



Modeling the Percentage of Poor People in West Nusa Tenggara Using Truncated Spline Nonparametric Regression

Maryam Febrianti^{a}, Nanda Aulia Sudiasmini^b, Mustika Hadijati^c*

^{a,b,c} Department of Statistics, Faculty of Mathematics and Natural Sciences, University of Mataram, Indonesia.

*Corresponding author: maryam.febrianti11@gmail.com

ABSTRACT

Nonparametric regression approaches have received significant attention due to their high flexibility and ability to model data without relying on specific functional form assumptions. One widely applied method is the truncated spline, which effectively handles data behavior changes within specific sub-intervals. This study aims to model the Percentage of Poor People in West Nusa Tenggara Province from 2017 to 2025 using independent variables Average Years of Schooling (X1), Open Unemployment Rate (X2), Per Capita Expenditure (X3), and Life Expectancy (X4). The best model selection was performed by evaluating combinations of 1 to 3 knot points, using the minimum Generalized Cross Validation (GCV) as the primary criterion. Based on the analysis, the optimal truncated spline nonparametric regression model utilizes 1 knot point and degree 1, yielding a GCV value of 12.34126 and a coefficient of determination (R^2) of 59.38%. The identified optimal knot points are 8.015 for X1; 3.07 for X2; 10310.5 for X3; and 68.095 for X4.

Keywords: Poverty; West Nusa Tenggara; Nonparametric Regression; Truncated Spline; GCV

1. Introduction

Poverty is a fundamental global issue its existence represents a crucial challenge faced by almost every nation in the world. As a multidimensional phenomenon, poverty involves an inability to fulfill the most fundamental needs in human life [1].

Substantially, this problem is often rooted in a lack of income and assets. This condition prevents individuals or groups from accessing acceptable primary standards of living, including food, clothing, housing, health, and adequate education. Furthermore, poverty is closely linked to limited access to employment and possess inadequate levels of education and health. This condition creates a cycle of vulnerability that requires serious attention through comprehensive statistical analysis to break the chains of poverty [2].

In a national context, although Indonesia continues to show development progress, regional poverty disparities remain stark. Nationally, the poverty percentage is recorded at 8.25% [3]. However, a more complex portrait appears in West Nusa Tenggara (NTB) Province, which consistently records poverty rates above the national average.

Based on the latest data, although there is a downward trend from 11.78% in the March 2025 period to 11.38% in September 2025, the persistence of these double-digit figures indicated dynamic and non-



uniform local economic characteristics across districts and cities [4]. These fluctuations show that regional poverty patterns exhibit irregular behavior. This situation requires a modeling approach that is more sensitive to changes in field patterns without being constrained by rigid linearity assumptions.

The fluctuating characteristics of poverty data cause conventional parametric regression to often fail in providing optimal result. As a solution, the nonparametric regression approach offers very high flexibility. The regression curve is assumed to be smooth and capable of automatically adapting to the distribution of the original data [5].

The fundamental advantage of this method is its ability to produce objective estimates without being influenced by researcher subjectivity. This allows the data to determine the shape of its own regression curve estimates [6]. One very effective instrument in the nonparametric family for analyzing data with specific behavioral changes at certain intervals is spline truncated nonparametric regression.

Theoretically, the spline truncated model consists of piecewise polynomials that are segmented yet remain continuous. Its main strength lies in the use of knot points as markers for specific changes in data patterns. In determining the optimal spline regression model, the selection of the model can be done by observing the minimum generalized cross validation (GCV) value.

The minimum GCV criterion ensures that the resulting model has a balance between data fit and curve smoothness. The relevance of using this method in mapping social and economic phenomena has also been strengthened by empirical evidence from various previous studies. Previous research applying a similar approach successfully proved the superiority of this method with a MSE value of 3.24 [7].

This relatively low MSE value serves as a foundation and strong motivation for the current research. This proves that the spline truncated approach is capable of minimizing prediction errors on data with complex patterns. By using the results of previous research and the minimum GSV criterion as a reference, this study seeks to model the latest poverty dynamics in the NTB region.

