theta}) = (I - V)y - G\boldsymbol{\theta}. \tag{13}$$

The corresponding first-order condition is obtained by differentiating the objective function with respect to the spline parameter vector:

$$\frac{\partial[Q(\boldsymbol{\theta})^T Q(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}} = 0. \tag{14}$$

2.2.4. Parameter Estimation

The parameter estimator was obtained by differentiating the objective function with respect to the spline parameter vector and setting the derivative equal to zero [35]:

$$\frac{\partial Q(\theta)}{\partial \theta} = 0, \tag{15}$$

The resulting estimator is

$$\widehat{\theta} = (G^T G)^{-1} G^T (I - V)y. \tag{16}$$

This estimator was subsequently used to estimate the spline component of the Hybrid Spline–Uniform Kernel model.

2.2.5. Construction of the Smoothing Matrix

Based on the estimated spline parameters, the smoothing matrix for the spline component is computed as [35]:

$$S = G(G^T G)^{-1} G^T (I - V). \tag{17}$$

The model smoothing matrix is then expressed as

$$Z = S + V. \tag{18}$$

The matrix represents the overall smoothing matrix that combines the truncated spline and Uniform kernel components.

2.2.6. Estimation of the Response Variable

The fitted values of the response variable are obtained as

$$\hat{y} = Zy. \tag{19}$$

The fitted values are subsequently used to compute the residuals and evaluate the predictive performance of the model.

2.2.7. Determination of the Optimal Knot Points and Bandwidth

The optimal knot points for the spline component were determined using a grid search procedure. Specifically, 20 candidate knot locations were generated at equally spaced intervals between the 10th and 90th percentiles of each standardized spline predictor. Models with one, two, and three knot points per spline predictor were evaluated, with the same

number of knots assigned to all three spline predictors in each model configuration. After determining the optimal knot configuration, the bandwidth parameters for the Uniform kernel component were selected using a grid search procedure. The bandwidth parameters, h_1 and h_2, were evaluated over a set of candidate values expressed in standardized units. For each candidate bandwidth combination, the Hybrid Spline–Uniform Kernel model was estimated while keeping the selected knot configuration fixed. The model performance for each candidate configuration was evaluated using the Generalized Cross-Validation (GCV) criterion, which provides a basis for selecting the smoothing level while accounting for model complexity [34–36]. The GCV criterion is calculated as follows:

$$GCV = \frac{MSE}{\left(1 - \frac{\text{tr}(Z)}{n}\right)^2}. \tag{20}$$

where MSE represents the mean squared error, tr(Z) denotes the trace of the overall smoothing matrix, and n is the number of observations. The optimal model configuration was determined based on the minimum GCV value.

2.2.8. Model Evaluation

The final model was evaluated using the Leave-One-Out Cross-Validation (LOOCV) procedure to obtain an unbiased estimate of its predictive performance. The prediction errors were assessed using the Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). The coefficient of determination (R^2) was also reported to describe the proportion of variability in the response variable explained by the model. Since the response variable was standardized and may contain zero or negative values, the Mean Absolute Percentage Error (MAPE) was not used as a performance measure. Model selection was primarily based on the GCV criterion, while LOOCV-based MSE and RMSE were used to assess predictive performance.

3. Results and Discussion

3.1. Overview of the Data

This study employed secondary data from 2024, with the 38 provinces of Indonesia serving as the units of analysis. The response variable was the number of malaria cases (Y), while the predictor variables included population density (X_1), the percentage of the population living in poverty (X_2), the percentage of integrated health service coverage (X_3), the Gini ratio (X_4), and the percentage of households with access to improved sanitation (X_5). Descriptive statistics for all study variables are presented in Table Table 2.

Table 2. Descriptive statistics of the study variables.

