BAREKENG: Journal of Mathematics and Its Applications
September 2026 Volume 20 Issue 3 Page 1821-1836
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss3pp1821-1836

APPLICATION OF TRUNCATED SPLINE AND FOURIER SERIES IN MODELING THE OPEN UNEMPLOYMENT RATE IN KALIMANTAN ISLAND

Aridha Pebriani Kusmiran[1*], Nur Salam[2], Agus Muslim[3]

[1,2]*Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Lambung Mangkurat*
Jln. A. Yani KM.36, Banjarbaru, 70713, Indonesia

[3]*Central Statistics Agency of Banjarbaru*
Jln. Pangeran Suriansyah Ujung, Banjarbaru, 70711, Indonesia

*Corresponding author's e-mail: * **aridha0904@gmail.com***

Article Info

Article History:
Received: 3rd June 2025
Revised: 11th December 2025
Accepted: 16th March 2026
Available online: 8th April 2026

Keywords:
Fourier series;
Kalimantan island;
Open unemployment rate;
Spline truncated.

ABSTRACT

Unemployment is a complex problem because it is one of the benchmarks for measuring the success of a region's economic development. According to the National Medium-Term Development Plan 2020-2024 and the Government Work Plan Summary 2023, Kalimantan Island is expected to have a maximum Open Unemployment Rate of 3.4%. However, provinces in Kalimantan Island still have an Open Unemployment Rate above 3.4%. Therefore, an analysis is needed to model the Open Unemployment Rate. The analytical method used is nonparametric regression, as the correlation pattern between Open Unemployment Rate and each predictor variable forms a random pattern. The estimators to be used are the truncated spline and fourier series. This study aims to compare the truncated spline and Fourier series methods in modeling the Open Unemployment Rate and determine the variables that affect Open Unemployment Rate on Kalimantan Island in 2023. Both the truncated spline and fourier series estimators have good flexibility. Based on the modelling results, it was concluded that the truncated spline method is better. The truncated spline model yields a lower GCV value (1,67545), a higher R^2 (0.57399), and a smaller MAE (0.821277) than the Fourier series model. Based on the analysis results, all predictor variables used, namely the Human Development Index, District Minimum Wage, population, and economic growth, significantly effect the Open Unemployment Rate with a coefficient of determination of 57.4%.

How to cite this article:

A. P. Kusmiran, N. Salam and A Muslim., "APPLICATION OF TRUNCATED SPLINE AND FOURIER SERIES IN MODELING THE OPEN UNEMPLOYMENT RATE IN KALIMANTAN ISLAND", *BAREKENG: J. Math. & App.*, vol. 20, no. 3, pp. 1821-1836, Sep, 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

The public prosperity reflects the success of national development. In improving public prosperity, one of the efforts that can be undertaken is to reduce the unemployment rate. Unemployment is a situation where someone who is able and wants to work cannot find a job that suits their qualifications and desires. Unemployment has negative impacts on individuals and society, such as reduced income, health and quality of life, as well as increased crime, social conflict and fiscal burdens [1]. One indicator that illustrates the level of unemployment is the Open Unemployment Rate. Based on the 2020-2024 National Medium-Term Development Plan, the Open Unemployment Rate target for the national level is in the range of 3.6-4.3% and a maximum of 3.4% for Kalimantan Island. However, data from the Central Statistics Agency in 2023 shows that 35 out of 56 regencies/cities and all provinces on Kalimantan Island have an Open Unemployment Rate above the target.

Unemployment can be reduced by identifying and addressing the factors that affect the unemployment rate, such as population, Human Development Index, Regency Minimum Wage, and economic growth, as shown in previous studies. Khoiriah et al. found that the population and minimum wage have a significant positive effect on the unemployment rate in Lampung Province [2]. Human Development Index has a positive and significant impact on the Open Unemployment Rate based on research by Wee et al. in Minahasa Regency, whereas economic growth has a negative and significant effect [3]. In addition, Rusydan et al. found that minimum wage has a significant positive effect on the Open Unemployment Rate in Sidoarjo Regency, whereas population has a negative and significant effect [4]. Meanwhile, Sinha showed that economic growth negatively affects India's unemployment rate in 1990-2023 [5].

The relationship between Open Unemployment Rate and the factors will be analyzed using nonparametric regression because the relationship pattern is random or unclear. The estimators used are the truncated spline estimator and the Fourier series estimator. The truncated spline estimator is an approach to the data plot while maintaining the curve's smoothness [6]. Meanwhile, the Fourier series has flexibility in modeling data patterns whose oscillation shape is unknown [7]. Some studies have compared spline regression and Fourier series. Mocodompis et al. modeled Gross Regional Domestic Product (GRDP) in Indonesia. and found that the spline model with three knots achieved a higher coefficient of determination (99%) compared to the Fourier series model (96%) [8]. Dani et al. modeled Case Fatality Rate due to dengue fever in Indonesia and found that the truncated spline model had a higher coefficient of determination (91.80%) than the Fourier series model (65.44%) [9]. Open Unemployment Rate modeling using truncated spline by Suhada et al. obtained a coefficient of determination of 94% [10], while Open Unemployment Rate modeling using the Fourier series method by Rahmania et al. obtained a coefficient of determination of 74.22% [11].

Despite the extensive findings regarding unemployment determinants in the previous studies, most relied on parametric approaches that assume linear relationships, potentially overlooking complex, non-standard patterns inherent in economic data. Furthermore, while comparative assessments of Spline and Fourier Series exist, they are either restricted to non-labor sectors (such as health and GRDP) or apply these estimators to unemployment data in isolation without a direct performance benchmark. To bridge this gap, this study applies nonparametric regression to handle data flexibility and conducts a rigorous, head-to-head comparison between Truncated Spline and Fourier Series to empirically determine the superior estimator for the specific context of Kalimantan Island.

While previous studies suggest that the truncated spline estimator tends to yield better results, this research seeks to validate this specifically for the Open Unemployment Rate on Kalimantan Island in 2023 and determine the significant affecting variables. The optimal parameters in both methods are defined using the Generalized Cross Validation (GCV) criteria, while the best method is selected based on the Mean Square Error (MSE), Mean Absolute Error (MAE), and coefficient of determination values. These evaluation metrics collectively assess both the accuracy and the explanatory power of each method. Consequently, the model with the lowest MSE and MAE, alongside the highest coefficient of determination, will be deemed the optimal estimator. The remainder of this paper first presents the research methods, followed by the results and discussion covering model development, comparison, and interpretation. The final part concludes the study.

2. RESEARCH METHODS

The research uses secondary data obtained from the Central Statistics Agency and the Provincial Employment Service on Kalimantan Island in 2023. The dataset consists of 56 observation units, covering 9 cities and 47 regencies. The response variable is the Open Unemployment Rate (y), representing the percentage of the labor force that is unemployed based on annual Central Statistics Agency publications. The predictor variables consist of the Human Development Index (HDI) (x_1), obtained from the annual human development report and reflecting achievements in health, education, and living standards; the Regency Minimum Wage (x_2), sourced from official provincial publications and statistical reports and expressed in hundred thousand rupiah; Population (x_3), obtained from BPS demographic statistics and reported in tens of thousands of residents; and Economic Growth (x_4), derived from regional Gross Domestic Product (GDP) statistics and expressed as the annual growth rate at constant prices. All variables are available completely for the 56 regions analyzed. The data processing procedure for the 56 observation units, ranging from the input stage to the drawing of conclusions is systematically illustrated in the Fig. 1 below.