Based on this background, this study aims to construct the best model for the percentage of poor people in NTB Province using a spline truncated nonparametric regression approach. In addition, this study is also directed toward performing predictions of the poverty percentage based on the model obtained.

2. Theoretical Framework

2.1. Poverty

Poverty is a multidimensional phenomenon that extends beyond mere income levels because it is closely related to various aspects of life such as environmental conditions, social structures, and levels of community participation [8]. Theoretically, poverty can be divided into natural poverty resulting from limited resources or low technological development and artificial or structural poverty occurring due to unequal institutional arrangements. Structural poverty is suffered by society because social structures hinder them from utilizing actually available income sources in a fair manner [9].

This condition is often triggered by a lack of asset ownership and low access to quality education. Furthermore, high unemployment rates are a major factor exacerbating the cycle of poverty because individuals lack access to decent and productive employment. The weak education system in many developing countries is unable to provide the skills needed by the community to participate fully in the modern economy, which results in persistent economic hardship that is difficult to break without intervention.

2.2. Nonparametric Regression

Nonparametric regression is a statistical approach used to determine the relationship between a response variable and a predictor variable when the functional form of the regression curve is unknown. Unlike the parametric approach which requires rigid functional form assumptions, this method provides the flexibility for the data to determine its own curve shape. The functional curve in this approach is assumed to be contained within a certain function space that possesses smooth properties.

The high flexibility in nonparametric regression allows it to capture complex data patterns without being forced into specific functional form assumptions. The general nonparametric regression model can be expressed through the following equation [6].

$$y_i = f(x_i) + \varepsilon_i \quad (1)$$

In this model, y_i represents the response variable at the i -th observation, while $f(x_i)$ is the regression function curve with an unknown specific form. The variable x_i acts as the independent variable, and ε_i is the random error assumed to be identically and independently distributed with a mean of zero and constant variance.

2.3. Spline Truncated Nonparametric Regression

Spline truncated nonparametric regression is an approach that has an excellent ability to handle data with behavioral patterns that change at certain sub-intervals. This method has the advantage of seeking its own estimation of data pattern movements through the existence of knot points symbolized by K_1, K_2, \dots, K_r [10]. Truncated spline itself consists of piecewise polynomials that are segmented at these knot points. The development of this function aims to overcome the weakness of standard global polynomial functions by maintaining basic polynomial properties but in a more locally adaptive form.

The univariable spline truncated function with degree m can be stated through a specific mathematical model. The univariable spline nonparametric regression model is presented in the following equation [11].

$$y_i = \sum_{j=0}^m \beta_j x_i^j + \sum_{k=1}^r \beta_{m+k} (x_i - K_k)_+^m + \varepsilon_i \quad (2)$$

In this equation, the variable y_i is the observation value, while β_j is the parameter for the polynomial model. The variable x_i is the predictor variable, and β_{m+k} is the parameter found in the truncated component. The symbol r indicates the number of knot points used in the model.

The component $(x_i - K_k)_+^m$ in the model is a truncated or piecewise function that provides flexibility to the curve. This function is defined with specific rules as follows.

$$\begin{aligned} (x_i - K_k)_+^m &= (x_i - K_k)^m, \text{ for } x_i \geq K_k \\ (x_i - K_k)_+^m &= 0, \text{ for } x_i < K_k \end{aligned}$$

The point K_k in this function is the knot point showing where changes in data behavior patterns occur. The use of this function ensures that the model can adapt to local data characteristics dynamically while maintaining overall curve continuity.

2.4. Generalized Cross Validation (GCV)

The best spline regression model is a model that possesses optimum knot points. These knot points are common joint points indicating changes in the behavioral patterns of the spline function curve at different intervals [12]. There are several methods that can be used to select optimum knot points in the spline approach, and one of the most commonly used methods is the GCV method. Determining the optimal knot points through this method is expected to produce a high coefficient of determination. Selecting the most appropriate knot points is done by seeking the minimum GCV value. This smallest value is a primary indicator that the resulting model has reached the most optimal balance point between data fit and curve smoothness. Mathematically, the GCV value for selecting optimum knot points can be expressed through the following equation [13].