Variable	Minimum	Mean	Maximum	Variance	Standard Deviation
Number of malaria cases (Y)	3.00	14,314.90	229,786.00	2,115,982,790.00	45,999.81
Population density (X_1)	5.00	678.20	16,165.00	6,809,210.00	2,609.45
Percentage of the population living in poverty (X_2)	3.80	10.67	29.66	40.5356	6.3668
Percentage of integrated health service coverage (X_3)	0.00	76.99	100.00	873.4091	29.5535
Gini ratio (X_4)	0.2350	0.3406	0.4310	0.002281	0.0478
Percentage of households with access to improved sanitation (X_5)	12.61	81.14	96.83	227.1340	15.0710

Based on Table Table 2, the descriptive statistics indicate considerable variation in the

study variables across the 38 provinces. The number of malaria cases exhibits a highly right-skewed distribution, as reflected by its mean of 14,314.90 and standard deviation of 45,999.81. Population density also shows substantial variation, with a maximum value of 16,165 persons/km^2. These characteristics highlight the presence of extreme observations that may influence the standardization process, knot selection, and model estimation. Therefore, the potential influence of these observations needs to be considered in the modeling procedure.

3.2. Visualization of Data Patterns and Estimator Selection

To investigate the underlying relationship between the response and predictor variables, scatter plots were generated. The relationships between each predictor variable and the response variable are illustrated in Figure Figure 1.

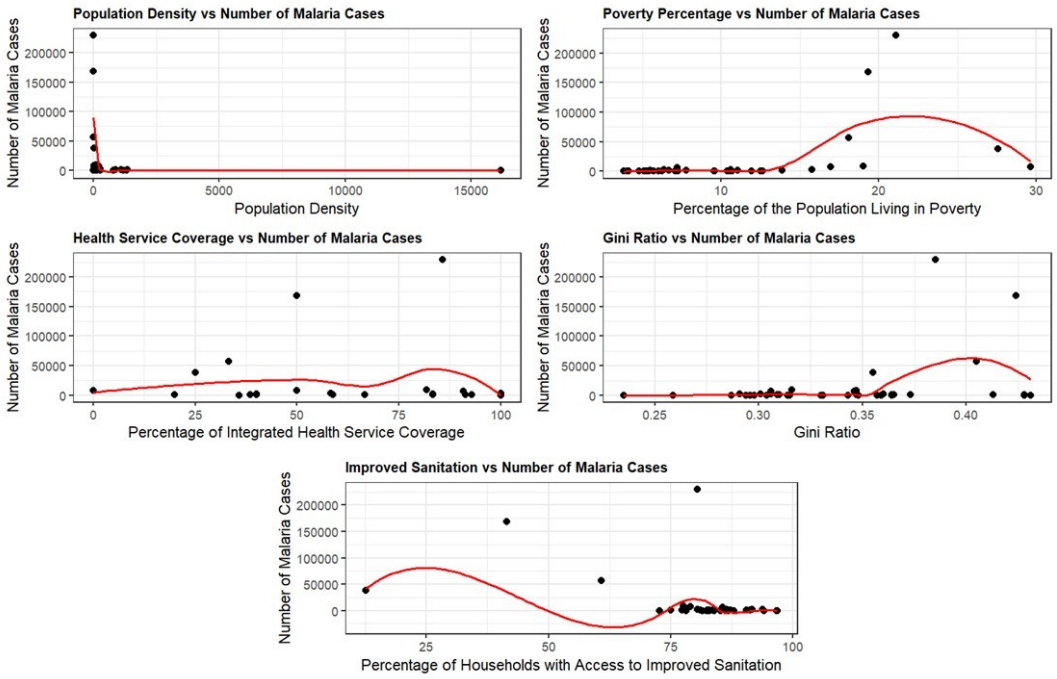


Figure 1. Research modeling framework.

Figure Figure 1 presents the relationships between the number of malaria cases and the five predictor variables. The number of malaria cases and population density are displayed on a logarithmic scale to improve the s, the percentage of the population living in poverty, and integrated health service coverage to the truncated spline component. Meanwhile, the Gini ratio and the percentage of households with access to improved sanitation are assigned to the Uniform kernel component. This specification is retained as the model structure examined in the present study.

3.3. Determination of the Optimal Knot Points and Bandwidth

3.3.1. Optimal Knot Points and Bandwidth for the One-Knot Model

The first step in constructing the Hybrid Spline–Uniform Kernel model was to determine the knot points and bandwidths that yielded the best estimation performance. At this stage, the predictor variables modeled using the spline approach, namely population density, the percentage of the population living in poverty, and the percentage of integrated health service coverage, were each fitted using a single knot point. Meanwhile, the two predictor variables modeled using the Uniform kernel approach, namely the Gini ratio and the percentage of

households with access to improved sanitation, employed bandwidths determined through an optimization procedure. The optimal combination of knot points and bandwidths was selected using the Generalized Cross Validation (GCV) criterion, with the best model corresponding to the minimum GCV value. The results of the optimization for the one-knot model are presented in Table Table 3.

