BAREKENG: Journal of Mathematics and Its Applications
March 2026 Volume 20 Issue 1 Page 0087–0104
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss1pp0087-0104

ANALYSIS OF FACTORS AFFECTING PNEUMONIA IN INDONESIAN TODDLERS USING NONPARAMETRIC REGRESSION WITH LEAST SQUARE SPLINE AND FOURIER SERIES METHODS

Toha Saifudin [1*], Suliyanto [2], Nabila Nurdin [3], Bryan Given Christiano Ginzel [4], Sabrina Salsa Oktavia [5], Jovansha Ariyawan [6], Mohammad Noufal Ubadah [7]

[1,2,3,4,5,6,7] *Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga*
Jln. Dr. Ir. H. Soekarno, Mulyorejo, Surabaya, Jawa Timur, 60115, Indonesia

*Corresponding author's e-mail: * **tohasaifudin@fst.unair.ac.id***

Article Info

Article History:
Received: 10th February 2025
Revised: 30th May 2025
Accepted: 25th July 2025
Available online: 24th November 2025

Keywords:
Fourier series method;
Health policy;
LS-Spline;
Nonparametric regression;
Pneumonia.

ABSTRACT

Pneumonia is the leading cause of death among children under five, with the highest prevalence in Indonesia found in West Papua Province (75%) and the lowest in North Sulawesi (0.3%). This study aims to analyze the factors influencing the prevalence of pneumonia in Indonesian toddlers using nonparametric regression approach by comparing Least Square Spline (LS-Spline) and Fourier Series. Data sourced from the Indonesian Ministry of Health website, consisting of 34 provinces in Indonesia in 2023, with one response variable (Y) and five predictor variables (X). The analyzed factors include the coverage of vitamin A supplementation, malnutrition rates, low birth weight prevalence, measles immunization coverage, and exclusive breastfeeding rates. The analysis was conducted by modeling with nonparametric Least Square Spline regression using up to three optimal knot points, then performing analysis using nonparametric regression with the Fourier series approach. The two methods were compared based on GCV and R^2, with the best model having lower GCV and higher R^2. The results showed that LS-Spline was better than Fourier Series, with a GCV value of 233.16 and a coefficient of determination of 92.5%. The findings reveal that the relationships between predictor factors and pneumonia prevalence are nonlinear, with varying influence patterns across different variable ranges. These results indicate that LS-Spline has a strong ability to explain data variability. The Fourier series is limited in this study because it is best suited for periodic data, unlike pneumonia data and its causal factors which do not show such patterns. The weakness of the Fourier Series in this study lies in its suitability for periodic data, while pneumonia cases and their causal factors do not follow such patterns. This study offers insights into health policy making to reduce pneumonia cases, improve their lives, in line with the SDGs target on Good Health and Well-being.

How to cite this article:

T. Saifudin, Suliyanto, N. Nurdin, B. G. C. Ginzel, S. S. Oktavia, J. Ariyawan and M. N. Ubadah., "ANALYSIS OF FACTORS AFFECTING PNEUMONIA IN INDONESIAN TODDLERS USING NONPARAMETRIC REGRESSION WITH LEAST SQUARE SPLINE AND FOURIER SERIES METHODS," *BAREKENG: J. Math. & App.*, vol. 20, iss. 1, pp. 0087-0104, Mar, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Pneumonia is an inflammation of the lung tissue that can be caused by viruses, bacteria, fungi, or parasites. In 2019, pneumonia caused 14% of all deaths in children under 5 years of age with a total of 740,180 deaths [1]. Pneumonia is the leading cause of death among toddlers in the world, even more than AIDS, malaria, and measles combined. The highest prevalence of pneumonia cases is in Southeast Asia at 39%, and Africa at 30% [2]. In the Southeast Asian region, there are 3 countries with the highest prevalence, namely the Philippines at 5.2%, Indonesia at 3.8%, and Malaysia at 1.8%. Cases of pneumonia among toddlers in Indonesia in 2020 were 309,838 people with East Java Province as the province with the highest rate of pneumonia cases, which was 76,929 people [3]. Every year, it is estimated that more than 2 million toddlers die from pneumonia (1 toddler every 20 seconds) out of a total of 9 million toddler deaths. Out of every 5 deaths of toddlers, 1 of them is caused by pneumonia.

There are several factors that affect pneumonia in toddlers. Previous studies that analyzed the effect of measles and vitamin A immunization on pneumonia using the Poisson regression test method concluded that measles immunization accompanied by vitamin A administration can be used as a preventive measure in an effort to reduce the incidence of pneumonia [4]. There were also other studies that examined the relationship between exclusive breastfeeding history, low birth weight (BBLR), and home physical condition with the incidence of pneumonia in toddlers, which was obtained that there was a significant relationship between exclusive breastfeeding history, BBLR history, and home physical condition with pneumonia in toddlers in the working area of the Kediri Health Center, West Lombok Regency [5]. As for another study that analyzed the relationship between nutritional status and the degree of pneumonia in toddlers, it was concluded that there was a meaningful relationship between nutritional status and the degree of pneumonia in toddlers in hospitals. Dr. M. Djamil Padang, namely most of the toddlers with severe pneumonia have poor and poor nutritional status [6]. In addition, there are studies that show that the use of the spline nonparametric regression method is appropriate to analyze factors that affect pneumonia in toddlers because R^2 is obtained at 97.224% [7]. Furthermore, a large-scale international study in China found that low birth weight, short breastfeeding duration, parental smoking, and damp housing conditions significantly increased the risk of pneumonia in children, highlighting the importance of both biological and environmental factors in pneumonia prevention efforts [8]. Supporting these findings, a case-control study in Indonesia also showed that nonexclusive breastfeeding, incomplete basic immunizations, indoor air pollution, low birth weight, and severe wasting are significant risk factors for pneumonia in children. Among these, exclusive breastfeeding was identified as the most dominant protective factor, while sex was found to modify its relationship with pneumonia incidence [9].

In study [10], research has been conducted on the analysis of factors affecting pneumonia in toddlers related to weather using the VAR and VECM methods. For further research, the authors analyzed factors influencing pneumonia in toddlers using different methods and variables to explore other aspects that have not been studied. This follow-up research with a different approach is expected to help uncover new relationships between variables. This study uses one of the nonparametric regression methods, namely the least square spline method because it is one of the methods that can explore the variables that affect pneumonia in Indonesian toddlers. LS-Spline is a method that adapts local changes in data using knots, which allows the model to follow changes in data behavior at different intervals [11]. The optimal determination of the position of the knots is carried out using the Generalized Cross Validation (GCV) method, which selects the number and position of knots by minimizing the function of the GCV. This approach ensures that the spline model balances between the accuracy of the data fitting and the smoothness to avoid overfitting [12]. The advantage of this method lies in its flexibility and ability to handle complex data with diverse patterns. Therefore, this method is very suitable for the analysis of variables that affect pneumonia in toddlers in Indonesia, with patterns and relationships between variables likely to differ in each region. The selection of the spline method in this study is based on its ability to capture non-linear relationships between predictor variables. The flexibility of this method allows for a more in-depth analysis of the interaction between factors such as nutritional status, immunization coverage, and exclusive breastfeeding. Traditional linear regression methods may not be able to capture the complexity of these relationships. Therefore, the spline method becomes more appropriate in this analysis, as it provides a better balance between model complexity and interpretability. This is supported by [13], who demonstrated that spline regression models offer higher explanatory power and better flexibility in modeling non-linear relationships compared to traditional linear or polynomial regression, while still maintaining interpretable coefficients.

In addition, further analysis was conducted using nonparametric series regression as a comparison to the results of the LS-Spline analysis. The use of the LS-Spline and Fourier Series methods is highly relevant in analyzing the factors influencing pneumonia among infants in Indonesia, due to their ability to capture complex, nonlinear relationships between causal variables and pneumonia prevalence relationships that cannot be adequately modeled using conventional parametric regression. The LS-Spline method was selected for its flexibility in adapting to local variations in the data by optimized knot points, determined by the Generalized Cross-Validation (GCV) method, which allows for accurate modeling while minimizing the risk of overfitting. In contrast, the Fourier Series method was applied as a comparative approach, particularly for assessing model performance in data that may exhibit oscillatory or periodic behavior. Therefore, the application of both methods serves to address the complexity of regional variations in pneumonia risk factors, while providing a more adaptive and data-driven modeling framework to inform health policy and intervention strategies.