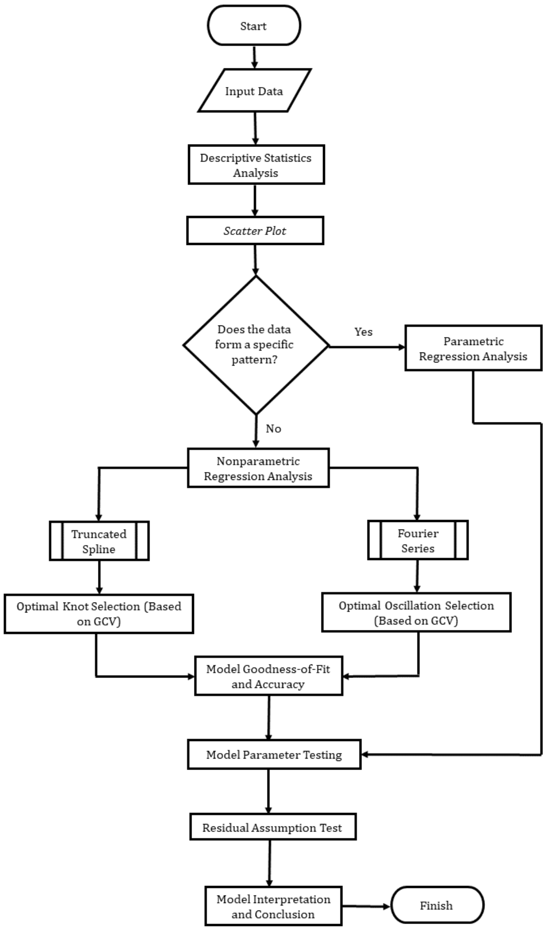

Figure 1. Research Flowchart

2.1 Spline Truncated Nonparametric Regression

A spline is a polynomial with distinct segments linked at knots and is continuous, making it more flexible than a regular polynomial [12]. Knot points are certain points that are used to divide the independent variable into different segments in the spline model [13]. The truncated spline function is obtained from the combination of a polynomial and a truncated function with a general regression model:

$$y_i = \sum_{j=0}^{m} \beta_j x_i^j + \sum_{k=1}^{r} \beta_{k+m}(x_i - K_k)_+^m + \varepsilon_i \,; i = 1,2, \dots, n, \tag{1}$$

where β_j is the polynomial model parameter ($j = 1, 2, \dots, m$), β_{k+m} is the truncated component parameter (k = 1,2, …, r), K_k is the k-th knot point, and x_i is the i-th observation.

2.2 Fourier Series Nonparametric Regression

The Fourier series is a flexible trigonometric polynomial that adapts to local properties of the data. The Fourier series is well used to explain curves that show sine and cosine waves [14]. The Fourier series estimator is often used there is a tendency to have a repeating pattern [15]. The Fourier series combines sine and cosine functions with linear functions to achieve efficiency in estimation [16]. In Fourier series modelling, the oscillation points work as essential parameters that determine the shape and characteristics of the estimated function [17]. The general equation of the Fourier series nonparametric regression model is as follows:

$$f(x_i) = \frac{1}{2}\beta_0 + \beta x_i + \sum_{q=1}^{Q} \alpha_q cosqx_i, \tag{2}$$

where β_0 is a constant, β is a regression parameter, and q is the oscillation point for $q = 1,2, \dots, Q$.

2.3 Generalized Cross Validation (GCV)

The selection of optimal knot points and oscillation points uses Generalized Cross Validation (GCV), which optimizes the balance between model fit and smoothness. The selection of GCV is based on having asymptotically optimal properties when compared to other methods such as Cross Validation (CV) and Unbiased Risk (UBR) [17][18]. GCV formula is expressed as:

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

where $\mathbf{I}$ is the identity matrix and $\mathbf{A_K} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$.

2.4 Model Parameter Test

Parameter tests are meant to significantly determine the influence of variables on the model. Parameter testing is carried out simultaneously and individually (partially).

Simultaneous Test

Simultaneous test is a test of significance to determine whether or not the independent variables in the model together have a significant influence on the dependent variable [19]. The hypotheses in the simultaneous test are as follows.

$H_0: \beta_1 = \beta_2 = \cdots = \beta_{p+r} = 0$
$H_1:$ at least $\beta_s \neq 0, s = 1, 2, \dots p + r$

The value $p + r$ is the number of parameters in the regression model, with the following test statistic.

$$F = \frac{MS_{regression}}{MS_{residual}}. \tag{4}$$

The rejection is of H_0 is when $F > F_{\alpha(p+r, n-(p+r)-1)}$ or if $p_{value} < \alpha$. If H_0 is rejected, it means that at least one parameter in the regression model is significant.

Partial Test

Partial test is used to determine the effect of each independent variable on the dependent variable [19], with the hypothesis:

$H_0: \beta_s = 0$
$H_1: \beta_s \neq 0, s = 1, 2, \dots p + r$

The test statistics in the partial significance test is as follows.

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}. \tag{5}$$

with $SE(\hat{\beta}_j)$ is the standard error from $(\hat{\beta}_j)$. H_0 is rejected if $|t| > t_{(\frac{\alpha}{2},(n-(p+r)-1)}$ or $p-value < \alpha$ which means that the parameter significantly affects the model.

2.5 Residual Assumption Test

The residual assumption test aims to determine whether the residuals of the regression model fulfil the identical, independent, and normally distributed assumptions.

Identical Assumptions

Identical Assumptions (heteroscedasticity) are satisfied if the residual variances are equal. The heteroscedasticity test aims to test whether in the regression model there is inequality in the variance of the residuals from one observation to another [19]. The Glejser test is one method to test for identical assumptions with the following test hypothesis.

$H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_n^2 = \sigma^2$
H_1: At least one $\sigma_i^2 \neq \sigma^2; i = 1,2,\ldots,n$

The test statistics used is as follows.

$$F = \frac{\frac{\sum_{i=1}^{n}(|\varepsilon_i| - |\bar{\varepsilon}|^2)}{v}}{\frac{\sum_{i=1}^{n}(|\varepsilon_i| - |\hat{\varepsilon}_i|^2)}{n-v-1}}, \quad (6)$$

v is the number of parameters of the Glejser model and $v = p + r$. If $F > F_{\alpha;(v,n-v-1)}$ or $p-value < \alpha$ then the decision is to accept H_0, which means that the residuals indicate that there is no case of heteroscedasticity (identical assumptions are satisfied).

Independent Assumptions

The independent assumption is an assumption that requires no correlation between residuals. This test can use the Durbin Watson test with the following hypothesis:

$H_0: \rho = 0$ (residuals are independent)
$H_1: \rho \neq 0$ (residuals are not independent)

The test statistics used is as follows.

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}. \quad (7)$$

The d value is compared with the lower limit (d_L) and upper limit (d_U) of the Durbin Watson table for decisions as follows.

$d < d_L$: reject H_0
$d > 4 - d_L$: reject H_0
$d_U < d < 4 - d_L$: accept H_0
$d_L \leq d \leq d_U$ and $4 - d_U \leq d \leq 4 - d_L$: inconclusive

Furthermore, the residual scatterplot against the fitted values also supports this conclusion, as the points are randomly dispersed without forming any visible pattern.

Normality Distribution

The Normal distribution test is used to determine whether the residuals are normally distributed or not [19]. The test uses the Kolmogorov-Smirnov test with the following hypothesis.