Table 3. GCV values for the one-knot model.

Knot			Bandwidth		GCV
$\boldsymbol{X_1}$	X_2	X_3	X_4	X_5	
-0.2215937	0.6834351	-1.1162630	0.2000000	1.2000000	0.5184287

Table Table 3 shows that the minimum GCV value of 0.5184287 was obtained using a single knot for each spline predictor, with standardized knot locations of -0.2215937, 0.6834351, and -1.1162630 for population density, the percentage of the population living in poverty, and integrated health service coverage, respectively. The corresponding optimal bandwidths for the Uniform kernel predictors were 0.2000000 and 1.2000000. These values were obtained using standardized variables. For interpretation, the knot locations are also converted back to their original measurement scales. Based on the GCV criterion, this combination was selected as the single-knot Hybrid Spline–Uniform Kernel model among the candidate models evaluated.

3.3.2. Determination of the Optimal Knot Points and Bandwidth for the Two-Knot Model

The next step was to determine the optimal knot points and bandwidths for the Hybrid Spline–Uniform Kernel model with two knot points. The optimal combination of knot points and bandwidths was selected using the Generalized Cross Validation (GCV) criterion, with the best model identified as the one yielding the minimum GCV value. The results of the optimization for the two-knot model are presented in Table Table 4.

Table 4. GCV values for the two-knot model.

Knot			Bandwidth		GCV
$\boldsymbol{X_1}$	X_2	X_3	X_4	X_5	
-0.2557580	-0.8385335	0.1372266	0.5684211	0.4000000	0.2641125
-0.2541293	1.7908280	0.4091065			

Table Table 4 shows that the minimum GCV value of 0.2641125 was obtained using two knot points for each spline predictor. The knot locations were -0.2557580 and -0.2541293 for population density, -0.8385335 and 1.7908280 for the percentage of the population living in poverty, and 0.1372266 and 0.4091065 for integrated health service coverage. The corresponding optimal bandwidths were 0.5684211 and 0.4000000 for the two Uniform kernel predictors. Although the two-knot model produced a lower GCV value than the one-knot model, the two knot locations for population density were very close to each other in the standardized scale. This proximity should be considered when interpreting the spline component, as the additional knot may provide limited additional flexibility. Therefore, the two-knot model was retained as a candidate based on its GCV value, while the adequacy of its knot configuration was further considered in the subsequent model evaluation.

3.3.3. Determination of the Optimal Knot Points and Bandwidth for the Three-Knot Model

The final step was to determine the optimal knot points and bandwidths for the Hybrid Spline–Uniform Kernel model with three knot points. As in the previous stages, the optimal

combination of knot points and bandwidths was determined using the Generalized Cross Validation (GCV) criterion, and the best model was selected based on the minimum GCV value. The optimization results for the three-knot model are presented in Table Table 5.

Table 5. GCV values for the three-knot model.

Knot			Bandwidth		GCV
$\boldsymbol{X_1}$	X_2	X_3	X_4	X_5	
-0.2562275	-1.0276410	-1.9791050			
-0.2557580	-0.9469090	0.2135289	2.3105260	1.3368420	0.0801371
-0.2551065	1.0462530	0.2724052			

Table Table 5 shows that the three-knot model produced a minimum GCV value of 0.0801371, with knot locations of -0.2562275, -0.2557580, and -0.2551065 for X1; -1.0276410, -0.9469090, and 1.0462530 for X2; and -1.9791050, 0.2135289, and 0.2724052 for X3. The corresponding optimal bandwidths were 2.3105260 and 1.3368420. Although the three-knot model yielded the lowest GCV value among the models considered, this result alone does not establish that it provides the best predictive performance. The closely spaced knot locations for X1 and the variation in the selected bandwidths across models indicate that the model selection results should be interpreted cautiously. Therefore, the three-knot model was retained as the selected model based on the GCV results, while acknowledging the need for further validation using an independent model assessment procedure.