This research is also carried out to support the Sustainable Development Goals (SDGs), precisely to support the third goal, namely "Good Health and Well-being" (SDG 3), especially target 3.2 which aims to end preventable infant and toddler deaths by 2030 [14]. By understanding the factors that affect pneumonia in toddlers, it is hoped that more effective preventive measures can be taken both through health policies and interventions in the community. This is expected to be able to reduce the incidence of pneumonia in toddlers and improve their quality of life.

2. RESEARCH METHODS

2.1 Data Sources and Research Variables

This approach is used to analyze factors affecting pneumonia in toddlers in Indonesia using the spline nonparametric regression method. The data used in this study is secondary data. Secondary data is a data source that does not directly provide data to data collectors, for example through documents [15]. This research comes from various sources to complement the shortcomings in the previous data. The data source used comes from the Ministry of Health (https://kemkes.go.id/) website, precisely from the 2023 Indonesian Health Profile [16]. The data used are from 34 provinces in Indonesia in 2023 for all variables, with one response variable (Y) and five predictor variables (X). The unit of measurement for each variable is a percentage. The data used included data on the percentage of pneumonia coverage in toddlers (Y), the percentage of toddlers receiving Vitamin A supplements (X_1), the percentage of toddlers with poor nutritional status (X_2), the percentage of infants with low-birth-weight (X_3), the percentage of toddlers who had received measles immunization (X_4), and the percentage of toddlers receiving exclusive breastfeeding (X_5). The variables used in this study were selected based on empirical findings from previous studies, specifically [4], [5], [6], and [7], which consistently demonstrate that these variables have a significant influence on the incidence of pneumonia in infants. Study [4] highlights the importance of measles immunization and vitamin A supplementation as effective preventive measures in reducing the incidence of pneumonia. Meanwhile, a study [5] found that a history of exclusive breastfeeding, low birth weight, and physical conditions at home have a meaningful association with the risk of pneumonia. Additionally, a study [6] confirms that poor nutritional status is strongly correlated with the severity of pneumonia in infants. Therefore, the selection of variables such as vitamin A coverage, nutritional status, prevalence of low birth weight, measles vaccination coverage, and exclusive breastfeeding in this study has a strong scientific basis and is relevant for further analysis using a nonparametric regression approach.

2.2 Nonparametric Spline Regression

Nonparametric regression using splines is a method in regression analysis that does not assume a specific functional form between the predictor and response variables. This method utilizes polynomial segments joined at knot points to form a flexible curve capable of capturing complex data patterns. The general model for spline regression can be expressed in Eq. (1) [17]:

$$y_i = f(x_i) + \varepsilon_i\,, \quad i = 1,2,\cdots,n, \tag{1}$$

where:

y_i : the response variable;
x_i : the predictor variable;

$f(x_i)$: the spline function;
ε_i : the random error term assumed to be normally distributed with a mean of 0 and variance σ^2.

A spline function of order m with knot points $k_1, k_2, \cdots, k_K$ can be represented in Eq. (2):

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_m x^m + \sum_{j=1}^{K} \beta_{m+j}\left(x - k_j\right)_+^m \tag{2}$$

where:
$\beta_0, \beta_1, \cdots, \beta_{m+K}$: parameters to be estimated
$(x - k_j)_+^m = \max\,(0, x - k_j)^m$: the truncated function.

Parameter estimation in spline regression is typically performed using methods such as Least Squares (LS) or Maximum Likelihood Estimation (MLE). The selection of optimal knot points is often carried out using Cross-Validation (CV) or information criteria like Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) [18]. With its ability to capture complex data patterns, nonparametric spline regression serves as a powerful tool in data analysis, particularly when the relationship between variables is unknown or when the data exhibits varying behaviour across specific intervals.

The best spline regression model is one with optimal knot points, which represent changes in the pattern of the function. One commonly used method for determining optimal knot points is Generalized Cross Validation (GCV). This method is asymptotically optimal, does not depend on the unknown population variance (σ^2), and is invariant to transformations. The optimal knot points can be determined by minimizing the GCV value. In general, the GCV method is expressed in Eq. (3) [19]:

$$GCV(\boldsymbol{K}) = \frac{MSE(\boldsymbol{K})}{[n^{-1} trace(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{K}))]^2} \tag{3}$$

where:
$\boldsymbol{K} = K_1, K_2, \ldots, K_r$: the knot points;
$\boldsymbol{I}$: the identity matrix;
n : the number of observations;

$$MSE(\boldsymbol{K}) = n^{-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 ;$$
$$\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T.$$

2.3 Fourier Series Nonparametric Regression

The Fourier series is a method for analyzing periodic functions. It decomposes periodic functions into their sine and cosine components. The Fourier series is a trigonometric polynomial with high flexibility, enabling it to adapt effectively to the local characteristics of the data. Given the multivariable nonparametric regression model in Eq. (4):

$$y_i = \mu\left(x_{1i}, x_{2i}, \cdots, x_{pi}\right) + \varepsilon_i = \sum_{j=1}^{p} f_j\left(x_{ji}\right) + \varepsilon_i, \quad i = 1,2, \cdots, n, \tag{4}$$

where:
y_i : the response variable for observation i;
x_{ji} : the predictor variable for observation i;
$f_j(x_{ji})$: the partial regression function estimates the nonlinear contribution of predictor j;
p : number of predictors;
ε_i : the random error term assumed to be normally distributed with a mean of 0 and variance σ^2.

The regression curve is then approximated using the Fourier series function as follows:

$$f_j\left(x_{ji}\right) = b_j x_{ji} + \frac{1}{2}\alpha_{0j} + \sum_{k=1}^{K} \alpha_{kj} \cos k x_{ji}, \quad j = 1,2, \cdots, p, \tag{5}$$

where:
b_j : the coefficient for the linear component of predictor x_j;
α_{kj} : the order cosine coefficient for predictor x_j;

K : oscillation parameter.

The estimated model using the Fourier series approach is as follows:

$$y_i = \beta_0 + b_1 x_{1i} + \sum_{k=1}^{K} \alpha_{k1} \cos kx_{1i} + b_2 x_{2i} + \sum_{k=1}^{K} \alpha_{k2} \cos kx_{2i} + \cdots + b_2 x_{2i} + \sum_{k=1}^{K} \alpha_{kp} \cos kx_{pi}\,. \quad (6)$$

The oscillation parameter (K) has a significant impact on the model in nonparametric regression using the Fourier series. The number of cosine wave oscillations in the model is determined by this parameter (K). The predicted curve's oscillations become denser and as K increases the model becomes more complicated and resembles the data patterns. As a result, the bias decreases while the variance tends to increase [20].

2.4 Steps of Data Analysis

The stages of analysis carried out in this study are as follows:

1. Describe the characteristics of the data.
   a. Collect data from the publication of the Ministry of Health.
   b. Describe the data of each variable of the indicator of the causative factor of pneumonia in toddlers to identify the characteristics of the data.
2. Nonparametric regression modeling using Least Squares Spline in RStudio.
   a. Displaying a visualization of data patterns with scatter plots to find out the pattern of relationships that occur between each response variable and each predictor variable.
   b. Modeling the response variables using a spline nonparametric regression model with various knot points.
   c. Selects the optimal knot point based on the minimum GCV value.
   d. Get the best spline regression model with optimal knot points.
   e. Conducting parameter significance tests simultaneously and partially.
   f. Conducting a homoscedasticity residual assumption test and a normality assumption test from the spline regression model.
   g. Create an interpretation of the Least Square Spline nonparametric regression model.
3. Nonparametric regression modeling using Fourier Series in RStudio.
   a. Modeling each predictor variable using a finite Fourier series expansion up to order $K = 3$.
   b. Testing multiple values of the oscillation parameter $(K = 1,2,3)$ to evaluate the model's fit.
   c. Selecting the optimal Fourier model based on the minimum GCV value and highest coefficient of determination (R^2).
   d. Formulate the regression equation estimate based on the optimal model results.
4. Comparing the performance of the Fourier model with the LS-Spline model to determine the better fit.

