

## Analysis of Open Unemployment Rates in Indonesia Based on GRDP and the Percentage of Poor Population Using Nonparametric B-Spline Regression

Nadila Riska Afifah<sup>1\*</sup>, Admi Salma<sup>2</sup>

<sup>1,2</sup>Department of Statistics, Padang State University, Indonesia

Email: <sup>1</sup>nadilariskaafifah2002@gmail.com, <sup>2</sup>admisalma1@fnipa.unp.ac.id

\*Corresponding author

### Abstract

Open Unemployment Rate (OUR) is a crucial indicator of the condition of the job market and the economy in Indonesia. This research is to modeling and analyzes the impact of GRDP and the percentage of poor population to the OUR in Indonesia using nonparametric B-Spline regression. The study applied B-Spline model due to the model's property of handling non-linear associations without imposing any distributional assumptions. The research a used secondary data sourced from BPS Indonesia in 2024, which included 38 provinces in Indonesia. The analysis outcomes show that optimal model is achieved when the order was set at 2 for both GRDP and the percentage of poor population with one knot set at GRDP (1.055) and two knots set at the percentage of poor population (6.813333 and 11.583333) which gave a minimum GCV produced at 1.577369. The model's coefficient of determination value of 0.7454 indicates that the model can explain 74.54% of the variation in the OUR is explained by GRDP and the percentage of poor population, with the remaining 25.46% is impacted by factors beyond the model.

**Keywords:** GRDP, Nonparametric B-Spline Regression, Open Unemployment Rate, The Percentage of Poor Population.

## 1. INTRODUCTION

Unemployment is on of the crucial problems that continues to impede Indonesia's economic growth. In addition, the issue remains a significant factor in the social stability of the state. According to Muhgni et al., (2020), the higher the unemployment rates, the greater the chances of criminal behavior. Unemployment can be described as individuals in the labor force but not yet in the job although actively searching for it [13]. The economic indicator used to measure the unemployment rate is known as the Open Unemployment Rate (OUR), it refers to the percentage of individuals but who are currently unemployed but available for work out of the total labor force [4]. The unemployment rate represents a socio-economic issue in both developed and developing countries, as the OUR determines a country's standard of living. The level of the OUR in Indonesia as of August 2024 was recorded at 4.91% [5], which remains much lower compared to the target set by the government, set at 4% [10]. The unemployment level not only shows the correlation between



the workforce availability and job opportunities in the labor market but also the level of poverty, income inequality, and the stability of the economy.

In order to minimize Indonesia's unemployment rate, it is important to examine any factors that determine the OUR, including the Gross Regional Domestic Product (GRDP) and the percentage of poor population. GRDP is a macroeconomic indicator that is closely related to the OUR because it represents the total value added generated by different economic sectors within an area [15]. Meanwhile, the percentage of the poor population demonstrates socio-economic disparities by reflecting the proportion of people whose average monthly per capita expenditure is less than the poverty levels.

According to Azizah et al., (2023), the OUR, GRDP, and the percentage of poor population are interrelated, with a higher GRDP tending to reduce the percentage of poor population due to increased employment opportunities and income, while a higher unemployment rate may lead to greater poverty because it reduces income and increases the risk of impoverishment. To examine the relationship between two or more variables, regression analysis can be employed. Nonparametric regression is a sort of regression analysis that does not require explicit assumptions, it provides greater flexibility because its regression curve can adjust to data that does not follow a specific pattern [2]. The B-Spline serves as the basis function for one of the nonparametric regression approaches, namely the spline estimator.

The B-Spline model overcomes the restrictions of truncated basis functions by allowing for greater orders, a large number of knot points, or tighter intervals [12]. Several previous studies have shown that the B-Spline model is capable of representing nonlinear patterns without requiring specific functional form assumptions and that it can effectively handle data patterns that fluctuate sharply—either increasing or decreasing—that parametric methods cannot address, as well as data with smooth characteristics [19]. Nasir et al., (2024) discovered that the B-Spline model provides a more accurate and flexible representation than conventional regression, providing deeper insights into the interplay of factors affecting poverty.

Several prior research have examined the correlation between the OUR and other social indicators. The research by Amaliah et al., (2025), used the B-Spline to analyze the OUR based on social indicators in Bangka Belitung Province, found that the Gross Enrollment Rate for Senior High School and the Poverty Depth Index significantly influenced the OUR. In addition, Erlangga (2019) examined the correlation between the OUR in Indonesia in 2017, indicating that GRDP and the Provincial Minimum Wage (PMW), significantly that both GRDP and PMW had a significant effect on the OUR using nonparametric B-Spline regression. Another study conducted by Wijaya (2018) in Aceh Province using truncated spline nonparametric regression showed that the Labor Force Participation Rate (LFPR) and the percentage of the poor population significantly affected the OUR.