3.3.4. Selection of the Best Model Based on the Generalized Cross Validation (GCV) Criterion

To identify the best-performing Hybrid Spline–Uniform Kernel model, the Generalized Cross Validation (GCV) values obtained from the one-knot, two-knot, and three-knot models were compared. The best model was selected based on the smallest GCV value, as it indicates better generalization performance in modeling the relationship between the response variable and the predictor variables. The comparison of GCV values for the three models is presented in Table Table 6.

Table 6. Comparison of GCV values for the hybrid spline–uniform kernel models.

Model	GCV
One-knot model	0.5184287
Two-knot model	0.2641125
Three-knot model	0.0801371

Based on Table Table 6, the three-knot model produced the smallest GCV value (0.0801371). This result indicates that the three-knot model provides a better fit for capturing the relationship between the number of malaria cases and the predictor variables than the one-knot and two-knot models. Consequently, the Hybrid Spline–Uniform Kernel model with three knot points was selected as the final model and subsequently used for parameter estimation.

3.4. Parameter Estimation of the Best Hybrid Spline–Uniform Kernel Model

After identifying the best model based on the Generalized Cross Validation (GCV) criterion, the next step was to estimate the parameters of the Hybrid Spline–Uniform Kernel model. Parameter estimation was performed using the three-knot model, as it produced the smallest GCV value among the one-knot, two-knot, and three-knot models. The estimated model parameters are presented in Table Table 7.

Table 7. Parameter estimates of the best hybrid spline–uniform kernel model.

Variable	Parameter	Estimate
Intercept	$\boldsymbol{\beta_0}$	460.1757533
$\mathbf{X_1}$	β_1	1804.3447085
	β_2	-9178.9917927
	β_3	5886.1324974
	β_4	1488.3285356
$\mathbf{X_2}$	β_5	-7.0471119
	β_6	75.3151455
	β_7	-83.5586885
	β_8	15.4321822
$\mathbf{X_3}$	β_9	0.5521721
	β_{10}	-0.5768269
	β_{11}	14.3393238
	β_{12}	-15.9908300
Bandwidth	α_1 (X_4)	2.3105260
	α_2 (X_5)	1.3368420

The parameter estimates of the selected Hybrid Spline–Uniform Kernel model are presented in Table Table 7. The model consists of three spline components for population density (X_1), poverty rate (X_2), and integrated health service coverage (X_3), as well as two Uniform Kernel components for the Gini ratio (X_4) and access to improved sanitation (X_5). These parameter estimates form the basis for constructing the fitted regression function and examining the individual contribution patterns of each predictor. To facilitate interpretation, the estimated spline and kernel components are further examined through component plots using the original measurement scales of the predictor variables.

3.5. Model Performance Evaluation

The performance of the Hybrid Spline–Uniform Kernel model was evaluated using four performance measures: Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). The evaluation results are presented in Table Table 8.

Table 8. Performance evaluation of the best hybrid spline–uniform kernel model.

Criterion	Value
MSE	0.05609136
RMSE	0.2368362
MAPE (%)	51.04229
$\boldsymbol{R^2}$	0.9423927

Based on Table Table 8, the selected Hybrid Spline–Uniform Kernel model produced an MSE of 0.05609136, an RMSE of 0.2368362, a MAPE of 51.04229%, and a coefficient of determination (R^2) of 0.9423927. The coefficient of determination indicates that approximately 94.24% of the variation in the standardized malaria case data is accounted for by the fitted model, while the remaining 5.76% represents unexplained variation. Since the MSE and RMSE were calculated using standardized variables, these values should be interpreted in relation to the standard deviation of the response rather than as errors in the original measurement units. Given that the standard deviation of malaria cases is approximately 46,000 cases, the RMSE corresponds to approximately 10,894 cases. Furthermore, the MAPE value of 51.04% indicates that the model still exhibits substantial relative prediction errors. Therefore, the

high coefficient of determination should not be interpreted as evidence of uniformly accurate predictions.

3.6. Discussion

The results demonstrate that the Hybrid Spline–Uniform Kernel nonparametric regression model with three knot points outperformed the one-knot and two-knot models. This finding is supported by the Generalized Cross Validation (GCV) value of 0.0801371, which is the smallest among all candidate models. A lower GCV value indicates better generalization performance in describing the relationship between the response variable and the predictor variables. The progressive decrease in GCV from the one-knot model (0.5184287) to the two-knot model (0.2641125) and finally to the three-knot model (0.0801371) suggests that increasing the number of knot points improves the model's flexibility in capturing changes in the relationship between malaria cases and the predictor variables.