3. RESULTS AND DISCUSSION

3.1 Result of Research

Pneumonia in toddlers is one of the health problems that occur in Indonesia. Pneumonia is recognized as one of the most significant respiratory tract infections, with the potential to cause serious complications if not treated appropriately and in a timely manner. Therefore, an analysis related to pneumonia in toddlers was conducted using Minitab and R software [21]. The percentage of pneumonia detection coverage in toddlers in 2023 is as follows:

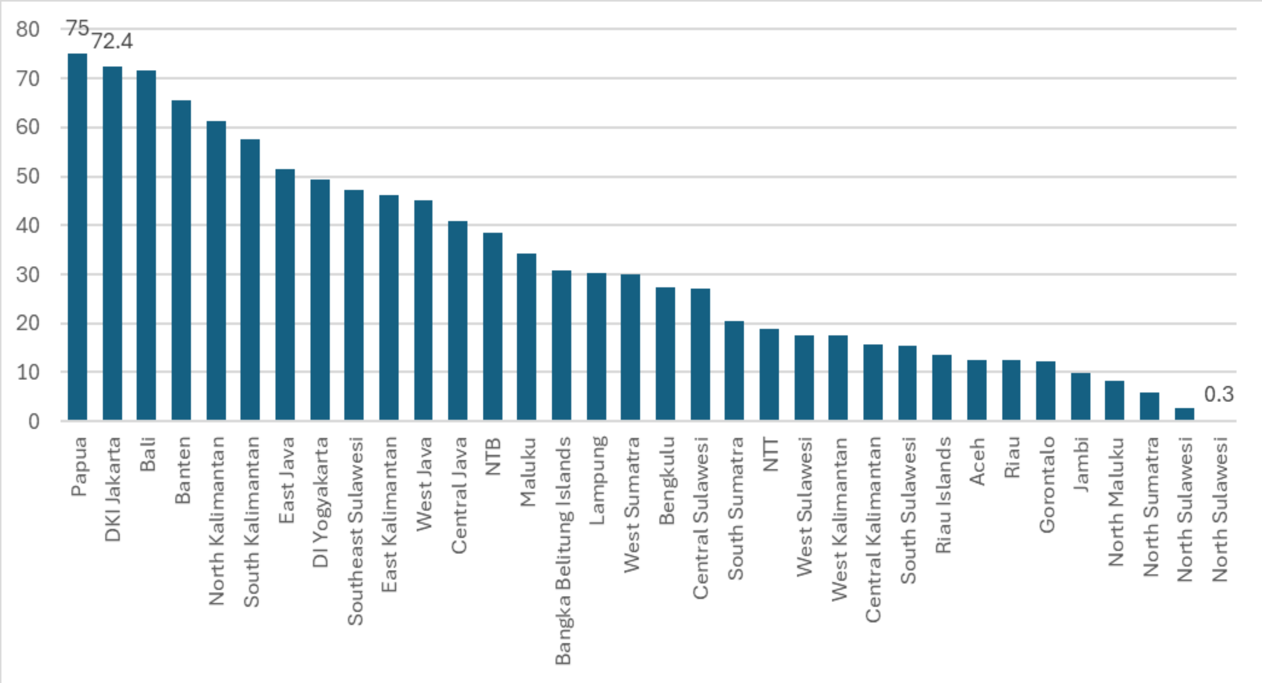

Figure 1. Percentage Coverage of Pneumonia Detection in Toddlers (2023)

Based on Fig. 1, it is known that the highest percentage of pneumonia in toddlers in Indonesia is in West Papua Province with 75 percent, which means that among 100 toddlers there are 75 toddlers who suffer from pneumonia. The province with the next highest percentage is DKI Jakarta Province with 72.4 percent and Bali Province with 71.6 percent. In addition, it is known that the lowest percentage of pneumonia in children under five in Indonesia is in North Sulawesi Province with a percentage of 0.3 percent. The characteristics of pneumonia in toddlers in 2023 and the factors that are suspected to affect it are as follows:

Table 1. Descriptive Statistics of Pneumonia in Toddlers (2023) and the Factors

Variable	Average	Variance	Minimum Value	Maximum Value
Y	31.92	469.39	0.30	75.00
X_1	63.21	246.16	26.31	90.49
X_2	4.088	3.442	1.200	7.900
X_3	4.591	4.144	1.000	8.10
X_4	68.83	39.72	49.72	81.32
X_5	57.39	308.01	10.90	81.10

Based on Table 1, it is known that the average percentage of pneumonia in children under five in Indonesia in 2023 is 31.92 percent with a variance of 469.39. This means that out of 100 toddlers in Indonesia, there are an average of 32 toddlers in each province who suffer from pneumonia. The percentage of pneumonia in toddlers is the highest in Indonesia at 75 percent, namely in West Papua Province and the lowest in Indonesia at 0.3 percent, namely in North Sulawesi Province. This extreme disparity, such as the 75% pneumonia rate in West Papua compared to only 0.3% in North Sulawesi may be influenced by several epidemiological and systemic factors. These include differences in healthcare access, the accuracy of case reporting, the availability of medical personnel, geographic challenges, and the public health infrastructure. In remote provinces like West Papua, limited access to health services may result in delayed diagnosis and higher reported prevalence during concentrated detection campaigns. Conversely, provinces with better infrastructure, such as North Sulawesi, may have more stable and continuous healthcare delivery, leading to lower recorded prevalence. These factors highlight the importance of contextualizing the data and addressing regional inequalities in health surveillance and intervention efforts.

Variable X_1 is the percentage of coverage of Vitamin A administration to toddlers in Indonesia in 2023 with an average of 63.21 percent with a variance of 246.16. This means that out of 100 toddlers in Indonesia, on average, there are 64 toddlers in each province who are given Vitamin A. The highest percentage of coverage of Vitamin A for toddlers in Indonesia in 2023 is 90.49 percent, namely in North Kalimantan

Province. Meanwhile, the lowest percentage of coverage for Vitamin A for toddlers in Indonesia in 2023 is 26.31 percent, namely in Papua Province.

The X_2 variable is the percentage of toddlers with undernourished status in 2023 with an average of 4.088 percent with a variance of 3.442. This means that out of 100 toddlers in Indonesia, there are an average of 5 toddlers with malnourished status in each province. The percentage of toddlers with the highest malnutrition status in Indonesia in 2023 is 7.9 percent, namely in East Nusa Tenggara Province. Meanwhile, the percentage of children under five with the lowest nutritional status in Indonesia in 2023 is 1.2 percent, namely in DKI Jakarta Province.

The X_3 variable is the percentage of babies with low birth weight in 2023 with an average of 4.591 percent with a variance of 4.144. This means that out of 100 babies in Indonesia, there will be an average of 5 babies with low birth weight in 2023 in each province. The highest percentage of babies with low birth weight in Indonesia in 2023 is 8.1 percent, namely in Gorontalo Province. Meanwhile, the percentage of babies with the lowest birth weight in Indonesia in 2023 is 1 percent, namely in DKI Jakarta Province.

The X_4 variable is the percentage of toddlers who have received measles immunization in 2023 with an average of 68.83 percent with a variance of 39.72. This means that out of 100 toddlers in Indonesia, there are an average of 69 toddlers who have received measles immunization in each province. The percentage of children under five who have received the highest measles immunization in Indonesia in 2023 is 81.32 percent, namely in Yogyakarta Province. Meanwhile, the percentage of children under five who have received the lowest measles immunization in Indonesia in 2023 is 49.72 percent, namely in Aceh Province.