$H_0: F_n(x) = F_0(x)$ (residuals are normally distributed)
$H_1: F_n(x) \neq F_0(x)$ (residuals are not normally distributed)

The test statistics used is as follows.

$$D = \sup_{x} |F_0(x) - S_n(x)|, \quad (8)$$

with $F_0(x)$ is the cumulative frequency distribution function, and $S_n(x)$ is the division between the number of observations equal to or less than $x(k)$ and the number of observations (n). If $D > D_\alpha$ or $p - value < \alpha$, then H_0 is rejected, which means that the residuals are not normally distributed. Otherwise, the residuals are considered normally distributed.

2.6 Model Goodness Measure

The model goodness measure will be gauged based on the R^2 value, and the model accuracy will be evaluated based on the Mean Squared Error (MSE) and Mean Absolute Error (MAE) values.

Coefficient of Determination (R^2)

Coefficient of Determination (R^2) measures the contribution of predictor variables to the response variables [20]. R^2 value can be calculated with using the following formula.

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{9}$$

with SSR is the regression sum of squares, SST is the total sum of squares, y_i is the i-th response observation, $\bar{y}$ is the average of the response variables, and $\hat{y}_i$ is the estimated value of the i-th response variable.

Mean Squared Error (MSE)

MSE is used to calculate the model's average squared error in prediction using the following formula.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \tag{10}$$

Mean Absolute Error (MAE)

MAE is the average absolute difference between the actual values and predicted values. The MAE value can be calculated with the following formula.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|. \tag{11}$$

3. RESULTS AND DISCUSSION

This chapter will discuss the results and analysis of the application of nonparametric regression by comparing the truncated spline method and Fourier series method in modeling the Open Unemployment Rate on Kalimantan Island in 2023.

3.1 Descriptive Statistics

An overview of the variables in this study can be seen in Table 1.

Table 1. Data Characteristics

	y (%)	x_1	x_2 (hundred thousand rupiah)	x_3 (tens of thousands of people	x_4 (%)
Mean	4.34	73.25	31.91	30.82	5.14
Variance	2.97	16.43	9.568802	390.31	13.25
Minimum	2.07	66.06	26.0647222	2.747	1.10
Maximum	8.92	82.61	40.5535662	85.063	29.85

Based on Table 1, the highest Open Unemployment Rate in Kalimantan Island in 2023 is in Pontianak City at 8.92%. while the lowest Open Unemployment Rate is in Pulang Pisau Regency and North Paser Regency at 2.07%. In general, the quality of human resources is in the high category, with the average Human Development Index (x_1) at 73.25. Meanwhile, the average Region Minimum Wage (x_2) is around Rp3,191,000, with a wide range between the minimum of Rp2,606,472 and the maximum of Rp4,055,356, indicating a significant difference across the regions. The variables exhibiting extreme variation are

Population (x_3) and Economic Growth (x_4). Population has a very broad range (from 27,470 to 850,630 people), which highlights substantial differences in labor supply. Similarly, Economic Growth demonstrates high volatility with an average of 5.14% but a range spanning from 1.10% to 29.85%, implying that the pace and quality of economic growth are highly uneven throughout the regions.

Regarding the data scale, standardization or normalization was not considered necessary for variables like Population (x_3) and Region Minimum Wage (x_2), as they are already presented in units of tens of thousands and hundreds of thousands, respectively. This approach was chosen to allow for the direct estimation and interpretation of non-linear thresholds. Rescaling the data would strip these thresholds of their direct empirical meaning, complicating immediate interpretation. Furthermore, the current scale is deemed adequate to ensure the numerical stability of the model . The data pattern between Open Unemployment Rate and each predictor variable is as follows.

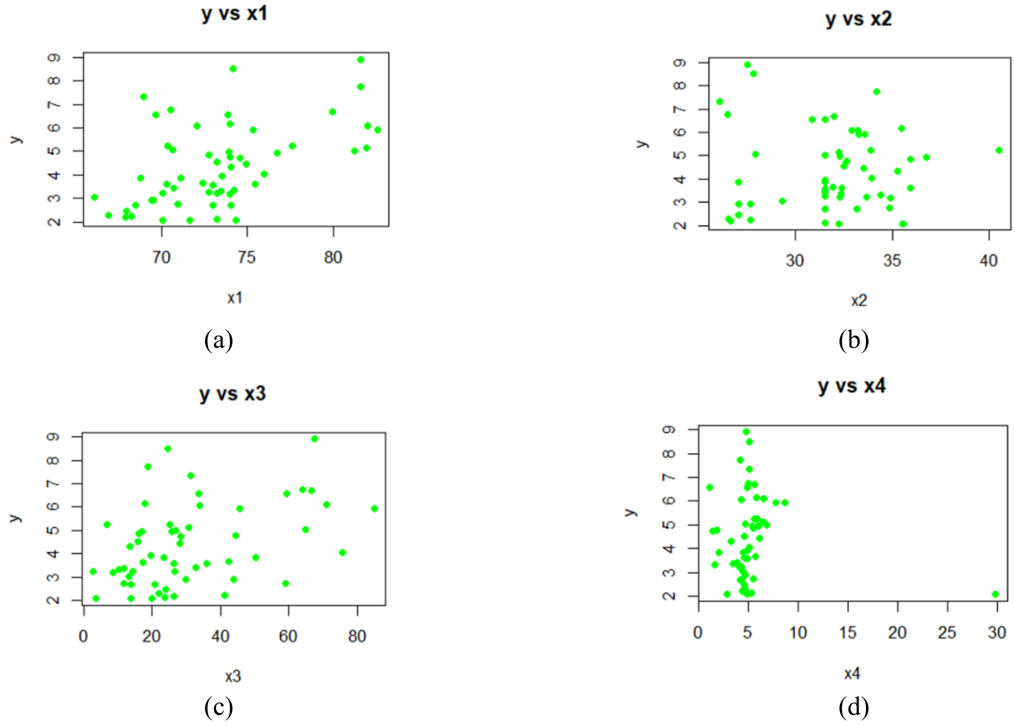

Figure 2. Scatter Plot of Open Unemployment Rate with Each Variable
(a) Scatter Plot Between Variables y and x1. (b) Scatter Plot Between Variables y and x2. (c) Scatter Plot Between Variables y and x3. (d) Scatter Plot Between Variables y and x4

The scatterplot of Open Unemployment Rate with each predictor variable in Fig. 2 shows a random pattern, so modeling will be applied using nonparametric regression.

3.2 Truncated Spline Nonparametric Regression Modeling

The optimal truncated spline model is determined based on the knot points with minimum GCV. This study tested one. two. and three knot points and the selection results are as follows.

Table 2. Comparison of GCV Values for Each Knot Point

Knot	GCV
1	2.14398
2	2.18794
3	1.67545

Based on Table 2, the optimal spline truncated nonparametric regression model uses three knot points as it yielded the smallest GCV value of 1.67545, with the following equation.

$$\hat{y}_i = -5.1548022 + 1.8613284x_{1i} - 5.2872918(x_{1i} - 68.8016)_+ + 6.2415425(x_{1i} - 70.183)_+ - 2.6104044(x_{1i} - 70.304)_+ - 4.3104251x_{2i} + 16.2783852(x_{2i} - 27.2228)_+ - 38.6922813(x_{2i} - 30.366)_+ + 26.7111639(x_{2i} - 31.2305)_+ - 0.1083037x_{3i} + 6.4468373(x_{3i} - 12.3498)_+ -$$

$39.7216624(x_{3i} - 15.98)_+ + 33.3894554(x_{3i} - 16.5699)_+ - 1.2965789x_{4i} + 12.4926949(x_{4i} - 3.3496)_+ - 14.4965327(x_{4i} - 4.1768)_+ + 3.2156656(x_{4i} - 4.2207)_+.$

3.3 Fourier Series Nonparametric Regression Modeling

The optimal Fourier series model is determined based on the oscillation point with the minimum GCV. This study tested one. two. and three oscillation points with the following selection results.