These findings are consistent with previous studies showing that selecting optimal knot points using the Generalized Cross Validation (GCV) criterion enables truncated spline models to provide a better fit for nonlinear data [8]. Similarly, other studies have reported that incorporating a truncated spline component into hybrid models enhances their ability to capture local structural changes, resulting in better performance than models relying on a single estimation approach [10]. These results further emphasize that selecting an appropriate number of knot points is a crucial step in developing an accurate nonparametric regression model.

In addition to the spline component, this study employed the Uniform kernel to model predictor variables exhibiting relatively smooth relationships, namely the Gini ratio and the percentage of households with access to improved sanitation. The selected bandwidths were 2.3105260 for the Gini ratio and 1.3368420 for improved sanitation. These values determine the width of the Uniform kernel windows, with relatively wide bandwidths allowing more observations to contribute to each local estimate. Given the standardized ranges of these predictors, the selected bandwidths may produce relatively smooth or nearly flat kernel component curves. Therefore, the contributions of these two predictors need to be examined through their estimated component plots. Bandwidth selection plays a fundamental role in kernel regression because it determines the trade-off between smoothness and the model's ability to capture the underlying data pattern. This finding is consistent with previous studies reporting that optimal bandwidth selection has a significant effect on the accuracy of kernel regression estimates [1]. It is also supported by studies demonstrating that kernel estimators provide consistent estimates when an appropriate kernel function and bandwidth are selected according to the characteristics of the data [12].

The best-performing model achieved an MSE of 0.05609136, an RMSE of 0.2368362, a MAPE of 51.04229%, and an value of 0.9423927. The R^2 value indicates that approximately 94.24% of the variation in malaria case counts across the 38 Indonesian provinces was accounted for by the fitted model, while the remaining 5.76% represents unexplained variation. However, this value reflects the in-sample goodness of fit and should not be interpreted as evidence of a causal relationship between the predictor variables and malaria incidence. Furthermore, the MAPE value of 51.04% indicates that the model still exhibits substantial relative prediction errors. The variation in malaria case counts across provinces, particularly the dominance of a few provinces with exceptionally high case counts, may contribute to this prediction error. Other factors, including climatic conditions, rainfall, temperature, humidity, and environmental characteristics, may also be relevant to malaria transmission but were not included in the model.

Overall, the findings indicate that the Hybrid Spline–Uniform Kernel approach provides a flexible framework for modeling malaria case counts across Indonesian provinces by incorporating truncated spline and Uniform kernel components. The truncated spline component represents potentially nonlinear relationships involving population density, poverty rate, and integrated health service coverage, while the Uniform kernel component accommodates the relationships involving the Gini ratio and access to improved sanitation. However, the present study evaluates only the proposed hybrid specification and does not include comparisons with alternative parametric, spline-based, kernel-based, or other nonparametric models. Therefore, the findings do not establish the superiority of the Hybrid Spline–Uniform Kernel approach over models based on a single estimator or other competing methods. Further comparative studies using the same dataset and leave-one-out cross-validation are needed to assess its relative predictive performance [17]. Therefore, the Hybrid Spline–Uniform Kernel model represents a promising alternative for modeling public health data characterized by complex and nonlinear relationships.

3.7. Study Limitations

This study has several limitations that should be considered when interpreting the results. First, the analysis was based on only 38 observations representing Indonesian provinces, resulting in a relatively small sample size for a model containing multiple spline coefficients and kernel components. Second, the distribution of malaria cases was highly heterogeneous across provinces, with a few provinces contributing disproportionately large case counts. This imbalance may influence parameter estimation and model performance.

Third, the response variable consisted of absolute malaria case counts rather than incidence rates adjusted for population size. Consequently, differences in provincial population sizes may affect the observed variation in malaria cases. Fourth, the cross-sectional ecological design limits the interpretation of the relationships between the predictors and malaria cases, and the findings cannot be used to establish individual-level or causal relationships. Fifth, several potentially influential climatic and environmental factors, including rainfall, temperature, humidity, and other environmental characteristics, were not included in the analysis.

Finally, model performance was evaluated using the available dataset without independent out-of-sample validation. Therefore, the reported performance metrics primarily reflect the model's fit to the observed data and may not fully represent its predictive performance for new observations. Future studies should consider larger datasets, population-adjusted malaria measures, additional climatic and environmental predictors, and cross-validation or independent validation procedures to assess the model's generalizability.