The X_5 variable is the percentage of babies who receive exclusive breastfeeding in 2023 with an average of 57.39 percent with a variance of 308.01. This means that out of 100 babies in Indonesia, on average, there are 58 babies who receive exclusive breastfeeding in each province. The highest percentage of babies receiving exclusive breastfeeding in Indonesia in 2023 is 81.1 percent, namely in West Nusa Tenggara Province. Meanwhile, the lowest percentage of babies who receive exclusive breastfeeding in Indonesia in 2023 is 10.9 percent, namely in West Papua Province.

Furthermore, the relationship pattern of response variables and predictor variables is identified by looking at scatterplots or scatter plots. Based on the results of the scatterplot, it can be known whether the pattern of the relationship between the response variable and the predictor forms a certain pattern or not. If it forms a certain pattern such as linear, quadratic, cubic or other patterns, it is included in the parametric component, while if it does not form a certain pattern, it is included in the non-parametric component. If the relationship pattern is known, it can be determined what method is appropriate to do the modeling.

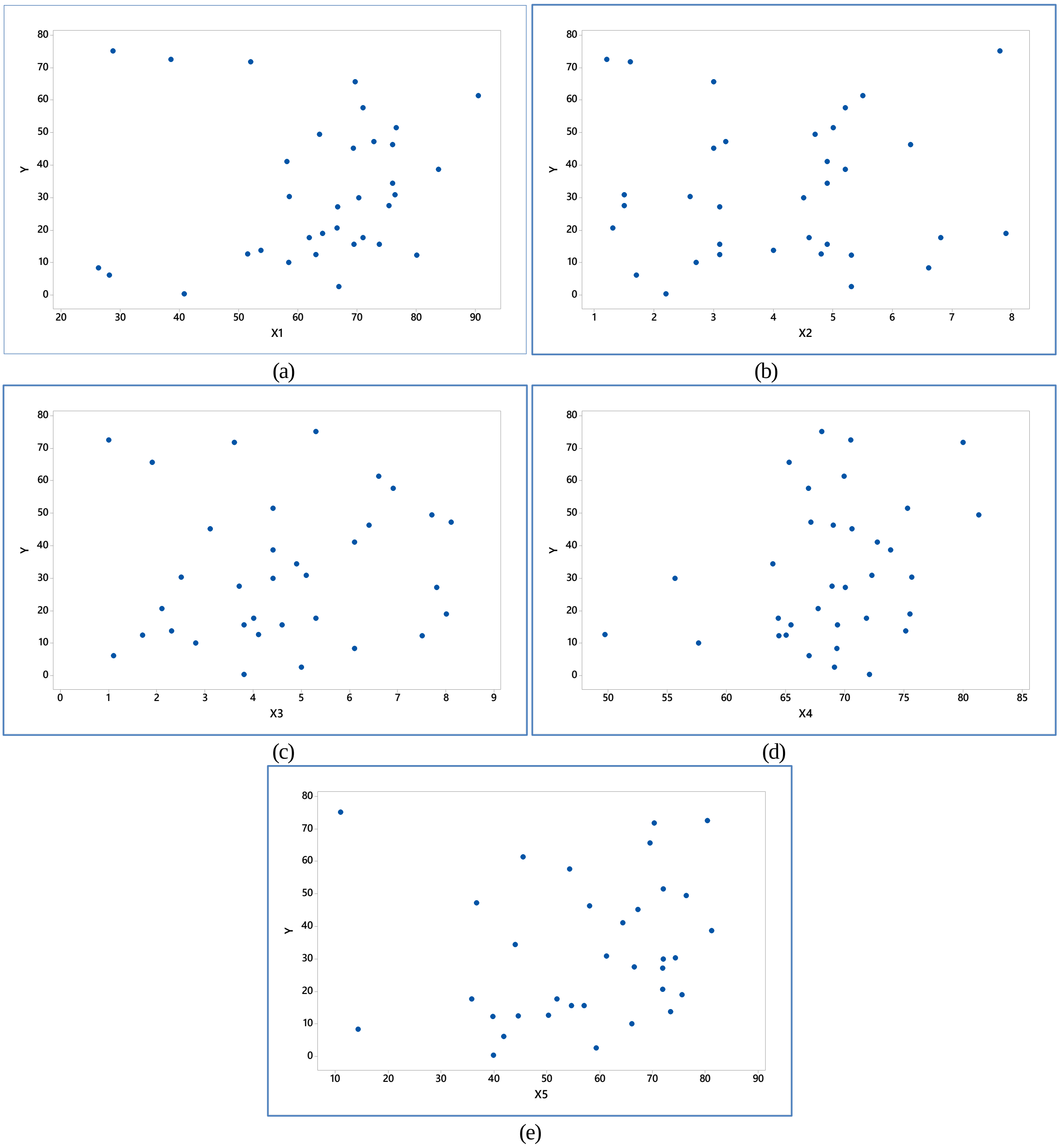

Figure 2. Scatterplot Between Variables (a) Y and X_1, (b) Y and X_2, (c) Y and X_3, (d) Y and X_4, (e) Y and X_5

Based on Fig. 2, it can be seen that all patterns of relationships between variables do not follow or do not form a certain pattern. The scatter plots show varying forms of association, with some indicating irregular distributions and fluctuations. These visual characteristics justify the use of nonparametric regression methods, which are more appropriate for capturing nonlinear and flexible relationships. Therefore, variables X_1, X_2, X_3, X_4, and X_5 can be used as nonparametric components and the analysis can proceed using a nonparametric spline regression approach. The selection of the optimal knot point is carried out to get the best model estimate. The knot point is a point of convergence where there is a pattern of behavior change in the data. The selection of the best model of nonparametric spline regression was obtained from choosing the optimum knot point, which has the lowest Generalized Cross Validation (GCV) value. In this study, order 1 is used and the determination of the optimal knot point starts with the calculation using 1, 2, 3, and a combination of 1 to 3 knot points.

Table 2. Optimum Knot Point with One Knot Point

No.	Knots					GCV
	X_1	X_2	X_3	X_4	X_5	
1.	32.86	1.88	1.72	52.94	18.06	472.76
2.	27.62	1.34	1.14	50.37	12.33	478.39
3.	35.48	2.16	2.01	54.23	20.93	482.67
4.	31.55	1.75	1.58	52.30	16.63	486.49
5.	36.79	2.29	2.16	54.88	22.36	487.29

Based on Table 2, the minimum GCV value for selecting the optimal knot point with 1 knot is 472.76, with each knot point for each variable being 32.86; 1.88; 1.72; 52.94; and 18.06.

Table 3. Optimum Knot Point with Two Knot Points

No.	Knots					GCV
	X_1	X_2	X_3	X_4	X_5	
1.	83.94	7.22	7.37	78.09	73.94	440.88
	89.18	7.76	7.95	80.67	79.67	
2.	86.56	7.49	7.66	79.38	76.80	448.17
	87.87	7.63	7.81	80.03	78.23	
3.	85.24	7.35	7.52	78.74	75.37	448.17
	87.86	7.62	7.81	80.03	78.23	
4.	85.24	7.35	7.52	78.74	75.36	448.17
	86.55	7.48	7.66	79.38	76.80	
5.	85.24	7.35	7.52	78.74	75.36	448.17
	89.17	7.76	7.95	80.67	79.66	

Based on Table 3, the minimum GCV value for selecting the optimal knot points with 2 knots is 440.88. The optimal knot points in the X_1 variable of the first-order is at 83.94 and 89.18. The optimal knots in the X_2 first-order variables are at 7.22 and 7.76. The optimal knot points in the X_3 first-order variable is at 7.37 and 7.95. The optimal knot points in the X_4 first-order variable is at 78.09 and 80.67. The optimal knot points in the X_5 variable of the first order is at 73.94 and 79.67.