The OUR, GRDP, and the percentage of the poor population have a linear correlation, especially when considering the region in Indonesia that has economically recovered in an unequal manner. accordingly, the purpose of this research is to model and examine the effect of GRDP and the percentage of the poor population on the OUR in Indonesia using nonparametric B-Spline regression. The result of this research shall serve as important for government concerning the formulation of strategies in promoting economic developments that address unemployment issues effectively.

## **2. MATERIAL AND METHODS**

### **2.1 Types and Sources of Data**

This research utilizes secondary data type, sourced from publications by Statistics Indonesia (BPS) in 2024 and 2025. The Open Unemployment Rate (OUR) is the dependent variable in this research, with the Gross Regional Domestic Product (GRDP) and the percentage of the poor population as independent variables.

## 2.2 Nonparametric Regression

Nonparametric regression is a method in regression analysis for modeling data patterns that do not follow a specific pattern [9]. Nonparametric regression is an alternative to parametric regression that does not impose rigid assumptions. The function in nonparametric regression is considered to be smooth within a certain function. The nonparametric regression equation is the following [11]:

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, 3, \dots, n \quad (2.1)$$

## 2.3 B-Spline Regression

The B-Spline model has criteria that need to be considered, such as determining the order, the total of knots, and the knots placements. The optimal B-Spline model is determined by the minimum Generalized Cross Validation (GCV) value [19]. Using an order  $m$  B-Spline function with  $k$  knots, the nonparametric regression model has the following equation [2]:

$$y_i = \sum_{h=1}^p \sum_{j=1}^{m+k} \beta_{hj} B_{(j-m),m}(x_{hi}) + \varepsilon_i \quad i = 1, 2, 3, \dots, n \quad (2.2)$$

To create an  $m$  order B-Spline function with knot points, where  $\alpha < u_1 < u_2 < \dots < u_k < b$ , the first step is to define additional knots totaling  $2m$ , given by:

$$u_{-(m-1)} < \dots < u_{-1} < u_0 < \dots < u_{(k+m)} \quad (2.3)$$

where  $u_{-(m-1)} = \dots = u_0 = a$  and  $u_{k+1} = \dots = u_{k+m} = b$ . Usually,  $a$  is taken as the minimum value of  $x$ , and  $b$  as the maximum value of  $x$  [6].

According to Ariesta et al., (2021), a recursive B-Spline function with order  $m$  at the  $k$  knot point can be defined as  $j = -(m-1), \dots, k$ , and the B-Spline basis function is expressed as follows:

$$B_{j,1}(x) = f(x) = \begin{cases} 1, & \text{if } u \leq x \leq u_{j+1} \\ 0, & \text{elsewhere} \end{cases} \quad (2.4)$$

The degree or order of a B-Spline is denoted by  $m$ . The types of B-Spline basis functions categorized based on the order  $m$  are divided into three types [14]:

a. For order  $m = 2$ , the B-Spline basis function is linear and defined as follows:

$$B_{j,2}(x) = \frac{x - u_j}{u_{j+1} - u_j} B_{j,1}(x) + \left(1 - \frac{u_{j+2} - x}{u_{j+2} - u_{j+1}}\right) B_{j+1,1}(x) \quad \text{in } j = -1, \dots, k \quad (2.5)$$

b. For order  $m = 3$ , the B-Spline basis function is quadratic and defined as follows:

$$B_{j,3}(x) = \frac{x - u_j}{u_{j+2} - u_j} B_{j,2}(x) + \left(1 - \frac{u_{j+3} - x}{u_{j+3} - u_{j+1}}\right) B_{j+1,2}(x) \quad \text{in } j = -2, \dots, k \quad (2.6)$$

c. For order  $m = 4$ , the B-Spline basis function is cubic and defined as follows:

$$B_{j,4}(x) = \frac{x - u_j}{u_{j+3} - u_j} B_{j,3}(x) + \left(1 - \frac{u_{j+4} - x}{u_{j+4} - u_{j+1}}\right) B_{j+1,3}(x) \quad \text{in } j = -3, \dots, k \quad (2.7)$$