Table 3. Comparison of GCV Values for Each Oscillation Point

Oscillation	GCV
1	76.81225
2	32.39693
3	15.96922

Based on Table 3, the optimal Fourier series nonparametric regression model is uses three oscillation points as it yielded the smallest GCV value of 15.96922, with the following equation.

$$\hat{y_i} = -12.38422 + 0.31125x_{1i} - 0.198007cosx_{1i} - 0.054811cos2x_{1i} - 0.303254cos3x_{1i} - 0.173335x_{2i} - 0.261691cosx_{2i} - 0.314414cos2x_{2i} + 0.092675cos3x_{2i} - 0.000487x_{3i} + 0.096462cosx_{3i} + 0.27459cos2x_{3i} + 0.611244cos3x_{3i} - 0.091979x_{4i} + 0.513571cosx_{4i} - 0.182313cos2x_{4i} - 0.618588cos3x_{4i}.$$

3.4 Best Model Selection

Table 4. Best Model Selection

Model	GCV	R^2	MSE	MAE
Truncated Spline	1.67545	0.57399	1.78162	0.821277
Fourier Series	15.96922	0.552414	1.30361	0.908985

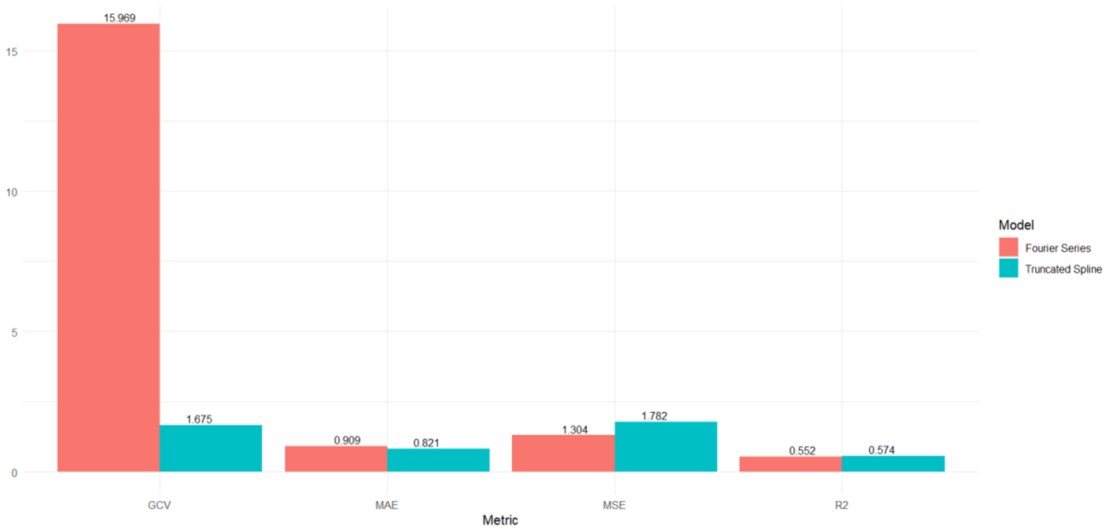

Figure 3. Performance Model

Based on Table 4 and Fig. 3, the Truncated Spline model is selected as the best model because it had the minimum GCV value (1.67545), the maximum coefficient of determination (0.57399), and the minimum Mean Absolute Error (MAE) (0.821277), compared to the Fourier Series model. The significantly higher GCV observed in the Fourier Series model reflects a fundamental mismatch between the model's assumptions and the data's nature. Fourier Series are inherently designed for oscillating patterns, this characteristic renders them inadequate to capture the irregular, threshold-driven dynamics of regional economic indicators, leading to a higher estimation error compared to the Truncated Spline method.

3.5 Model Parameter Test

Concurrent Test

The simultaneous test of the truncated spline nonparametric regression model with three knots at =0.05 shows the results as listed in Table 5 below.

Table 5. ANOVA Table

	df	Sum of Squares	Mean Square	F_{test}	p_{value}
Regression	16	93.61884	5.851178	3.284196	0.0012279
Residuals	39	69.48304	1.781616		
Total	55	163.1019			

Based on Table 5, the value of $F_{test}(3.284196) > F_{0.05(16;39)}(1.910658)$ and the value of p_{value} (0.0012279) < α. Which shows that H_0 is rejected or there is at least one parameter that has a significant effect on the response variable.

Partial Test

The partial test of the truncated spline nonparametric regression model with three knots at α = 0.05 shows the results in Table 6 below.

Table 6. Partial Test Results

Variables	Parameters	Estimator	t_{test}	p_{value}
Constant	β_0	−5.1548022	−37.3875913	3.678304×10^{-32}*
	β_1	1.8613284	4.7438901	2.800389×10^{-5}*
x_1	β_2	−5.2872918	−3.9677246	0.0003015644*
	β_3	6.2415425	13.9474252	9.643968×10^{-17}*
	β_4	−2.6104044	−1.7247059	0.09249879
	β_5	−4.3104351	−4.3423558	9.711343×10^{-5}*
x_2	β_6	16.2783852	5.3213872	4.52265×10^{-6}*
	β_7	−38.6922813	−73.5323215	1.893872×10^{-43}*
	β_8	26.7111639	14.2755902	4.501067×10^{-43}*
	β_9	−0.1083037	−0.7282662	0.470801
x_3	β_{10}	6.4468373	6.7528221	4.689782×10^{-8}*
	β_{11}	−39.7216624	−110.0312135	3.04573×10^{-50}*
	β_{12}	33.3894554	63.2850696	6.288799×10^{-41}*
	β_{13}	−1.2965789	−2.0222809	0.05004422
x_4	β_{14}	12.4926949	6.5437657	9.11912×10^{-8}*
	β_{15}	−14.4965327	−27.3980679	4.36329×10^{-27}*
	β_{16}	3.2156656	1.7185920	0.09361878

Notes: (*) Significant at the real level α(0.05)

Parameters are said to be insignificant if $|t_{test}| < t_{(0.025.39)}(2.022691)$ and $p_{value} > \alpha$. The partial test results show that the four predictor variables affect the response variable significantly, indicating that there are significant parameters in the four predictor variables. The majority of the parameters in the nonparametric truncated spline model were found to be statistically significant, validating that the non-linear effects captured by the knot thresholds for HDI (x_1), Region Minimum Wage (x_2), Population (x_3), and Economic Growth (x_4) are statistically reliable and meaningful in explaining the Open Unemployment Rate.

3.6 Residual Assumptions Test

Identical Test

The identical test was applied using the Glejser test. and the results are presented in Table 7 below.