Table 4. Optimal Knot Point with Three Knot Points

No.	Knots					GCV
	X_1	X_2	X_3	X_4	X_5	
1.	60.36	4.75	4.76	66.48	48.14	233.16
	78.70	6.66	6.79	75.51	68.20	
	85.24	7.35	7.52	78.74	75.36	
2.	60.36	4.75	4.76	66.48	48.14	249.87
	78.70	6.66	6.79	75.51	68.20	
	83.93	7.21	7.37	78.09	73.93	
3.	61.67	4.89	4.91	67.13	49.58	251.79
	78.70	6.66	6.79	75.51	68.20	
	85.24	7.35	7.52	78.74	75.36	
4.	60.36	4.75	4.76	66.48	48.14	253.68
	78.70	6.66	6.79	75.51	68.20	
	86.55	7.48	7.66	79.38	76.80	
5.	60.36	4.75	4.76	66.48	48.14	253.75
	77.39	6.53	6.65	74.87	66.77	
	89.17	7.76	7.95	80.67	79.66	

Based on Table 4, the minimum GCV value for selecting the optimal knot points with 3 knots is 233.16. The optimal knot points in the X_1 variable of the first-order is at 60.36; 78,70; and 85.24. The optimal knots in the X_2 first-order variables are at 4.75; 6.66; and 7.35. The optimal knot points in the X_3 first-order variable is at 4.76; 6,79; and 7.52. The optimal knot points in the X_4 first-order variable is at 66.48; 75.51; and 78.74. The optimal knot points in the X_5 variable of the first order is at 48.14; 68,20; and 75.36.

Table 5. Optimum Knot Point with Combination Knot Point

No.	Knots					GCV
	X_1	X_2	X_3	X_4	X_5	
1.	60.36 78.70 85.24	4.75 6.66 7.35	4.76 6.79 7.52	66.48 75.51 78.74	48.14 68.20 75.36	233.16
2.	60.36 78.70 85.24	4.75 6.66 7.35	4.76 6.79 7.52	78.09 80.67	48.14 68.20 75.36	270.94
3.	32.86	7.22 7.76	4.76 6.79 7.52	78.09 80.67	73.94 79.67	324.5

Based on Table 5 the determination of the optimal knot point starting from one knot point, two knot points, three knot points, to a combination of knot points, it was obtained that the optimal knot point in each predictor variable was with three knot points and the minimum GCV obtained was 233.16. After obtaining the optimal knot point based on the smallest GCV value, a parameter significance test will be carried out to find out whether the predicted variable has a significant influence on the response variable, namely the percentage of pneumonia in toddlers. Parameter testing is divided into two stages, namely simultaneous and partial testing [22].

Concurrent testing was carried out to test whether the parameter had a significant influence on the response variable by involving all parameters of the regression model. Here are the hypotheses in concurrent testing of regression model parameters:

H_0 : $\beta_1 = \beta_2 = \cdots = \beta_{20} = 0$;
H_1 : at least one $\beta_j \neq 0,\ j = 1, 2, \cdots, 20$.

Table 6. Simultaneous Test Results of Parameters

Source	Df	SS	MS	F-value	p-Value
Regression	20	14331.03	716.551	8.03731	0.0002
Error	13	1158.991	89.153		
Total	33	15490.02			

Based on the ANOVA results in Table 6, a F-value of 8.03731 with *a* p-value of 0.0002 was obtained. With a significance level of 5%, a decision to reject H_0 was obtained because the p-value < 5%. It can be concluded that there is at least one significant parameter in the spline nonparametric regression model. In addition, an R^2 value of 92.52% was obtained so that the model is a good model in modeling the percentage of pneumonia in toddlers in provinces in Indonesia.

Partial testing was carried out to find out whether individual parameters had a significant influence on the response variable, namely the percentage of pneumonia in toddlers. The hypothesis on the partial test is as follows:

H_0 : $\beta_j = 0$;
H_1 : $\beta_j \neq 0,\ j = 1, 2, \cdots, 20$.

Table 7. Partial Test Results and Parameter Assessment

Variable	Parameters	Coefficient	t-count	p-value	Decision
	β_0	154.282	2.739	0.016	Significant
X_1	β_1	0.076	0.208	0.837	Insignificant
	β_2	2.881	4.119	0.001	Significant
	β_3	-17.622	-4.700	0.000	Significant
	β_4	30.118	3.667	0.002	Significant
X_2	β_5	7.267	2.959	0.011	Significant
	β_6	-40.386	-5.807	6.100	Insignificant
	β_7	429.642	4.920	0.000	Significant
	β_8	-806.839	-4.199	0.001	Significant
X_3	β_9	-13.085	-3.946	0.001	Significant
	β_{10}	46.380	6.101	3.771	Insignificant
	β_{11}	-116.847	-4.690	0.00	Significant

Variable	Parameters	Coefficient	t-count	p-value	Decision
	β_{12}	122.913	3.069	0.008	Significant
X_4	β_{13}	-1.272	-1.698	0.113	Insignificant
	β_{14}	3.106	1.780	0.098	Insignificant
	β_{15}	36.000	4.623	0.000	Significant
	β_{16}	-79.299	-4.007	0.001	Significant
X_5	β_{17}	-1.477	-3.805	0.002	Significant
	β_{18}	2.580	4.519	0.005	Significant
	β_{19}	-3.215	-1.609	0.131	Insignificant
	β_{20}	14.669	3.622	0.003	Significant

Based on Table 7, it can be concluded that all variables have significant parameters with a significance level of 5%. There were 15 significant parameters and 6 insignificant parameters in the spline nonparametric regression model. This can be interpreted that the percentage of coverage of Vitamin A administration (X_1), the percentage of toddlers with poor nutritional status (X_2), the percentage of low-birth-weight babies (X_3), the percentage of toddlers who have received measles immunization (X_4), and the percentage of infants who receive exclusive breastfeeding (X_5) shows a statistically significant effect on the percentage of pneumonia in toddlers in each province in Indonesia.

After the parameter significance test is carried out, a residual assumption test will be carried out to find out whether the residual has met several assumptions. The residuals of the spline nonparametric regression model must meet the assumptions of identical, independent, and normally distributed (IIDN). One of the regression assumptions that must be met is the homogeneity of variance from residual (homoscedasticity). Homoskedasticity or identical means that the residual variance between the residuals must be the same or no heteroscedasticity occurs. One way to detect heteroscedasticity is to use the Glejser test. The hypotheses used in the Glejser test are as follows:

H_0 : $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_{34}^2 = \sigma^2$;
H_1 : at least one $\sigma_i^2 \neq \sigma^2,\ i = 1, 2, \cdots, 34$.

Table 8. Identical Assumption Test Results (Glejser Test)

Type	Df	SS	RMS	F-Value	p-Value
Regression	5	1188.880	237.776	1.960	0.116
Error	28	3397.247	121.330		
Total	33	4586.127			

Based on Table 8, p-value of 0.116 was obtained, which means that the value was more than alpha (α = 0.05), so the decision was to reject H_0. Thus, the conclusion is there is no heteroscedasticity. In other words, the error variance is homogeneous. Then the identical assumption has been fulfilled.

Furthermore, a normality test was carried out which aimed to test whether the residues were normally distributed or not. This test uses the Kolmogorov-Smirnov method. The hypothesis for the Kolmogorov-Smirnov test is as follows:

H_0 : Normally distributed residual;
H_1 : Residual not normally distributed.