4. Conclusion

This study developed a Hybrid Spline–Uniform Kernel nonparametric regression model to examine the variation in malaria case counts across 38 Indonesian provinces. The model combined truncated spline components for population density, poverty rate, and integrated health service coverage with Uniform kernel components for the Gini ratio and access to improved sanitation. Based on the GCV criterion, the three-knot specification yielded the lowest GCV value of 0.0801371 among the configurations evaluated in this study. The corresponding knot locations, expressed in their original measurement units, were approximately 9.6, 10.8, and 12.5 persons/km^2 for population density; 4.1%, 4.6%, and 17.3% for the poverty rate; and 18.5%, 83.3%, and 85.0% for integrated health service coverage. The selected bandwidths were 2.3105260 for the Gini ratio and 1.3368420 for access to improved sanitation. The

model achieved an R^2 of 0.9423927 and a MAPE of 51.04229%, indicating that although the model accounted for a substantial proportion of the variation in the observed data, considerable relative prediction errors remained. These findings highlight the relevance of examining population density, poverty, health service coverage, income inequality, and sanitation when studying differences in malaria case counts across provinces. However, the results represent associations at the provincial level and should not be interpreted as causal effects. The findings may provide an analytical basis for further research on regional malaria patterns, while model validation and comparisons with alternative approaches are still needed to establish its predictive performance and practical applicability.

Declaration of Artificial Intelligence (AI) Use. The authors declare that AI-assisted tools were used solely to support language refinement, grammar checking, and improvements in manuscript readability. All research activities, data analysis, interpretation of results, and scientific conclusions were conducted and verified by the authors, who take full responsibility for the content of this manuscript.

Author Contributions. **Nur Salsabila Wero**: Conceptualization, methodology, data curation, formal analysis, visualization, and writing original draft. **Rahmat Hidayat**: Conceptualization, methodology, supervision, validation, and writing review and editing. Rika Saniarti: Data curation and investigation. **BQ. Nurkhalishah K**: Data curation and investigation. **Ruliana Ruliana**: Methodology and validation. **Aswi Aswi**: Supervision and validation.

Acknowledgment. The authors express their sincere gratitude to everyone who provided valuable contributions and assistance throughout the research process and manuscript preparation. They also extend their appreciation to the editors and reviewers for their constructive feedback and valuable suggestions, which helped enhance the quality of this manuscript.

Funding. This research received no external funding.

Data availability Statement. The data used in this study are secondary data obtained from official publications of Statistics Indonesia (BPS) and the Ministry of Health of the Republic of Indonesia for 2024. The compiled dataset covering 38 Indonesian provinces and the R code used for data preprocessing, standardization, knot selection, bandwidth optimization, model estimation, and performance evaluation are provided as supplementary materials. These materials are intended to facilitate the reproducibility of the research findings.

Conflict of Interest. The authors declare no conflicts of interest regarding the research or the publication of this article.

References

[1] M. I. Efendi, I. N. Budiantara, and V. Ratnasari, "Pemilihan bandwidth optimal pada regresi nonparametrik kernel menggunakan metode unbiased risk (ubr)," *Jurnal Gaussian*, vol. 14, no. 2, pp. 619–630, 2025, doi:10.14710/j.gauss.14.2.619-630.

[2] W. Sanusi, R. Syam, and R. Adawiyah, "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*, vol. 2, no. 1, pp. 70–81, 2019, doi:10.35580/jmathcos.v2i1.12460.

[3] R. Hidayat, I. N. Budiantara, B. W. Otok, and V. Ratnasari, "An extended model of penalized spline with the addition of kernel functions in nonparametric regression model," *Applied Mathematics & Information Sciences*, vol. 13, no. 3, pp. 453–460, 2019, doi:10.18576/amis/130318.

[4] D. C. Wongkar, Ruliana, and M. F. S., "Analisis regresi nonparametrik spline truncated untuk menganalisis faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di provinsi sulawesi selatan," *VAR-*

IANSI: Journal of Statistics and Its Application on Teaching and Research, vol. 5, no. 2, pp. 55–63, 2023, doi:10.35580/variansiunm101.