Table 7. ANOVA Table Glejser Test

	df	Sum of Squares	Mean Square	F_{test}	p_{value}
Regression	16	6.632085	0.4145053	0.6448756	0.826938
Residuals	39	25.06795	0.642769		
Total	55	31.70003			

Based on Table 7, it is known that the test statistic is F_{test} (0.826938) and p_{value} (0.6448756). The value of $F_{test} < F_{0.05(16.39)}$ (1.910658) and the value of $p_{value} > \alpha(0.05)$, so it is concluded that H_0 is accepted in the sense that the residuals do not indicate a case of heteroscedasticity. This shows that the residuals have fulfilled the identical assumption.

Independent Test

The independent test is applied using the Durbin Watson test to determine whether the residuals are independent or not using the following Durbin Watson formula.

$$d = \frac{167.7744}{69.48304} = 2.37143.$$

Based on the calculation results. the Durbin-Watson value is 2.37143. This shows that $d < (4 - d_L)$ with d_L is 1.3815 or $d > d_U$ with d_U is 1.7678, so the decision H_0 is accepted. meaning that the residuals are independent or there is no autocorrelation.

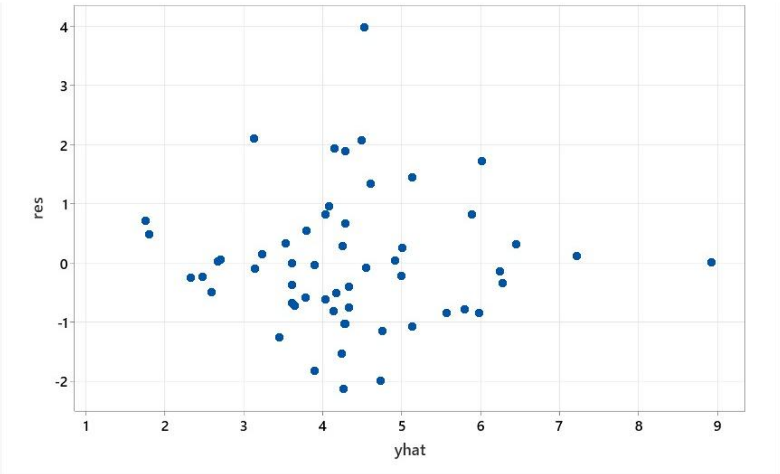

Figure 4. Scatterplot of Residuals vs Fitted Values

The scatterplot of residuals versus fitted values in Fig. 4 shows that the points are spread randomly around zero without forming a clear pattern. This indicates that the assumption of homoscedasticity (constant variance of errors) is satisfied, suggesting that the model does not suffer from major non-linearity problems.

Normal Distribution Test

The normal distribution test was carried out with the Kolmogorov-Smirnov test with the following results.

Table 8. Kolmogorov-Smirnov Test Results

D	p-value
0.11539	0.0608

Based on Table 8, shows that the p-value is 0.0608 which means $p_{value} > \alpha(0.05)$. Thus, the null hypothesis H_0 is accepted, meaning the model residuals follow a normal distribution. In addition to the statistical test, a visual inspection of the residuals was conducted using the Normal Q-Q Plot. The plot is presented in the following figure.

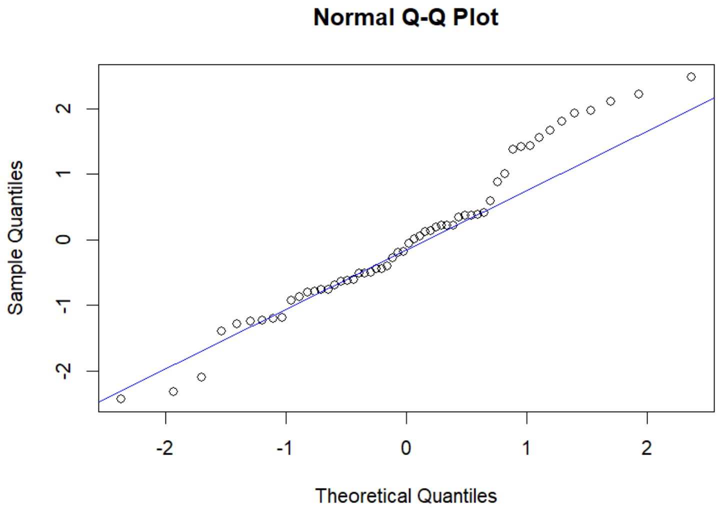

Figure 5. Normal Q-Q Plot

The visual normality test using the Normal Q-Q Plot in Fig. 5 shows that the data points are distributed around the diagonal line and follow its direction. This indicates that the regression model satisfies the normality assumption. This visual result is also consistent with the Kolmogorov-Smirnov statistical test conducted previously.

3.7 Interpretation

The following are the variables that significantly affect the Open Unemployment Rate.

Variable x_1 (Human Development Index)

$$\hat{y} = \begin{cases} -5.1548022 + 1.8613284x_1; x_1 < 68.8016 \\ 358.619333 - 3.4259634x_1;\ 68.8016 \leq x_1 < 70.183 \\ -83.363016 + 2.7894616x_1;\ 70.183 \leq x_1 < 70.304 \\ 100.158854 + 0.1790572x_1; x_1 \geq 70.304 \end{cases}.$$

Assuming the variables x_2. x_3. and x_4 are constant. the effect of HDI on Open Unemployment Rate varies by HDI range as follows:

HDI < 68.8016: A one-unit increase in HDI increases Open Unemployment Rate by 1.86%.
68.8016 ≤ HDI < 70.183: A one-unit increase in HDI decreases Open Unemployment Rate by 3.43%.
70.183 ≤ HDI < 70.304: A one-unit increase in HDI increases Open Unemployment Rate by 2.79%.
HDI ≥ 70.304: A one-unit increase in HDI increases Open Unemployment Rate by 0.18%.

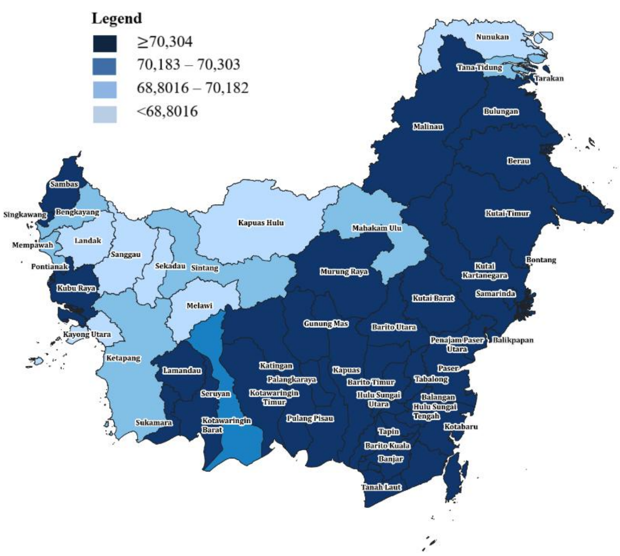

Figure 6. Distribution Map of Human Development Index

Fig. 6 illustrates the spatial distribution of the HDI across the Kalimantan region. The distribution map reveals a disparity in development levels, where the darker blue areas predominantly located in East Kalimantan and select cities in South and Central Kalimantan indicate regions with a higher HDI (HDI ≥ 70.304). Conversely, the lighter shaded areas, particularly in parts of West Kalimantan and the interior border regions, suggest relatively lower HDI scores. This highlights a gap in social and economic well-being between the coastal/industrialized zones and the rural interior.