The critical area of this test is to reject H_0 if the p-value < 0.05. The following are the results of the normality test on residuals:

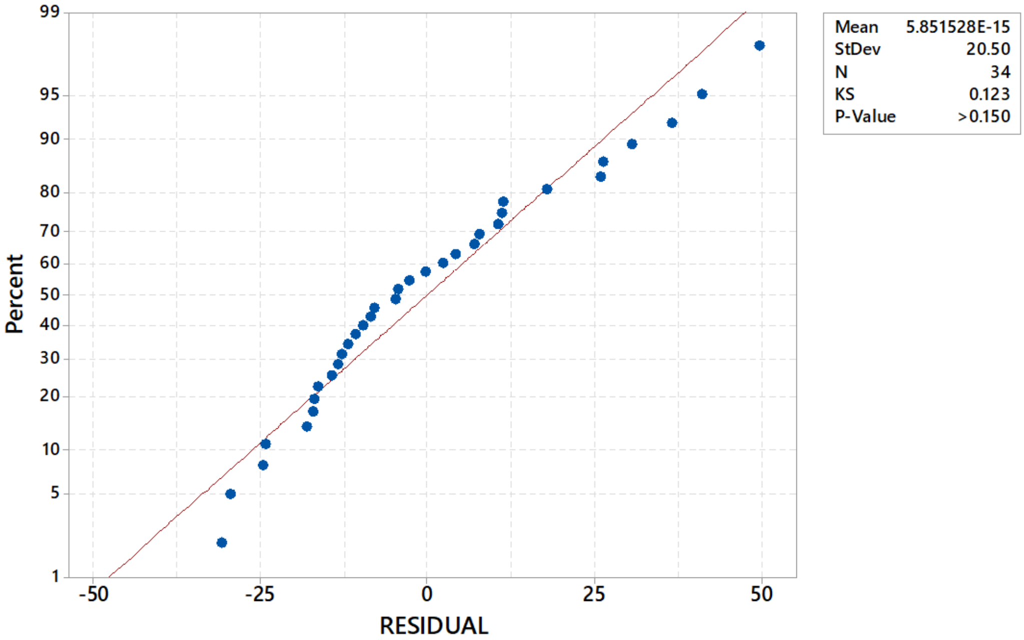

Figure 3. Results of the Normality Assumption Test

Based on Figure 3, it was found that the p-value (> 0.150) was greater than alpha (0.05), so the decision was to reject H_0. So, the conclusion is that residual data is normally distributed. This means that the assumption of normality has been fulfilled.

Following the spline approach study, a nonparametric regression method with a Fourier series approach was used for a follow-up analysis. The Fourier series analysis was performed to ascertain whether the model generated by the spline approach is better than other methods. For the analysis, $K = 1, 2, 3$ values were used. The results of the Fourier series analysis are as follows:

Table 9. Analysis Results with Fourier Series

K	GCV	R^2
1	4126.3969	21.65%
2	1172.1259	49.92%
3	451.7365	65.69%

Based on the analysis results in the Table 9, the optimal K value is K = 3, with the minimum GCV value of 451.7365 and an R^2 of 65.69%. The best model generated based on pneumonia data in toddlers in Indonesia is as follows:

$$\begin{aligned}\hat{y}_i = {} & -63{,}20 + 0.31x_{1i} + 3.40\cos x_{1i} + 6.61\cos 2x_{1i} - 13.56\cos 3x_{1i} + 3.84x_{2i} - \\ & 14.76\cos x_{2i} - 9.29\cos 2x_{2i} - 12.37\cos 3x_{2i} - 1.40x_{3i} + 10.80\cos x_{3i} + \\ & 15.30\cos 2x_{3i} - 0.96\cos 3x_{3i} + 0.83x_{4i} + 2.47\cos x_{4i} - 12{,}79\cos 2x_{4i} + \\ & 1.87\cos 3x_{4i} + 0.14x_{5i} + 6.44\cos x_{5i} - 2.41\cos 2x_{5i} - 12.29\cos 3x_{5i}\end{aligned} \quad (7)$$

When compared with the results using spline, the GCV and R^2 values are as follows:

Table 10. Comparison of Spline and Fourier Models

Model	GCV	R^2
LS-Spline	233.16	92.5%
Fourier Series	451.7365	65.69%

Based on Table 10, a comparison can be made between the performance of the spline model with three knot points for each predictor variable and the Fourier Series model with the optimal K value of 3. It can be concluded that the better model is the one using the nonparametric regression spline method with three knot points in each predictor variable, as it has a minimum GCV value of 233.16 and a higher R^2 of 92.5%.

3.2 Discussion

After testing the parameters of the LS-Spline regression model and all residual assumptions have been met, the regression model that has been obtained can be interpreted. It is known that the value of the determination coefficient or R^2 of the LS-Spline regression model is 92.5%. This means that the variables of the percentage of coverage of Vitamin A (X_1), the percentage of toddlers with poor nutritional status (X_2), the percentage of underweight babies (X_3), the percentage of toddlers who have received measles immunization (X_4), and the percentage of infants who receive exclusive breastfeeding (X_5) are able to explain the variability of the percentage of pneumonia cases of 92.5%. With this R^2 value, it can be said that the resulting LS-Spline regression model is a good model and suitable for modeling. The model obtained can be written as follows:

$$\begin{aligned}\hat{y} = {} & 154.282 + 2.881(x_1 - 60.363)_+ - 17.622(x_1 - 78.7)_+ + 30.118(x_1 - \\ & 85.249)_+ + 7.267x_2 + 429.642(x_2 - 6.669)_+ - 806.839(x_2 - 7.353)_+ - \\ & 13.085x_3 - 116.847(x_3 - 6.795)_+ + 122.913(x_3 - 752)_+ + 36(x_4 - \\ & 75.515)_+ - 79.299(x_4 - 78.740)_+ - 1.477x_5 + 2.58(x_5 - 48.148)_+ + \\ & 14.669(x_5 - 75.369)_+\end{aligned} \tag{8}$$

Model interpretation for significant variables was carried out to determine how much influence it had on the percentage of pneumonia in toddlers in 34 provinces in Indonesia. It was found that the five identified predictor variables had significantly affected the response variable, namely the percentage of pneumonia in toddlers.

Assuming other variables are constant, the effect of the percentage of coverage of Vitamin A on the percentage of pneumonia in toddlers is as follows:

$$\hat{y} = 2.881(x_1 - 60.363)_+ - 17.622(x_1 - 78.7)_+ + 30.118(x_1 - 85.249)_+, \tag{9}$$

with

$$\hat{y} = \begin{cases} 0, & x_1 < 60.363; \\ -173.909 + 2.881x_1, & 60.363 \le x_1 < 78.7; \\ 1212.976 - 14.741x_1, & 78.7 \le x_1 < 85.249; \\ 1354.632 + 15.377x_1, & x_1 \ge 85.249. \end{cases} \tag{10}$$

If the percentage of coverage of Vitamin A is between 60.363 and 78.7 percent, a 1 percent increase in Vitamin A is likely to increase the percentage of pneumonia in toddlers by 2.881 percent. In the range of 78.7 to 85.249 percent, a 1 percent increase in Vitamin A administration actually tended to reduce the percentage of pneumonia in toddlers by 14.741 percent. However, if the percentage of giving Vitamin A is more than 85.249 percent, an increase of 1 percent in the administration of Vitamin A tends to increase the percentage of pneumonia in toddlers by 15.377 percent.

Assuming other variables are constant, the effect of the percentage of toddlers with poor nutritional status on the percentage of pneumonia in toddlers is as follows:

$$\hat{y} = 7.267x_2 + 429.642(x_2 - 6.669)_+ + 122.913(x_2 - 7.353)_+, \tag{11}$$

with

$$\hat{y} = \begin{cases} 7.267x_2, & x_2 < 4.755; \\ 7.267x_2, & 4.755 \le x_2 < 6.669; \\ -2861.42 + 436.909x_2, & 6.669 \le x_2 < 7.353; \\ -3765.20 + 559.822x_2, & x_2 \ge 7.353. \end{cases} \tag{12}$$

If the percentage of toddlers with undernourished status is less than 4.755 percent, an increase of 1 percent of toddlers with undernourished status tends to increase the percentage of pneumonia cases in toddlers by 7.267 percent. For the percentage of toddlers with undernourished status between 4.755 and 6.669 percent, an increase of 1 percent of toddlers with undernourished status tends to reduce the percentage of pneumonia in toddlers by 7.267 percent. In the range of 6,669 to 7,353 percent, an increase of 1 percent of toddlers with poor nutritional status tended to increase the percentage of pneumonia in toddlers by 436.909 percent. However, if the percentage of toddlers with nutritional status is approximately 7.353 percent, an increase of 1 percent of toddlers with poor nutritional status tends to increase the percentage of pneumonia in toddlers by 559.822 percent.