[5] N. A. Schuster, J. J. M. Rijnhart, J. W. R. Twisk, and M. W. Heymans, "Modeling non-linear relationships in epidemiological data: The application and interpretation of spline models," *Frontiers in Epidemiology*, vol. 2, p. 975380, 2022, doi:10.3389/fepid.2022.975380.

[6] R. Hidayat, M. Ilyas, and Yuliani, "Multiple nonparametric regression approach for regional mapping by considering the significance of the independent variables," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 3, pp. 552–561, 2024.

[7] E. Yavuz and M. Şahin, "Investigation of parametric, non-parametric and semiparametric methods in regression analysis," *Sakarya University Journal of Science*, vol. 26, no. 6, pp. 1111–1116, 2022, doi:10.16984/saufenbilder.1147135.

[8] S. Anggraeni, Sifriyani, and Q. Q. A'yun, "Tuberculosis case model using gcv and ubr knot selection methods in truncated spline nonparametric regression," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 3, pp. 1565–1574, 2023, doi:10.30598/barekengvol17iss3pp1565-1574.

[9] R. Hidayat, M. Ilyas, and Yuliani, "Spline model in the case of cervical cancer patient resilience," *Trends in Sciences*, vol. 20, no. 8, p. 6170, 2023, doi:10.48048/tis.2023.6170.

[10] 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, p. 102652, 2024, doi:10.1016/j.mex.2024.102652.

[11] 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," *Communications in Statistics - Theory and Methods*, vol. 50, no. 17, pp. 3942–3953, 2021, doi:10.1080/03610926.2019.1710201.

[12] S. Bouzebda and N. Taachouche, "Rates of the strong uniform consistency for the kernel-type regression function estimators with general kernels on manifolds," *Mathematical Methods of Statistics*, vol. 32, no. 1, pp. 27–80, 2023, doi:10.3103/S1066530723010027.

[13] F. Lu, Q. An, and Y. Yu, "Nonparametric learning of kernels in nonlocal operators," *Journal of Peridynamics and Nonlocal Modeling*, vol. 6, no. 3, pp. 347–370, 2024, doi:10.1007/s42102-023-00105-9.

[14] A. M. Sukran, I. N. Budiantara, and V. Ratnasari, "The curve estimation nonparametric regression multiresponse mixed with truncated spline, fourier series, and kernel," *Mandalika Mathematics and Educations Journal*, vol. 7, no. 2, pp. 766–787, 2025, doi:10.29303/jm.v7i2.9188.

[15] A. M. Kahar, S. Suparti, and A. R. Hakim, "Analisis indeks harga saham gabungan dan faktor pengaruhnya menggunakan pemodelan regresi semiparametrik kernel dilengkapi gui-r," *Jurnal Gaussian*, vol. 12, no. 1, pp. 30–41, 2023, doi:10.14710/j.gauss.12.1.30-41.

[16] P. R. Sihombing and A. Famalika, "Penerapan analisis regresi nonparametrik dengan pendekatan regresi kernel dan spline," *Jurnal Ekonomi dan Statistik Indonesia*, vol. 2, no. 2, pp. 172–181, 2022, doi:10.11594/jesi.02.02.05.

[17] R. Hidayat, I. Muthahharah, and Resmawan, "Kernel-truncated spline: Estimator fleksibel untuk regresi nonparametrik," *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 13, no. 2, 2025, doi:10.37905/euler.v13i2.33062.

[18] B. W. Otok, M. S. Akbar, and Raupong, "Estimasi spline dan mars menggunakan kuadrat terkecil," *Jurnal Matematika, Statistika dan Komputasi*, vol. 4, no. 1, pp. 1–11, 2007, doi:10.20956/jmsk.v4i1.3319.

[19] M. A. Fitri, M. F. F. Mardianto, and E. Ana, "Modeling environmental quality index in indonesia with multivariate adaptive regression splines," *International Journal of Innovative Science and Research Review*, vol. 7, no. 5, pp. 8350–8353, May 2025.

[20] A. Kurniawan, A. Z. Widyawati, F. A. Pratiwi, N. Faizun, and R. Meliana, "Pemodelan prevalensi angka kesakitan malaria berdasarkan persentase sanitasi layak dengan pendekatan estimator least square spline," *Jurnal Kesehatan Lingkungan Indonesia*, vol. 23, no. 3, pp. 287–293, 2024, doi:10.14710/jkli.23.3.287-293.