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{trace}[I - A(K_1, K_2, \dots, K_r)])^2} \quad (3)$$

In this equation, I is the identity matrix and n is the total number of observations used. The component $A(K_1, K_2, \dots, K_r)$ is the hat matrix defined as $X(X^T X)^{-1} X^T$. The use of this criterion ensures that the chosen model has stable prediction performance and is free from estimation bias.

The magnitude of error in the modeling is expressed through the Mean Square Error (MSE) value. The MSE value related to the selection of knot points is shown through the following equation [11].

$$MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

Through achieving the minimum GCV value, researchers can ensure that the constructed spline regression model is the best representation of data dynamics in the field. This allows the obtained estimation results to have a high level of confidence and be scientifically accountable for further analysis.

3. Methods

The data utilized in this research are secondary data obtained officially through reports from the Central Bureau of Statistics (BPS) of West Nusa Tenggara (NTB) Province. The observation units in this study include ten districts and cities located within the West Nusa Tenggara Province region. The variables involved in this research consist of a response variable and predictor variables that will be analyzed to observe their relationship patterns toward the poverty level in the area.

The data analysis technique applied in this research is spline truncated nonparametric regression. The nonparametric regression approach is used as an alternative when the form of the regression curve is not known with certainty or does not follow a specific functional pattern. This method has a very high level of flexibility in performing regression curve estimation because the data are expected to find their own estimation form objectively.

The use of the spline truncated method in this research is based on its advantages in handling data behaviors that change at certain sub-intervals. This model tends to be able to follow the direction of data pattern movements dynamically so that the estimation results become more accurate. Besides having high flexibility, the spline method is also very superior in providing good visual representations in describing social and economic phenomena that are fluctuating in nature.

The following are the analysis steps used in this research.

- Data Collection and Processing

Secondary data collection is conducted through official reports of the Central Bureau of Statistics (BPS) of West Nusa Tenggara (NTB) Province by setting the percentage of the poor population as the variable y . The predictor variables used in sequence include the mean years of schooling (x_1), the open unemployment rate (x_2), the per capita expenditure (x_3), and life expectancy (x_4). All of these data are then processed systematically to ensure information validity before performing the nonparametric regression modeling stage.

- Development of the Modeling Program

The development of the modeling program is carried out by compiling computational algorithms using R software to support the entire range of nonparametric regression analysis. This stage includes writing program codes designed to search for the most optimal combinations of knot points automatically to produce accurate estimations.

- Spline Truncated Nonparametric Regression Modeling

The modeling stage is focused on determining the most optimal number and location of knot points to construct the spline truncated regression function. The selection of these knot points is based on achieving the minimum Generalized Cross Validation (GCV) value to ensure the accuracy and the goodness of fit of the resulting model.

- Model Application on Poverty Percentage Data in West Nusa Tenggara

The selected spline truncated model is then applied to the data of the percentage of the poor population in ten districts and cities within the West Nusa Tenggara Province region.

- Model Significance Testing

Model significance testing is conducted through simultaneous and partial parameter tests. This stage aims to ensure that the resulting model is valid and reliable in explaining the statistical relationships between variables.

- Poverty Percentage Prediction Based on the Obtained Model

Prediction is performed to estimate poverty figures based on the optimal model parameters that have been obtained through a series of previous statistical processes.

4. Results and Discussion

4.1. Descriptive Statistics

Descriptive statistical analysis was performed to provide an initial picture of the distribution and characteristics of each variable used in this study. The variables analyzed consist of one response variable, namely the Percentage of Poor People (Y), and four predictor variable such as Average Years of Schooling (X1), Open Unemployment Rate (X2), Per Capita Expenditure (X3), and Life Expectancy (X4). The descriptive statistics for all variables are presented in Table 1 below.