Assuming other variables are constant, the effect of the percentage of low-birth-weight babies on the percentage of pneumonia in toddlers is as follows:

$$\hat{y} = -13.085x_3 - 116.847(x_3 - 6.795)_+ - 806.839(x_3 - 7.520)_+, \tag{13}$$

with,

$$\hat{y} = \begin{cases} -13.085x_3, & x_3 < 4.767; \\ -13.085x_3, & 4.767 \le x_3 < 6.795; \\ 793.98 - 129.932x_3, & 6.795 \le x_3 < 7.520; \\ 6861.41 - 936.777x_3, & x_3 \ge 7.520. \end{cases} \tag{14}$$

If the percentage of toddlers with low-birth-weight status is less than 4.767 percent, an increase of 1 percent of toddlers with low-birth-weight status tends to decrease the percentage of pneumonia cases in toddlers by 13.085 percent. For the percentage of toddlers with low-birth-weight status between 4.767 and 6.795 percent, a 1 percent increase in toddlers with low-birth-weight status tends to increase the percentage of pneumonia in toddlers by 13.085 percent. In the range of 6.795 to 7.520 percent, a 1 percent increase in toddlers with low-birth-weight status tended to reduce the percentage of pneumonia in toddlers by 129.932 percent. If the percentage of toddlers with low-birth-weight status is more than 7.520 percent, an increase of 1 percent of toddlers with poor nutritional status tends to reduce the percentage of pneumonia in toddlers by 936.777 percent.

Coefficients such as a 129.932 percent or 936.777 percent decrease do not mean that the actual percentage of pneumonia cases dropped by that amount. Instead, these values reflect how steep the slope of the model curve is when the proportion of infants with low birth weight (LBW) increases by 1 percentage point. Such large figures arise because the spline method divides the data into local segments, and within certain ranges, the relationship between variables can appear extremely sharp. These values should be interpreted as indicating the intensity and direction of the relationship—not literal percentage changes in pneumonia cases.

Assuming other variables are constant, the effect of the percentage of toddlers who have received measles immunization on the percentage of pneumonia in toddlers is as follows:

$$\hat{y} = 36(x_4 - 75.515)_+ - 79.299(x_4 - 78.740)_+ \tag{15}$$

with,

$$\hat{y} = \begin{cases} 0, & x_4 < 66.487; \\ 0, & 66.487 \le x_4 < 75.515; \\ -2718.54 + 36x_4, & 75.515 \le x_4 < 78.740; \\ 6244 - 79.229x_4, & x_4 \ge 78.740. \end{cases} \tag{16}$$

If the percentage of toddlers with poor nutritional status is in the range of 75.515 to 78.740 percent, an increase of 1 percent of toddlers who have received measles immunization tends to increase the percentage of pneumonia in toddlers by 36 percent. However, if the percentage of toddlers who have received measles immunization is more than 78.740 percent, an increase of 1 percent of toddlers with less nutritional status tends to reduce the percentage of pneumonia in toddlers by 79.229 percent.

Assuming other variables are constant, the effect of the percentage of toddlers who have received exclusive breastfeeding on the percentage of pneumonia in toddlers is as follows:

$$\hat{y} = -1.477x_5 + 2.58(x_5 - 48.148)_+ + 14.669(x_5 - 75.369)_+, \tag{17}$$

with

$$\hat{y} = \begin{cases} -1.477x_5, & x_5 < 48.148; \\ -124.222 + 1.103x_5, & 48.148 \le x_5 < 68.206; \\ -124.222 + 1.103x_5, & 68.206 \le x_5 < 75.369; \\ -1229.81 + 15.772x_5, & x_5 \ge 75.369. \end{cases} \tag{18}$$

If the percentage of toddlers who have received exclusive breastfeeding is less than 48.148 percent, an increase of 1 percent of toddlers who have received exclusive breastfeeding tends to decrease the percentage of pneumonia cases in toddlers by 1.477 percent. For the percentage of toddlers who have received exclusive breastfeeding between 48.148 and 68.206 percent, an increase of 1 percent of toddlers who have received exclusive breastfeeding tends to increase the percentage of pneumonia in toddlers by 1.103 percent. In the range of 68,206 to 75,369 percent, an increase of 1 percent of toddlers who had received exclusive breastfeeding tended to reduce the percentage of pneumonia in toddlers by 1.103 percent. However, if the percentage of toddlers who have received exclusive breastfeeding is more than 75.369 percent, an increase of 1 percent of toddlers who have received exclusive breastfeeding tends to increase the percentage of pneumonia in toddlers by 15.772 percent.

The results of this study align with previous findings, which show that the coverage of vitamin A supplementation, the percentage of undernourished toddlers, the percentage of low-birth-weight infants, the percentage of toddlers who have received measles immunization, and the percentage of infants who received exclusive breastfeeding are all significantly associated with pneumonia in toddlers. A study by [5] states that exclusive breastfeeding and a history of low birth weight significantly influence pneumonia in toddlers in West Lombok. Furthermore, the relationship between nutritional status, vitamin A coverage, and measles immunization is consistent with findings by [7], who also reported that these variables are significant factors in pneumonia in toddlers in Surabaya. By utilizing data from a broader coverage area, including all provinces in Indonesia, this study contributes new insights by demonstrating that the effects of these factors on pneumonia in Indonesian toddlers are non-linear and vary depending on the level of prevalence. Vitamin A supplementation, poor nutritional status, low birth weight (BBLR), measles immunization, and exclusive breastfeeding show varying effects depending on their coverage levels. At moderate coverage levels, vitamin A and measles immunization appear to have a protective effect. In contrast, at very low or very high levels, they are associated with an increase in pneumonia. Low-birth-weight babies and toddlers with poor nutritional condition were identified as important risk factors, and at some prevalence levels, the rates of pneumonia increased sharply. Exclusive breastfeeding, on the other hand, generally has a protective effect, although pneumonia rates tend to rise at very high coverage levels. These results highlight the value of locally specific health treatments and show the complex interactions between variables that call for a more adaptable analytical approach, such as the nonparametric regression employed in this study.

4. CONCLUSION

Based on the results of the analysis and discussion that has been carried out, it was confirmed that West Papua has the highest pneumonia prevalence among toddlers at 75%, while North Sulawesi has the lowest with a percentage of 0.3%. The optimal knot point was obtained, namely with three knot points in each predictor variable with a GCV value of 233.16. The value of the determination coefficient of the LS-Spline regression model is 92.5%. This means that the variables of the percentage of coverage of Vitamin A administration (X_1), the percentage of toddlers with poor nutritional status (X_2), the percentage of low-birth-weight babies (X_3), the percentage of toddlers who have received measles immunization (X_4), and the percentage of infants who receive exclusive breastfeeding (X_5) are able to explain the variability of the percentage of pneumonia cases in toddlers of 92.5%. When compared to the model using the Fourier series method with K = 3, which yields an R^2 value of 65.69%, this value is lower than that obtained using the spline method. With this R^2 value, it can be said that the resulting LS-Spline regression model is a good model and suitable for modeling. The best model obtained in this study indicates that the percentage of pneumonia in toddlers in Indonesia is influenced by various factors with different relationship patterns in each percentage range, as demonstrated by a coefficient of determination of 92.5%. Vitamin A administration tends to increase pneumonia in some ranges, but decreases it in others, depending on the level of coverage. Toddlers with undernourished status or low birth weight have a complex relationship, with significant effects on the increase or decrease of pneumonia based on a certain percentage range. Measles immunization in toddlers and exclusive breastfeeding also have a varied effect, both increasing and decreasing pneumonia in toddlers, depending on the proportion of the population. This relationship reflects the non-linear dynamics between child health factors and pneumonia in toddlers.

Author Contributions

Toha Saifudin: Conceptualization, Funding Acquisition, Methodology, Supervision, Validation, Writing - Review and Editing. Suliyanto: Conceptualization, Writing-Review and Editing, Supervision. Nabila Nurdin: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing - Original Draft, Writing - Review and Editing. Bryan Given Christiano Ginzel: Data Curation, Formal Analysis, Investigation, Methodology, Visualization, Software, Writing - Original Draft. Sabrina Salsa Oktavia: Resources, Visualization, Writing - Original

Draft. Jovansha Ariyawan: Resources, Visualization. Mohammad Noufal Ubadah: Resources, Visualization. All authors discussed the results and contributed to the final manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgment

The authors expressed their gratitude to the Ministry of Health of the Republic of Indonesia for providing publications with relevant data so that they could be used in the preparation of this research.