In general, an increase in HDI tends to reduce Open Unemployment Rate. However. under certain conditions. an increase in HDI actually causes an increase in Open Unemployment Rate. This can occur because the availability of suitable jobs does not match the rise in education, so graduates take longer to find work. Regions with low HDI (HDI < 68.8016), specific mid-range HDI (70.183 ≤ HDI < 70.304) and high HDI (HDI ≥ 70.304), the Open Unemployment Rate rises. Improvements in the HDI yield a more educated workforce. However. in regions with a low-to-middle HDI or at a certain mid-range threshold, the growth of formal employment opportunities commensurate with qualifications has lagged behind the increase in the graduate supply. The HDI range of 68.8016 ≤ HDI < 70.183 represents a critical phase where the improved quality of human resources transforms into productive capital and effectively addressing the problem of educated unemployment.

Variable x_2 (Region Minimum Wage)

$$\hat{y} = \begin{cases} -5.1548022 - 4.3104251x_2; x_2 < 27.2228 \\ -448.310734 + 11.9679601x_2;\ 27.2228 \le x_2 < 30.366 \\ 726.61908 - 26.7243212x_2;\ 30.366 \le x_2 < 31.2305 \\ -107.583924 - 0.0131573x_2; x_2 \ge 31.2305 \end{cases}.$$

Assuming the variables x_1. x_3. and x_4 are constant. the effect of Region Minimum Wage on Open Unemployment Rate varies by Disctrict/City Minimum Wage range as follows:

Region Minimum Wage < 27.2228: A one-unit increase in Region Minimum Wage decreases the Open Unemployment Rate by 4.31%.
27.2228 ≤ Region Minimum Wage < 30.366: A one-unit increase in Region Minimum Wage increases the Open Unemployment Rate by 11.97%.
30.366 ≤ Region Minimum Wage < 31.2305: A one-unit increase in Region Minimum Wage decreases the Open Unemployment Rate by 26.72%.
Region Minimum Wage ≥ 31.2305: An increase in Region Minimum Wage by one unit decreases the Open Unemployment Rate by 0.013%.

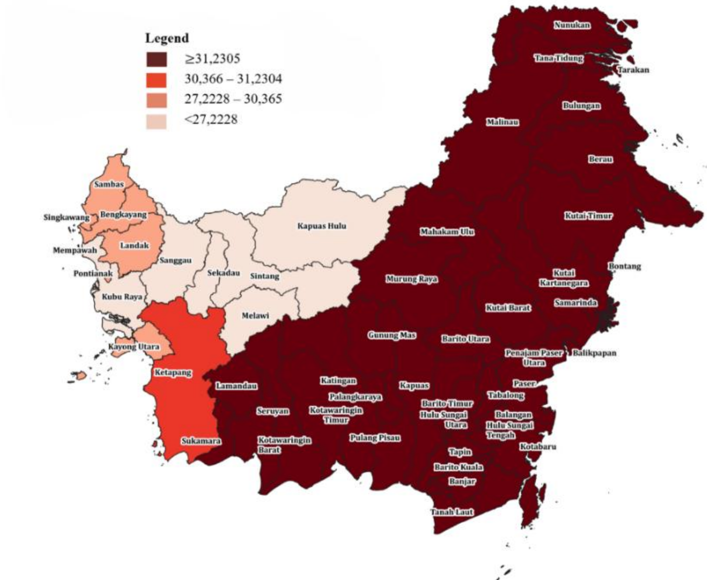

Figure 7. Distribution Map of Region Minimum Wage

Fig. 7 illustrates the regional variation in minimum wage levels throughout Kalimantan. A distinct geographical divide is evident in the data, the eastern, central, and northern provinces (shaded in dark red) exhibit significantly higher minimum wages compared to the western region (West Kalimantan), which is predominantly characterized by the lightest shading. This distribution likely correlates with the concentration of capital-intensive industries, such as mining and oil extraction, in East and Central Kalimantan. In contrast, the western region relies more heavily on traditional agriculture and plantation sectors, which generally maintain lower wage standards.

In general, an increase in Region Minimum Wage tends to reduce the Open Unemployment Rate. This can happen if the rise in Region Minimum Wage is accompanied by an increase in labor productivity. investment. and growth in the business sector that creates jobs. The main factors at play are increased consumption. investment. entrepreneurship. and labor quality. which overall help to reduce unemployment. However. in the medium Region Minimum Wage range. wage increases becoming a significant burden for for small and medium-sized businesses and labor-intensive industries. which subsequently leads to automation or layoffs and a sharp increase in the Open Unemployment Rate.

Variable x_3 (Total Population)

$$\hat{y} = \begin{cases} -5.1548022 - 0.1083037x_3; x_3 < 12.3498 \\ -84.771953 + 6.3385336x_3;\ 12.3498 \leq x_3 < 15.98 \\ 549.980212 - 33.3831288x_3;\ 15.98 \leq x_3 < 16.5699 \\ -3.279725 + 0.0063266x_3; x_3 \geq 16.5699 \end{cases}.$$

Assuming the variables x_1, x_2, and x_4 are constant. the effect of total population on Open Unemployment Rate varies by total population range as follows:

Total population < 12.3498: A one-unit increase in population decreases the Open Unemployment Rate by 0.108%.
12.3498 ≤ total population < 15.98: A one-unit increase in population increases the Open Unemployment Rate by 6.34%.
15.98 ≤ total population < 16.5699: A one-unit increase in population decreases the Open Unemployment Rate by 33.38%.
Total population≥ 16.5699: A one-unit increase in population increases the Open Unemployment Rate by 0.006%.

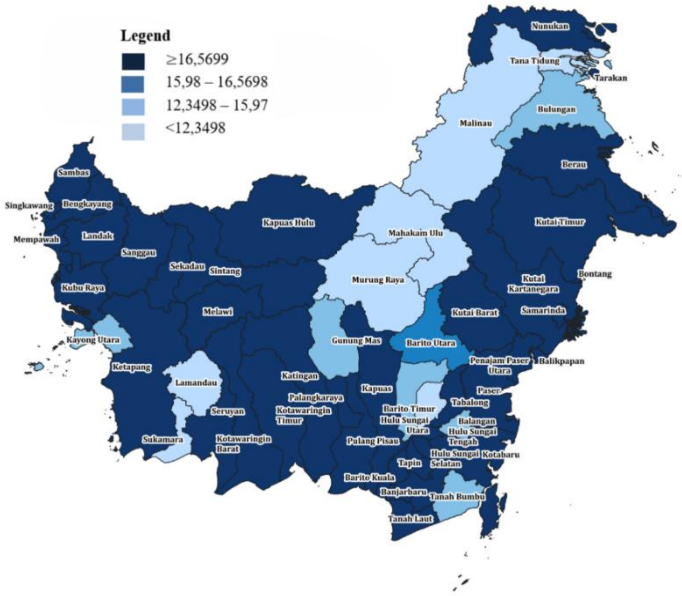

Figure 8. Distribution Map of Total Population

Fig. 8 illustrates the demographic distribution across Kalimantan. The map highlights a pattern of coastal urbanization, where the highest population densities (indicated by the darkest blue shades) are concentrated in the coastal regencies of West, South, and East Kalimantan. In contrast, the central interior of the island remains sparsely populated, as shown by the lighter hues. This spatial pattern reflects a settlement preference for coastal urban areas, which possess significantly more comprehensive infrastructure access and public facilities compared to the interior regions.