## 2.4 Parameter Estimation of B-Spline Model

The B-Spline model in equation (2.2) can be written as follows :

$$y_i = \beta_{11}B_{(1-m),m}(x_{1i}) + \beta_{12}B_{(2-m),m}(x_{1i}) + \cdots + \beta_{1(m+k)}B_{k,m}(x_{1i}) + \cdots \quad (2.8)$$

$$+ \beta_{p1}B_{(1-m),m}(x_{pi}) + \beta_{p2}B_{(2-m),m}(x_{pi}) + \cdots + \beta_{p(m+k)}B_{k,m}(x_{pi}) + \varepsilon_i$$

The B-Spline model in equation (2.8) can be expressed in matrix form [14]:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} B_{(1-m),m}(x_{11}) & \cdots & B_{k,m}(x_{11}) & \cdots & B_{(1-m),m}(x_{p1}) & \cdots & B_{k,m}(x_{p1}) \\ B_{(1-m),m}(x_{12}) & \cdots & B_{k,m}(x_{12}) & \cdots & B_{(1-m),m}(x_{p2}) & \cdots & B_{k,m}(x_{p2}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_{(1-m),m}(x_{1n}) & \cdots & B_{k,m}(x_{1n}) & \cdots & B_{(1-m),m}(x_{pn}) & \cdots & B_{k,m}(x_{pn}) \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{(m+k)1} \\ \vdots \\ \beta_{1p} \\ \beta_{2p} \\ \vdots \\ \beta_{p(m+k)} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.9)$$

can be written as:

$$\mathbf{y} = \mathbf{B}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.10)$$

The least squares spline method is used to derive the parameter estimate  $B_\lambda$ . The estimator  $\hat{\beta}_\lambda$  is derived through the minimization of the Residual Sum of Squares (RSS), resulting in the following expression:

$$\hat{\beta} = (B^T B)^{-1} B^T \mathbf{y} \quad (2.11)$$

Thus, the estimated model for B-Spline function in nonparametric regression is expressed as follows [2]:

$$\begin{aligned} \hat{y} &= B\hat{\beta} \\ &= B((B^T B)^{-1} B^T \mathbf{y}) \\ &= B(B^T B)^{-1} B^T \mathbf{y} \\ &= S\mathbf{y} \end{aligned}$$

where the matrix  $S\mathbf{y} = B(B^T B)^{-1} B^T \mathbf{y}$  is symmetric (Eubank, 1999, as cited in [2]). The B-Spline model estimation can be written as follows [12]:

$$\hat{y}_i = \sum_{h=1}^p \sum_{j=1}^{m+k} \hat{\beta}_{hj} B_{(j-m),m}(x_{hi}) \quad (2.13)$$

The equation can be defined as:

$$(2.12)$$

$$\hat{y}_i = \hat{\beta}_{11}B_{(1-m),m}(x_{1i}) + \hat{\beta}_{12}B_{(2-m),m}(x_{1i}) + \cdots + \hat{\beta}_{1(m+k)}B_{k,m}(x_{1i}) + \cdots + \hat{\beta}_{p1}B_{(1-m),m}(x_{pi}) + \hat{\beta}_{p2}B_{(2-m),m}(x_{pi}) + \cdots + \hat{\beta}_{p(m+k)}B_{k,m}(x_{pi})$$

### 2.5 Selection of Optimal Knots

The optimal B-Spline model is determined based on knot points corresponding to the lowest GCV value. A lower GCV value achieved during the computation process shows that B-Spline model with the chosen knot points configuration produces the best model performance. The formula for computing the GCV value is expressed as follows [17]:

$$GCV(\lambda) = \frac{MSE(\lambda)}{(n^{-1}\text{trace}[I - S_\lambda])^2} \quad (2.15)$$

where  $n$  represents the number of observations,  $I$  is an  $n \times n$  matrix, and  $S_\lambda = B_\lambda(B_\lambda^T B_\lambda)^{-1}B_\lambda^T$ .

### 2.6 Coefficient of Determination

The coefficient of determination ( $R^2$ ) measures how well the independent variables explains the variation in the dependent variable in the model. The higher the  $R^2$  value, or the closer it is to 1, the better the independent variables are at explaining the variability of the dependent variable [17]. The  $R^2$  helps evaluate the model's effectiveness in representing the correlation among the variables and indicates how accurately the model can predict outcomes. The coefficient of determination can be calculated using the following formula (Sembiring, 2003) as cited in [2]:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (2.16)$$