[21] A. Avichena and R. Anggriyani, "Analisis penyakit malaria akibat infeksi plasmodium sp terhadap darah manusia," *EKOTONIA: Jurnal Penelitian Biologi, Botani, Zoologi dan Mikrobiologi*, vol. 8, no. 1, pp. 30–37, 2023, doi:10.33019/ekotonia.v8i1.4128.

[22] Kementerian Kesehatan Republik Indonesia, *Profil Kesehatan Indonesia 2024*. Jakarta: Kementerian Kesehatan Republik Indonesia, 2024.

[23] U. Kalsum, D. R. Pertiwi, A. L. Veronica, and A. Wulandari, "Determinan yang berhubungan dengan kejadian malaria di indonesia tahun 2016," *Jurnal Kesmas Jambi*, vol. 2, no. 1, pp. 81–91, 2018, doi:10.22437/jkmj.v2i1.6545.

[24] H. Hasyim, P. Dale, D. A. Groneberg, U. Kuch, and R. Müller, "Social determinants of malaria in an endemic area of indonesia," *Malaria Journal*, vol. 18, no. 1, p. 134, apr 2019, doi:10.1186/s12936-019-

2760-8.
[25] M. Y. Matdoan, M. W. Talakua, and R. J. Djami, "Pemodelan regresi quantil dengan kernel smoothing pada faktor-faktor yang mempengaruhi penyebaran API malaria di indonesia," *Uniqbu Journal of Exact Sciences*, vol. 1, no. 2, pp. 1–9, aug 2020.
[26] Badan Pusat Statistik Provinsi Jawa Barat, *Provinsi Jawa Barat dalam Angka 2024*. Bandung: Badan Pusat Statistik Provinsi Jawa Barat, 2024.
[27] I. Muthahharah and R. Hidayat, "The impact of malnutrition on infant mortality rate in indonesia: A spline regression approach," *Journal of Mathematics, Computations, and Statistics*, vol. 8, no. 2, pp. 474–481, 2025, doi:10.35580/jmathcos.v8i2.75361.
[28] B. Funke and M. Hirukawa, "On uniform consistency of nonparametric estimators smoothed by the gamma kernel," *Annals of the Institute of Statistical Mathematics*, vol. 77, no. 3, 2025, doi:10.1007/s10463-024-00923-8.
[29] R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd ed. New York: Marcel Dekker, 1999.
[30] D. R. Arifanti and R. Hidayat, "Statistical inference on nonparametric spline models and its applications," *Journal of Physics: Conference Series*, vol. 2123, no. 1, p. 012023, 2021, doi:10.1088/1742-6596/2123/1/012023.
[31] M. Hirukawa, I. Murtazashvili, and A. Prokhorov, "Uniform convergence rates for nonparametric estimators smoothed by the beta kernel," *Scandinavian Journal of Statistics*, vol. 49, no. 3, pp. 1353–1382, 2022, doi:10.1111/sjos.12573.
[32] W. Härdle, *Applied Nonparametric Regression*, ser. Econometric Society Monographs. Cambridge: Cambridge University Press, 1990, no. 19, doi:10.1017/CCOL0521352487.
[33] P. Speckman, "Kernel smoothing in partial linear models," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 50, no. 3, pp. 413–436, 1988, doi:10.1111/j.2517-6161.1988.tb01738.x.
[34] D. Ariesta, N. Gusriani, and K. Parmikanti, "Estimasi parameter model regresi nonparametrik B-Spline pada angka kematian maternal," *Jurnal Matematika UNAND*, vol. 10, no. 3, pp. 342–354, 2021, doi:10.25077/jmu.10.3.342-354.2021.
[35] N. Chamidah, B. Lestari, T. Saifuddin, R. Rulaningtyas, P. Wardhani, and I. N. Budiantara, "Estimating the number of malaria parasites on blood smears microscopic images using penalized spline nonparametric poisson regression," *Communications in Mathematical Biology and Neuroscience*, vol. 2024, pp. 1–19, 2024, doi:10.28919/cmbn/8737.
[36] P. Craven and G. Wahba, "Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation," *Numerische Mathematik*, vol. 31, no. 4, pp. 377–403, 1979, doi:10.1007/BF01404567.