An increase in population can have different impacts on the Open Unemployment Rate. If not accompanied by an increase in employment and labor skills. an increase in population tends to increase the

Open Unemployment Rate. However, an increase in population can also reduce the Open Unemployment Rate if accompanied by good economic growth and an increase in entrepreneurship that creates more jobs. The balance between population, employment opportunities, and labor quality is the main factor in determining its impact on the Open Unemployment Rate. Supporting this finding, the effect of population on the Open Unemployment Rate is non-linear. In low and high population, the population creates economies of scale, attracts investment, and stimulates entrepreneurship, thereby reducing the Open Unemployment Rate. Conversely, in the medium and very high population, the population exceeds the current job absorption capacity, leading to a sharp increase in the Open Unemployment Rate.

Variable x_4 (Economic Growth)

$$\hat{y} = \begin{cases} -5.1548022 - 1.2965789x_4; x_4 < 3.3496 \\ -47.000332 + 11.196116x_4;\ 3.3496 \leq x_4 < 4.1768 \\ 13.548786 - 3.3004167x_4;\ 4.1768 \leq x_4 < 4.2207 \\ -0.023574 - 0.0847511x_4; x_4 \geq 4.2207 \end{cases}.$$

Assuming the variables x_1, x_2, and x_3 are constant. the effect of economic growth on Open Unemployment Rate varies by economic growth range as follows:

Economic growth< 3.3496: A one-unit increase in economic growth decreases the Open Unemployment Rate by 1.29%.
3.3496 ≤ economic growth < 4.1768: A one unit increase in economic growth increases the Open Unemployment Rate by 11.19%.
4.1768 ≤ economic growth < 4.2207: A one-unit increase in economic growth decreases the Open Unemployment Rate by 3.3%.
Economic growth ≥ 4.2207: An increase in economic growth by one unit decreases the Open Unemployment Rate by 0.08%.

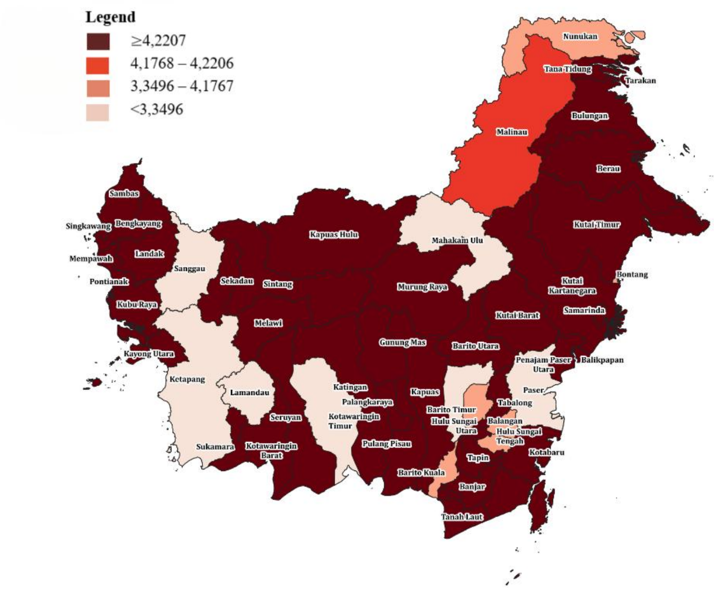

Figure 9. Distribution Map of Economic Growth

Fig. 9 illustrates the variability in economic growth rates across the districts of Kalimantan. The map indicates that high economic growth, characterized by darker red hues, is not spatially uniform but is instead clustered within specific zones, predominantly in North Kalimantan and parts of Central Kalimantan. In contrast, a significant portion of West Kalimantan and scattered districts in the south exhibit lower growth rates as evidenced by the lighter shading. This visualization underscores the heterogeneous economic performance across regions, a disparity likely driven by varying degrees of local industrial development.

In general, an increase in economic growth tends to reduce the Open Unemployment Rate because a growing economy will be able to create more jobs and improve people's welfare. Economic improvement can increase people's purchasing power, which encourages increased demand in various sectors. To meet this demand. companies will expand production and increase their workforce. However. these findings reveal that this relationship is non-linear. In the medium economic growth range. growth fails to keep pace with the growth of the labor force, consequently leading to a sharp increase in the Open Unemployment Rate.

Conversely, in the low, high, and very high economic growth ranges, an increase in economic growth causes a decrease in the Open Unemployment Rate, with the very high economic growth range achieving a sufficient scale of growth to begin massively creating employment opportunities. This finding underscores that the quality of growth is more important than the quantity, necessitating that investment policies be directed toward sectors with high labor elasticity.

4. CONCLUSION

The best truncated spline nonparametric regression model is obtained with three knot points, while the best Fourier series model is obtained with three oscillation points. Based on the GCV, R^2, and MAE values. the truncated spline model is superior to the fourier series model in modelling Open Unemployment Rate on Kalimantan Island in 2023. The truncated spline model has GCV (1.67545) and MAE (0.821277) values that are smaller than the GCV (15.96922) and MAE (0.908985) values in the Fourier series model. Meanwhile, the value of in the truncated spline model (57.399%) is greater than the Fourier series model (0.908985). Variables that have a significant effect on Open Unemployment Rate in this study are Human Development Index, Region Minimum Wage, total population, and economic growth.

Improvements in human development reflected through higher levels of education. better health outcomes. and greater adjusted per-capita expenditure play a crucial role in enhancing labor absorption and reducing socio-economic vulnerability. The influence of the minimum wage variable suggests that minimum wage policies should ensure fair pay without reducing firms' hiring capacity. population size can affect the labor force and competition for employment. which may affect unemployment levels when labor demand does not expand proportionally to population increases. Meanwhile. economic growth shapes the labor market's ability to absorb workers, implying that fluctuations in economic activity may lead to variations in unemployment levels.

Author Contributions

Aridha Pebriani Kusmiran: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Software, Validation, Visualization, Writing-Original Draft. Nur Salam: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Resources, Validation, Writing - Review and Editing. Agus Muslim: Conceptualization, Funding Acquisition, Investigation, Resources, Software, Validation, Writing - Review and Editing. All authors had approved the final version and 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 express sincere appreciation to the reviewers for their insightful comments and constructive critiques. which have significantly improved the accuracy. clarity. and overall quality of this manuscript. The authors also extend their gratitude to the editorial team for their professional management of the review process and their commitment to maintaining the high scholarly standards of the journal.

Declarations

The authors declare no conflicts of interest to report study.

Declaration of Generative AI and AI-assisted Technologies

Generative AI tools were used solely for language refinement and stylistic improvement. The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. The authors take full responsibility for the content and confirm that all final text was reviewed for linguistic accuracy by an English-language expert.

REFERENCES

[1] D. Allen and B. H. Prabowo. "SOCIAL AND ECONOMIC IMPACT OF UNEMPLOYMENT . POVERTY . AND ECONOMIC GROWTH IN INDONESIA : AN EMPIRICAL STUDY WITH A REGRESSION APPROACH." vol. 12. no. 1. pp. 69–76. 2024.

[2] A. A. Khoiriah and R. Azizah. "THE EFFECT OF MINIMUM WAGE. POPULATION. HUMAN DEVELOPMENT INDEX ON THE OPEN UNEMPLOYMENT RATE IN LAMPUNG PROVINCE." *Symmetry & Sigma: Journal of Mathematical Structures and Statistical Patterns*. vol. 1. no. 2. pp. 132–147. 2024.doi: https://doi.org/10.58989/symmerge.v1i2.26

[3] D. K. P. Wee. G. M. V. Kawung. and I. Masloman. "ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT PENGANGGURAN TERBUKA DAN KEMISKINAN DI KABUPATEN MINAHASA." *jurnal Berkala Ilmiah Efisiensi*. vol. 24. no. 2. pp. 133–145. 2024.