## 3. RESULTS AND DISCUSSION

### 3.1 Data Exploration

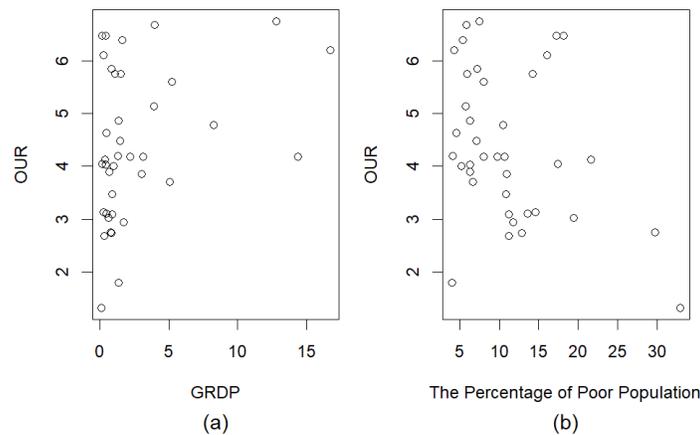
This study employs nonparametric B-Spline regression to analyze the OUR (Y) in Indonesia based on GRDP ( $X_1$ ) and the percentage of the poor population ( $X_2$ ), consisting of 38 observations. Table 3.1 shows the descriptive statistics for each variable.

**Table 3. 1.** Descriptive statistics of variables

Variable	Mean	Min	Max
Y	4.38	1.32	6.75
X1	2.63	0.12	16.71
X2	11.15	4	32.97

Based on Table 3.1, the OUR across Indonesian provinces in 2024 averages 4.38%, with the lowest rate recorded in Papua Pegunungan Province at 1.32% and the highest in West Java Province at 6.75%. The GRDP in Indonesia during the same year averages 2.63%, ranging from 0.12% in Papua Pegunungan Province to 16.71% in DKI Jakarta. Meanwhile, the percentage of the poor population across 38 provinces averages 11.15%, with the lowest in Bali Province at 4% and the highest in Papua Pegunungan Province at 32.97%.

The pattern of the correlation between the OUR and both GRDP and the percentage of the poor population in Indonesia in 2024 is illustrated through a scatter plot, as presented in Figure 3.1.



**Figure 3. 1.** (a) Plot of the correlation between the OUR and GRDP, (b) Plot of the correlation between the OUR and the percentage of the poor population.

Figure 3.1 show that the plot does not exhibit a specific pattern, indicating that applying a parametric regression analysis would face challenges in meeting the linearity assumption. Therefore, the data are analyzed using a nonparametric regression approach, which does not require assumption testing in its application. In this study, the nonparametric approach employed is the B-spline method.

### 3.2 Optimal Knot Points

Nonparametric B-Spline regression analysis involves three important aspects in model construction: the selection of the spline order, the number of knots, and the placements of those knots. In this study, spline order ranging from 2 to 4 were considered, following de Boor (2001), who noted that low-order splines are computationally stable and practically effective for modeling. This choice is also consistent with Ruppert et al., (2003), who recommended using low-degree splines—typically linear to cubic (1–4)—because model flexibility is primarily governed by knot placement rather than spline order. To generate variations in spline order and number of knots, the number of knots points used for each combination of spline orders across the independent variable consists of one knot, two knots, and a combination of one and two knots. This strategy allows the model to capture different levels of smoothness across predictors, depending on the complexity of their relationships with the response variable. The selection of optimal knot points is carried out to obtain the best B-spline model estimation. An optimal B-spline model is achieved by determining the number and placement of knots across several spline order, the optimal knots points will be selected based on the smallest GCV value.

The knot points and GCV values for each combination of order and number of knots are presented in Table 3.2 through Table 3.6. All numerical calculations were obtained using R Studio, which applied the formulations in equations (2.3), (2.5), (2.6), (2.7) and (2.15). this process produced the optimal knot points configuration along with the corresponding GCV values in the analysis.

**Table 3. 2.** GCV and knot points for one knot in X1 and one knot in X2

Combination of Order		Knot Points		GCV
		X1	X2	
2	2	1.055	10.13	2.170908
2	3	1.055	10.13	1.874362
<b>2</b>	<b>4</b>	<b>1.055</b>	<b>10.13</b>	<b>1.677202</b>
3	2	1.055	10.13	2.312293

## JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Nadila Riska Afifah, Admi Salma

Combination of Order		Knot Points		GCV
		X1	X2	
3	3	1.055	10.13	1.949063
3	4	1.055	10.13	1.792816
4	2	1.055	10.13	2.480075
4	3	1.055	10.13	2.076878
4	4	1.055	10.13	1.926409

Table 3.2, which presents the results for one knot in X1 and one knot in X2, shows that the smallest GCV value produces the optimal B-Spline model, where when X1 is order 2 and X2 is order 4, the lowest GCV value of 1.677202 in Table 3.2 is achieved with the knot point of X1 = 1.055 and X2 = 10.13.