Declarations

The authors declare no competing interest.

REFERENCES

[1] World Health Organization, "PNEUMONIA IN CHILDREN," *World Health Organization*, [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/pneumonia [Accessed: 12 September 2024].

[2] Y. Ayele, "DETERMINANTS OF COMMUNITY-ACQUIRED PNEUMONIA AMONG 2-59 MONTHS OLD CHILDREN ATTENDING HEALTH FACILITY IN HOSSAENA TOWN, ETHIOPIA," *Biomedical Journal of Scientific & Technical Research*, vol. 43, no. 2, Apr. 2022, doi: https://doi.org/10.26717/BJSTR.2022.43.006867.

[3] Ministry of Health of the Republic of Indonesia, "INDONESIA HEALTH PROFILE 2020," *Ministry of Health of the Republic of Indonesia*, [Online]. Available: https://www.kemkes.go.id/id/profil-kesehatan-indonesia-2020 [Accessed: 12 September 2024].

[4] R. Widoyo, "MEASLES IMMUNIZATION AND VITAMIN A FOR PREVENTION OF PNEUMONIA IN INDONESIA," *Kesmas: National Public Health Journal*, vol. 11, no. 1, p. 46, Aug. 2016, doi: https://doi.org/10.21109/kesmas.v11i1.841.

[5] M. Erliandani, R. I. P. Priono, S. Ruqayyah, and A. F. Benvenuto, "THE RELATIONSHIP BETWEEN THE HISTORY OF EXCLUSIVE BREASTFEEDING, LOW BIRTH WEIGHT, AND THE PHYSICAL CONDITION OF THE HOUSE WITH THE INCIDENCE RATE PNEUMONIA IN TODDLERS," *Jambura Journal of Health Sciences and Research*, vol. 5, no. 2, pp. 746–754, Apr. 2023, doi: https://doi.org/10.35971/jjhsr.v5i2.19141.

[6] M. Nurnajiah, R. Rusdi, and D. Desmawati, "THE RELATIONSHIP BETWEEN NUTRITIONAL STATUS AND THE DEGREE OF PNEUMONIA IN TODDLERS IN HOSPITALS. DR. M. DJAMIL PADANG," *Jurnal Kesehatan Andalas*, vol. 5, no. 1, Jan. 2016, doi: https://doi.org/10.25077/jka.v5i1.478.

[7] A. C. D. Nugroho, *ANALYSIS OF FACTORS AFFECTING PNEUMONIA IN TODDLERS IN THE CITY SURABAYA USES SPLINE NONPARAMETRIC REGRESSION*. Thesis, Faculty of Mathematics and Natural Sciences, Sepuluh Nopember Institute of Technology, 2015. [Online].

[8] Y. Zhuge *et al.*, "RESIDENTIAL RISK FACTORS FOR CHILDHOOD PNEUMONIA: A CROSS-SECTIONAL STUDY IN EIGHT CITIES OF CHINA," *Environment International*, vol. 116, pp. 83–91, Jul. 2018, doi: https://doi.org/10.1016/j.envint.2018.03.022.

[9] V. N. Sutriana, M. N. Sitaresmi, and A. Wahab, "RISK FACTORS FOR CHILDHOOD PNEUMONIA: A CASE-CONTROL STUDY IN A HIGH PREVALENCE AREA IN INDONESIA," *Clinical and Experimental Pediatrics*, vol. 64, no. 11, pp. 588–595, Nov. 2021, doi: https://doi.org/10.3345/cep.2020.00339.

[10] R. Amelia, D. Wahyuni, and E. Julianti, "IMPLEMENTATION OF VECTOR AUTOREGRESSIVE (VAR) AND VECTOR ERROR CORRECTION MODEL (VECM) METHOD IN PNEUMONIA PATIENTS WITH WEATHER ELEMENTS IN PANGKALPINANG CITY", *BAREKENG: J. Math. & App.*, vol. 18, no. 3, pp. 1447-1458, Jul. 2024, doi https://doi.org/10.30598/barekengvol18iss3pp1447-1458

[11] N. Chamidah and B. Lestari, *NONPARAMETRIC REGRESSION ANALYSIS WITH R SOFTWARE*, Surabaya: Airlangga University Press, 2022

[12] A. T. R. Dani, L. Ni'matuzzahroh, V. Ratnasari, and I. N. Budiantara, "MODELING OF NONPARAMETRIC REGRESSION SPLINE TRUNCATED ON LONGITUDINAL DATA," *Inferensi*, vol. 4, no. 1, p. 47, Mar. 2021, doi: https://doi.org/10.12962/j27213862.v4i1.8737.

[13] 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, Aug. 2022, doi: https://doi.org/10.3389/fepid.2022.975380.

[14] Indonesian Ministry of National Development Planning, "SDGS - KEMENTERIAN PPN/BAPPENAS?", Bappenas, [Online]. Available: https://sdgs.bappenas.go.id [Accessed: 12 September 2024].

[15] Sugiyono, S. (2013). EDUCATIONAL RESEARCH METHODS QUANTITATIVE, QUALITATIVE AND R&D APPROACHES. Bandung. Alfabeta.

[16] Ministry of Health of the Republic of Indonesia, "INDONESIA HEALTH PROFILE 2023", *Ministry of Health of the Republic of Indonesia*, [Online]. Available: https://www.kemkes.go.id/id/profil-kesehatan-indonesia-2023 [Accessed: 10 September 2024].

[17] N. A. Kurniawati and I. N. Budiantara, "MODELING THE OPEN UNEMPLOYMENT RATE IN WEST JAVA PROVINCE USING TRUNCATED SPLINE NONPARAMETRIC REGRESSION," *Jurnal Sains dan Seni ITS*, vol. 8, no. 2, Feb. 2020, doi: https://doi.org/10.12962/j23373520.v8i2.48239.

[18] W. Nurmalisa and I. M. Nur, "NONPARAMETRIC SPLINE REGRESSION MODELING WITH OPTIMAL KNOT SELECTION USING UNBIASED RISK AND GENERALIZED CROSS-VALIDATION METHODS: A CASE STUDY OF THE HUMAN DEVELOPMENT INDEX IN CENTRAL JAVA," *Proc. National Student Seminar at UNIMUS*, pp. 344-355, Oct. 5, 2019.

[19] U. Rohma, A. A. R. Fernandes, S. Astutik, and S. Solimun, "DEVELOPMENT OF NONPARAMETRIC PATH FUNCTION USING HYBRID TRUNCATED SPLINE AND KERNEL FOR MODELING WASTE-TO-ECONOMIC VALUE BEHAVIOR", *BAREKENG: J. Math. & App.*, vol. 19, no. 1, pp. 331-344, Jan. 2025, doi: https://doi.org/10.30598/barekengvol19iss1pp331-344.

[20] N. P. A. M. Mariati, *FOURIER SERIES AND TRUNCATED SPLINE REGRESSION MODELING IN MULTIVARIABLE REGRESSION (APPLICATION: POVERTY DATA IN PAPUA PROVINCE).* Thesis, Sepuluh Nopember Institute of Technology, 2015. [Online].

[21] R Core Team, "R: A LANGUAGE AND ENVIRONMENT FOR STATISTICAL COMPUTING," 2023, *R Foundation for Statistical Computing, Vienna, Austria*. [Online]. Available: https://www.r-project.org/

[22] R. L. Toruan, *SIMULTANEOUS HYPOTHESIS TESTING OF PARAMETER IN TRUNCATED SPLINE NONPARAMETRIC REGRESSION MODEL FOR LONGITUDINAL DATA (APPLICATION ON DATA OF ECONOMIC GROWTH IN PAPUA PROVINCE IN 2011–2016).* Thesis, Sepuluh Nopember Institute of Technology, 2018. [Online].