[4] R. M. Rusydan and R. S. Wijaya. "IMPACT OF ECONOMIC GROWTH. MINIMUM WAGE. LABOR FORCE PARTICIPATION RATE. AND POPULATION SIZE ON THE OPEN UNEMPLOYMENT." *Journal of Business Management and Economic Development*. vol. 2. no. 03. pp. 1186–1198. 2024. doi: https://doi.org/10.59653/jbmed.v2i03.911.

[5] J. K. Sinha. "IMPACT OF CAPITAL EXPENDITURE AND ECONOMIC GROWTH ON UNEMPLOYMENT IN INDIA." vol. 8. no. 10. pp. 2271–2286. 2024. doi: https://doi.org/10.26855/jhass.2024.10.008.

[6] F. H. Junianto. A. A. R. Fernandes. S. Solimun. and R. B. Hamdan. "DEVELOPMENT OF SEMIPARAMETRIC PATH ANALYSIS MODELING TRUNCATED SPLINE: DETERMINANTS OF INCREASED REGIONAL ECONOMIC GROWTH." *BAREKENG: Jurnal Ilmu Matematika dan Terapan*. vol. 18. no. 2. pp. 0949–0960. 2024. doi: https://doi.org/10.30598/barekengvol18iss2pp0949-0960.

[7] A. S. A. Sadikin. A. A. Reswara. M. F. F. Mardianto. and A. Kurniawan. "COMPARISION OF RICE PRICE PREDICTION RESULTS IN EAST JAVA USING FOURIER SERIES ESTIMATOR AND GAUSSIAN KERNEL ESTIMATOR SIMULTANEOUSLY." *BAREKENG: Jurnal Ilmu Matematika dan Terapan*. vol. 18. no. 3. pp. 1963–1974. 2024. doi: https://doi.org/10.30598/barekengvol18iss3pp1963-1974.

[8] M. E. Mocodompis. D. T. Salaki. and D. Hatidja. "PENERAPAN REGRESI NONPARAMETRIK SPLINE DAN DERET FOURIER UNTUK MEMODELKAN PRODUK DOMESTIK REGIONAL BRUTO DI INDONESIA." *Jurnal Ilmiah Sains*. vol. 23. no. 2. pp. 168–174. 2023. doi: https://doi.org/10.35799/jis.v23i2.48868.

[9] A. T. R. Dani and N. Y. Adrianingsih. "PEMODELAN REGRESI NONPARAMETRIK DENGAN ESTIMATOR SPLINE TRUNCATED VS DERET FOURIER." *Jambura Journal of Mathematics*. vol. 3. no. 1. pp. 26–36. 2021. doi: https://doi.org/10.34312/jjom.v3i1.7713.

[10] A. Suhada. Syafriandi. D. Vionanda. and F. Fitri. "MODELING OPEN UNEMPLOYMENT RATE IN WEST SUMATERA PROVINCE USING TRUNCATED SPLINE REGRESSION." *UNP Journal of Statistics and Data Science*. vol. 1. no. 1. pp. 39–44. 2023. doi: https://doi.org/10.24036/ujsds/vol1-iss1/3.

[11] R. Rahmania. S. Sifriyani. and A. T. R. Dani. "MODELING OPEN UNEMPLOYMENT RATE IN KALIMANTAN ISLAND USING NONPARAMETRIC REGRESSION WITH FOURIER SERIES ESTIMATOR." *BAREKENG: Jurnal Ilmu Matematika dan Terapan*. vol. 18. no. 1. pp. 0245–0254. 2024. doi: https://doi.org/10.30598/barekengvol18iss1pp0245-0254.

[12] A. F. Al Barra and D. R. S. Saputro. "KNOT OPTIMIZATION FOR BI-RESPONSE SPLINE NONPARAMETRIC REGRESSION WITH GENERALIZED CROSS-VALIDATION (GCV)." *Barekeng*. vol. 19. no. 1. pp. 271–280. 2025. doi: https://doi.org/10.30598/barekengvol19iss1pp271-280.

[13] M. Fatekurohman. S. N. Khasanah. and Y. S. Dewi. "NONPARAMETRIC REGRESSION MODELING USING THE SPLINE APPROACH TO STUNTING CASES IN INDONESIA." *Barekeng*. vol. 19. no. 2. pp. 697–708. 2025. doi: https://doi.org/10.30598/barekengvol19iss2pp697-708.

[14] M. A. D. Octavanny. I. N. Budiantara. H. Kuswanto. and D. P. Rahmawati. "MODELING OF CHILDREN EVER BORN IN INDONESIA USING FOURIER SERIES NONPARAMETRIC REGRESSION." *Journal of Physics: Conference Series*. vol. 1752. no. 1. 2021. doi: https://doi.org/10.1088/1742-6596/1752/1/012019.

[15] L. Ni'matuzzahroh and A. T. R. Dani. "NONPARAMETRIC REGRESSION MODELING WITH MULTIVARIABLE FOURIER SERIES ESTIMATOR ON AVERAGE LENGTH OF SCHOOLING IN CENTRAL JAVA IN 2023." *Inferensi*. vol. 7. no. 2. p. 73. 2024. doi: https://doi.org/10.12962/j27213862.v7i2.20219.

[16] I. W. Sudiarsa. N. P. A. M. Mariati. N. M. S. Sanjiwani. and D. P. Pramesuari. "ESTIMATOR AND APPLIED MIXED KERNEL AND FOURIER SERIES MODELLING IN NONPARAMETRIC REGRESSION." *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*. vol. 8. no. 2. pp. 622–633. 2024. doi: https://doi.org/10.22437/jiituj.v8i2.36246.

[17] M. I. Salim. Adnan Sauddin. and M. Ichsan Nawawi. "MODEL REGRESI NONPARAMETRIK DERET FOURIER PADA KASUS TINGKAT PENGANGGURAN TERBUKA DI SULAWESI SELATAN." *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)*. vol. 10. no. 2. pp. 48–56. 2022. doi: https://doi.org/10.24252/msa.v10i2.30993.

[18] M. Maharani and D. R. S. Saputro. "GENERALIZED CROSS VALIDATION (GCV) IN SMOOTHING SPLINE NONPARAMETRIC REGRESSION MODELS." *IOP Conference Series: Earth and Environmental Science*. vol. 1808. no. 1. 2021. doi: https://doi.org/10.1088/1742-6596/1808/1/012053.

[19] N. C. Iswanto. Soerahman. A. Naim. and R. Saputra. "THE INFLUENCE OF OCCUPATIONAL SAFETY AND HEALTH (K3) ON EMPLOYEE WORK PRODUCTIVITY WITH MULTIPLE LINEAR REGRESSION METHODS AT PT. UDM." *International Journal of Science and Society*. vol. 5. no. 5. pp. 39–52. 2023. doi: https://doi.org/10.54783/ijsoc.v5i5.866.

[20] D. Chicco. M. J. Warrens. and G. Jurman. "THE COEFFICIENT OF DETERMINATION R-SQUARED IS MORE INFORMATIVE THAN SMAPE. MAE. MAPE. MSE AND RMSE IN REGRESSION ANALYSIS EVALUATION." *PeerJ Computer Science*. vol. 7. pp. 1–24. 2021. doi: https://doi.org/10.7717/peerj-cs.623.