**Table 3. 3.** GCV and knot points for one knot in X1 and two knots in X2.

Combination of Order		Knot Points			GCV
		X1	X2		
<b>2</b>	<b>2</b>	<b>1.055</b>	<b>6.813333</b>	<b>11.58333</b>	<b>1.577369</b>
2	3	1.055	6.813333	11.58333	1.720722
2	4	1.055	6.813333	11.58333	1.738438
3	2	1.055	6.813333	11.58333	1.67065
3	3	1.055	6.813333	11.58333	1.823322
3	4	1.055	6.813333	11.58333	1.865669
4	2	1.055	6.813333	11.58333	1.782745
4	3	1.055	6.813333	11.58333	1.946855
4	4	1.055	6.813333	11.58333	2.006569

Based on Table 3.3, for one knot in X1 and two knots in X2, it is found that when X1 is of order 2 and X2 is of order 2, with the knot point of X1 located at 1.055 and the knot points of X2 located at 6.813333 and 11.583333, the lowest GCV value obtained is 1.577369, indicating that the optimal knot combination occurs at these points.

**Table 3. 4.** GCV and knot points for two knots in X1 and one knots in X2

Combination of Order		Knot Points			GCV
		X1	X2		
2	2	0.71	1.57	10.13	2.311798
2	3	0.71	1.57	10.13	1.892361
<b>2</b>	<b>4</b>	<b>0.71</b>	<b>1.57</b>	<b>10.13</b>	<b>1.743715</b>
3	2	0.71	1.57	10.13	2.479763
3	3	0.71	1.57	10.13	2.06574
3	4	0.71	1.57	10.13	1.907029
4	2	0.71	1.57	10.13	2.651183
4	3	0.71	1.57	10.13	2.238756
4	4	0.71	1.57	10.13	2.075682

Based on Table 3.4, for two knots in X1 and one knot in X2, the lowest GCV value is 1.743715, indicating that the optimal knot combination occurs when X1 is of order 2 and X2 is of order 4, with the knot points of X1 at 0.71 and 1.57 and the knot point of X2 at 10.13.

**Table 3. 5.** GCV and knot points for two knots in X1 and two knots in X2

Combination of Order		Knot Points				GCV
		X1		X2		
<b>2</b>	<b>2</b>	<b>0.71</b>	<b>1.57</b>	<b>6.813333</b>	<b>11.58333</b>	<b>1.612069</b>
2	3	0.71	1.57	6.813333	11.58333	1.749042
2	4	0.71	1.57	6.813333	11.58333	1.795134
3	2	0.71	1.57	6.813333	11.58333	1.76194
3	3	0.71	1.57	6.813333	11.58333	1.920557
3	4	0.71	1.57	6.813333	11.58333	1.97614
4	2	0.71	1.57	6.813333	11.58333	1.917029
4	3	0.71	1.57	6.813333	11.58333	2.099123
4	4	0.71	1.57	6.813333	11.58333	2.16725

Based on Table 3.5, for two knots in X1 and two knots in X2, the lowest GCV value obtained is 1.612069. This indicates that the optimal knot combination occurs when X1 is of order 2 and X2 is of order 2, with the knot points of X1 at 0.71 and 1.57 and the knot points of X2 at 6.813333 and 11.583333.

### 3.3 Optimal B-Spline Model

The optimal B-Spline model is conducted based on pairs of order and knot points that have minimum GCV value. Table 3.6 presents the minimum GCV values for each combination of the number of knot points for each variable.

**Table 3. 6.** Minimum GCV values for each combination of order and knot points

Combination of Order		Knot Points				GCV
		X1		X2		
2	4	1.055		10.13		1.677202
<b>2</b>	<b>2</b>	<b>1.055</b>		<b>6.813333</b>	<b>11.58333</b>	<b>1.577369</b>
2	4	0.71	1.57	10.13		1.743715
2	2	0.71	1.57	6.813333	11.58333	1.612069

According to Table 3.6, the minimum GCV of 1.577369 is obtained when both X1 and X2 are of order 2, with one knot for X1 at point 1.055 and two knots for X2 at point 6.813333 and 11.583333. The estimated B-Spline model based on equation (2.8), representing the optimal fit based on the minimum GCV, is given by the following equation:

$$\hat{y}_i = \beta_{11}B_{-1,2}(x_{1i}) + \beta_{12}B_{0,2}(x_{1i}) + \beta_{13}B_{1,2}(x_{1i}) + \beta_{21}B_{-1,2}(x_{2i}) + \beta_{22}B_{0,2}(x_{2i}) + \beta_{23}B_{1,2}(x_{2i}) + \beta_{24}B_{2,2}(x_{2i}) \quad (3.1)$$

### 3.4 B-Spline Model Parameter Estimation

Based on the estimated B-Spline model in equation (3.1), the parameter estimates of the model derived from equation (2.12) were obtained using R Studio as follows:

**Table 3. 7.** Parameter estimates of the best B-Spline model

Variable	Parameter	Parameter Estimates
X1	$\beta_{11}$	-0.1928
	$\beta_{12}$	2.0639
	$\beta_{13}$	2.0536
X2	$\beta_{21}$	2.7257

$\beta_{22}$	-0.7169
$\beta_{23}$	4.1793
$\beta_{24}$	-2.0534

The following is an expression for the best B-Spline model's estimated results:

$$\hat{y}_i = -0.1928B_{-1,2}(x_{1i}) + 2.0639B_{0,2}(x_{1i}) + 2.0536B_{1,2}(x_{1i}) + 2.7257B_{-1,2}(x_{2i}) \\ + (-0.7169)B_{0,2}(x_{2i}) + 4.1793B_{1,2}(x_{2i}) + (-2.0534)B_{2,2}(x_{2i})$$

Based on the obtained model, the  $R^2$  value is 0.7454, shows that 74.54% of the variation in Indonesia's Open Unemployment Rate across provinces in 2024 can be explained by GRDP and the percentage of the poor population, while the remaining 25.46% is impacted by other variables beyond the model. This indicates that the B-Spline model used is quite robust and representative in describing the correlation between GRDP and the percentage of poor population to the OUR.

### 3.5 Evaluation of Predicted Results Against Actual Data

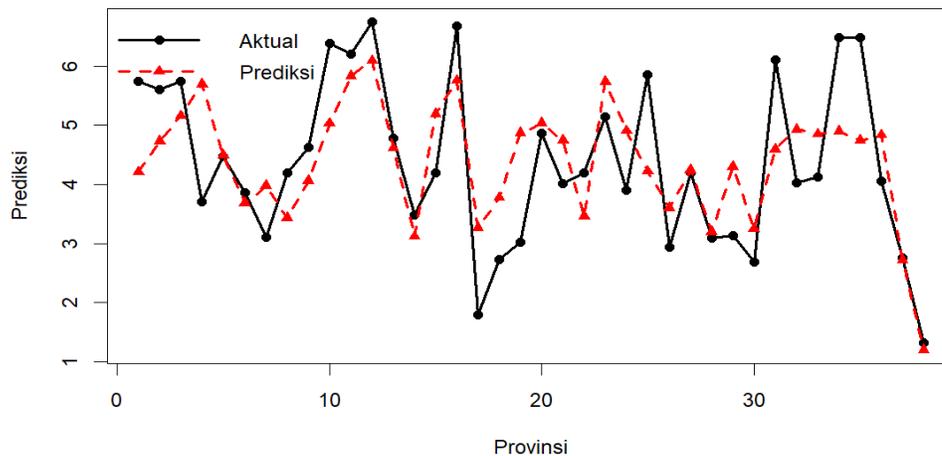
The comparison of predicted outcomes with actual data is conducted to evaluate how accurately the nonparametric B-Spline model predicting the Open Unemployment Rate (OUR) based on GRDP and the percentage of the poor population. This evaluation allows for an understanding of how well the model captures the true correlation patterns among the variables in the observed data.

**Table 3. 8.** Actual and predicted OUR values in Indonesia

Provinsi	Actual Data	Predicted Data	Residual
Aceh	5.75	4.20	1.55
Sumatera Utara	5.6	4.73	0.87
Sumatera Barat	5.75	5.16	0.59
Riau	3.7	5.70	-2.00
Jambi	4.48	4.49	-0.01
Sumatera Selatan	3.86	3.68	0.18
Bengkulu	3.11	3.98	-0.87
Lampung	4.19	3.43	0.76
Kep. Bangka Belitung	4.63	4.06	0.57
Kep. Riau	6.39	5.03	1.36
DKI Jakarta	6.21	5.84	0.37
Jawa Barat	6.75	6.09	0.66
Jawa Tengah	4.78	4.61	0.17
DI Yogyakarta	3.48	3.12	0.36
Jawa Timur	4.19	5.19	-1.00
Banten	6.68	5.75	0.93
Bali	1.79	3.26	-1.47
Nusa Tenggara Barat	2.73	3.77	-1.04
Nusa Tenggara Timur	3.02	4.87	-1.85
Kalimantan Barat	4.86	5.04	-0.18
Kalimantan Tengah	4.01	4.75	-0.74
Kalimantan Selatan	4.2	3.46	0.74

Kalimantan Timur	5.14	5.73	-0.59
Kalimantan Utara	3.9	4.91	-1.01
Sulawesi Utara	5.85	4.22	1.63
Sulawesi Tengah	2.94	3.60	-0.66
Sulawesi Selatan	4.19	4.24	-0.05
Sulawesi Tenggara	3.09	3.20	-0.11
Gorontalo	3.13	4.30	-1.17
Sulawesi Barat	2.68	3.24	-0.56
Maluku	6.11	4.59	1.52
Maluku Utara	4.03	4.93	-0.90
Papua Barat	4.13	4.85	-0.72
Papua Barat Daya	6.48	4.90	1.58
Papua	6.48	4.74	1.74
Papua Selatan	4.05	4.84	-0.79
Papua Tengah	2.75	2.72	0.03
Papua Pegunungan	1.32	1.19	0.13

The comparison between the predicted estimates OUR and the actual OUR data is illustrated in the following curve.



**Figure 3. 2.** Plot of predicted and actual data

According to Figure 3.2, the black line represents the actual OUR data, while the red dashed line represents the predicted OUR data. In general, the predicted curve follows the overall pattern of the observed data, although discrepancies between the observed and predicted values are present at several points. These deviations indicate that the predicted values do not exactly coincide with the observed data, particularly at certain levels of OUR. Furthermore, Table 3.8 shows that the residuals do not exhibit a systematic pattern and are randomly distributed around zero, suggesting that the model adequately captures the fluctuating behavior of the OUR data. Therefore, the B-spline regression model, supported by an appropriate selection of spline order and knot placement, provides a satisfactory fit to the data.

#### 4. CONCLUSION AND RECOMMENDATION

Based on the analysis of the correlation between GRDP and the percentage of the poor population on the OUR of provinces in Indonesia in 2024, it was concluded that the optimal number and placement of knots based on the minimum GCV value of 1,577369, for one knot for GRDP located at 1,055 and two knots for the percentage of the poor population located at 6,813333 and 11,583333, both with an order of 2. The resulting model equation is as follows:

$$\hat{y}_i = -0.1928B_{-1,2}(x_1) + 2.0639B_{0,2}(x_1) + 2.0536B_{1,2}(x_1) + 2.7257B_{-1,2}(x_2) + (-0.7169)B_{0,2}(x_2) + 4.1793B_{1,2}(x_2) + (-2.0534)B_{2,2}(x_2)$$

Based on the B-Spline modeling results, considering the number and placement of knots for each variable and the model order, the coefficient of determination ( $R^2$ ) is 74.54%, indicating that GRDP and the percentage of the poor population significantly influence the OUR of provinces in Indonesia in 2024. This method serves as a reference and consideration for identifying factors that affect OUR in Indonesia. For future research, it is recommended to include other social indicators beyond GRDP and the percentage of the poor population to better identify factors that may have a greater impact on OUR in Indonesia.

#### REFERENCES

- [1] Amaliah, K. L. N., Amalya, D., Selviana, N., Afrilita, M., Putri, D. S. A., & Sulistiana, I., 2025. Analisis Tingkat Pengangguran Terbuka Berdasarkan Indikator Sosial di Provinsi Kepulauan Bangka Belitung Menggunakan Pendekatan Regresi B-Spline. *Jurnal Gaussian*. Vol. 14, No. 2, 401–410. <https://doi.org/10.14710/j.gauss.14.2.401-410>
- [2] Ariesta, D., Gusriani, N., & Parmikanti, K., 2021. Estimasi Parameter Model Regresi Nonparametrik B-Spline Pada Angka Kematian Maternal. *Jurnal Matematika UNAND*. Vol 10, No. 3, 342–354.
- [3] Azizah, L. N., Pasaribu, J. R. S., Hutagulung, I., Purba, A. A., & Sinaga, S. A., 2023. Analisis Pengaruh PDRB dan Pengangguran Terhadap Kemiskinan di Indonesia Tahun 2018-2022. *JoSES: Journal of Sharia Economic Scholar*. Vol. 2 No. 1, 25–32.
- [4] Badan Pusat Statistik. 2024. *Keadaan Angkatan Kerja di Indonesia Agustus 2024*. Badan Pusat Statistik.
- [5] Badan Pusat Statistik. 2025. *Tingkat Pengangguran Terbuka Menurut Provinsi (Persen), 2024*. Badan Pusat Statistik.
- [6] Budiantara, I. N., Suryadi, F., Otok, B. W., & Guritno, S., 2006. Pemodelan B-Spline Dan MARS pada Nilai Ujian Masuk Terhadap IPK Mahasiswa Jurusan Disain Komunikasi Visual UK. PETRA Surabaya. *Jurnal Teknik Industri*. Vol 8, No. 1, 1–13. <https://doi.org/10.9744/jti.8.1.1-13>
- [7] de Boor, C. 2001. *A Practical Guide to Splines* (Revised Ed). Springer-Verlag, New York.
- [8] Erlangga, L. T. 2019. *Analisis Tingkat Pengangguran Terbuka di Indonesia menggunakan Regresi Nonparametrik B-spline*. Universitas Muhammadiyah Semarang, Semarang.
- [9] Husain, H., Dewi, A. F., & Wardani, A. E., 2024. Pemodelan Prevalensi Stunting Indonesia Menggunakan Regresi Nonparametrik Spline Truncated. *Journal of Analytical Research, Statistics and Computation*. Vol. 3, No. 1, 1–13.
- [10] Kemnaker. 2023. *Rencana Tenaga Kerja Jangka Panjang: Menuju Indonesia Emas 2045*. Satudata Kemnaker. <https://satudata.kemnaker.go.id/publikasi/118>. [11 November 2025]
- [11] Muhgni, M., Fadly, F., Adnan, A., & Harison., 2020. Pemodelan Tingkat Pengangguran Terbuka di Pulau Sumatera dengan Menggunakan Regresi Nonparametrik Spline. *Jurnal Sains Matematika Dan Statistika*. Vol. 6, No. 1.
- [12] Nasir, R. S., Wahid, M. A., Padang, D. R. A., Anna, I., & Raupong., 2024. Analisis

- Kemiskinan di Sulawesi Selatan dengan Regresi Nonparametrik Berbasis B-Spline. *Indonesian Journal of Applied Statistics*. Vol. 7, No. 1, 27–40. <https://doi.org/10.13057/ijas.v7i1.80716>
- [13] Prayogo, S. A., & Satria, D., 2020. Analisis faktor – faktor yang memengaruhi tingkat pengangguran terbuka kabupaten/kota di provinsi jawa timur tahun 2014-2018. *Jurnal Ilmiah Mahasiswa FEB*. Vol. 8, No. 2.
- [14] Rahmawati, A. S., Ispriyanti, D., & Warsito, B., 2017. Pemodelan Kasus Kemiskinan di Jawa Tengah Menggunakan Regresi Nonparametrik Metode B-Spline. *Jurnal Gaussian*. Vol. 6, No. 1, 11–20. <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- [15] Romhadhoni, P., Faizah, D. Z., & Afifah, N., 2018. Pengaruh Produk Domestik Regional Bruto ( PDRB ) Daerah Terhadap Pertumbuhan Ekonomi dan Tingkat Pengangguran Terbuka di Provinsi DKI Jakarta. *Jurnal Matematika Integratif*. Vol. 14, No. 2, 115–121. <https://doi.org/10.24198/jmi.v14.n2.2018.115-121>
- [16] Ruppert, D., Wand, M. P., & Carrol, R. J., 2003. Semiparametric Regression. In *Regression Modeling*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511755453>
- [17] Wardani, P. K., & Ratna, M., 2022. Pemodelan Terhadap Faktor-Faktor yang Mempengaruhi Pertumbuhan Ekonomi di Nusa Tenggara Timur Menggunakan Pendekatan Regresi Nonparametrik Spline. *Jurnal Sains Dan Seni ITS*. Vol. 11, No. 3, 2337–3520. <https://doi.org/10.12962/j23373520.v11i3.77735>
- [18] Wijaya, A. F. H., 2018. *Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Pengangguran Terbuka (TPT) Di Provinsi Aceh dengan Regresi Nonparametrik Spline Truncated*. Institut Teknologi Sepuluh Nopember, Surabaya. <https://api.semanticscholar.org/CorpusID:126778402>
- [19] Wulan, P. R., & Rifai, N. A. K., 2023. Penerapan Regresi Nonparametrik B-Spline pada Model Tingkat Pengangguran Terbuka Berdasarkan Tingkat Partisipasi Angkatan Kerja dan PDRB di Provinsi Jawa Barat. *Bandung Conference Series: Statistics*. Vol. 3, No. 2, 294–302. <https://doi.org/10.29313/bcss.v3i2.8095>