Table 1. Descriptive Statistics of Variables

Variable	Minimum	Maximum	Mean	Std. Deviation
Y	7.15	32.1	14.0	4.99
X1	5.54	11.0	7.96	1.46
X2	0.38	6.83	3.31	1.20
X3	8,006	16,851	10,768	2,004
X4	65.0	75.1	69.2	2.89

Based on Table 2, the response of variable Percentage of Poor People (Y) has a mean of 14.0% and a standard deviation of 4.99, indicating substansial variability across districts and cities. The range between the minimum (7.15%) and maximum (32.1%) values is strikingly wide, which spanning nearly 25 percentage points. Such a wide spread signals that poverty is distributed non-uniformly across NTB, with certain areas experiencing poverty levels more than four times that of the best-performing areas.

Among the predictor variables, Per Capita Expenditure (X3) exhibits the largest absolute variability, with a mean of Rp10,768 (in thousand of rupiah per month), with standard deviation of 2,400 and a range from 8,006 to 16,851. This suggests profound economic stratification within the province, some districts have per capita expenditures twice as high as others, reflecting stark disparities in living standard and economic. The Average Years of Schooling (X1) across ten districts and cities in West Nusa Tenggara Barat Province ranged from a minimum of 5.54 years to a maximum of 11. years, with a mean of 7.96 and a standard deviation of 1.46. This variation reflects the disparity in educational attainment levels among regions in NTB. The response of variable Open Unemployment Rate (X2) displays the widest relative spread among the predictors, the range between 0.38% to 6.83% with a mean of 3.31%. The coefficient of variation for variable Open Unemployment Rate (X2) is approximately 36.3% suggesting it is the most heterogeneous predictor relative to its mean. By contrast, Life Expectancy (X4) ranges from 65.0 to 75.1 years with relatively small standard deviation of 2.89, suggesting that while health outcomes vary, they do so less drastically than economic indicators.

4.2. Scatterplot Analysis

A scatterplot analysis was performed to examine the pattern of the relationship between the response variable Y and each predictor variable.

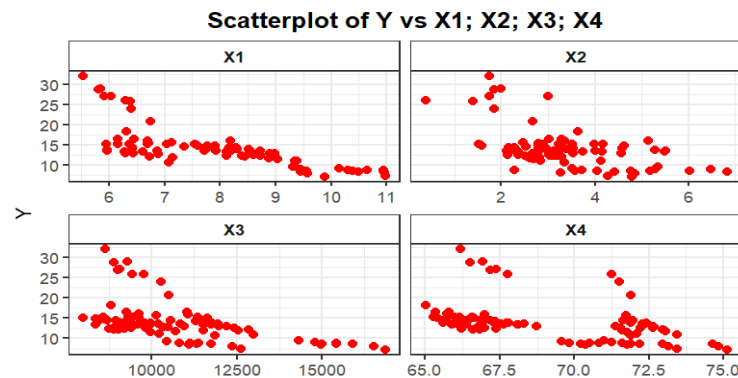


Figure 1. Scatterplots

Based on the scatterplots, it is evident that the relationship between the poverty percentage (Y) and each predictor variable does not follow a specific linear or parametric functional form. The data points exhibit irregular, non-uniform patterns that change at certain intervals. This behavior suggests the presences of structural changes in the data pattern, making parametric regression approaches insufficient for capturing the true relationships among variables.

These irregular patterns justify the use of the truncated spline nonparametric regression approach, which is specifically designed to handle data with behavioral changes at particular sub-intervals through the use of knot points. The flexibility of this approach allows the regression curve to adapt to the local characteristics of the data without imposing rigid functional form assumptions.

4.3. Modeling with Truncated Spline Nonparametric Regression

Model selection in truncated spline nonparametric regression out by determining the optimal number and location of knot points based on the minimum Generalized Cross Validation (GCV) criterion. The evaluation was conducted for models with 1, 2, and 3 knot points at degrees 1 to 3. The GCV comparison results for the one-knot model are presented in Table 2.

Table 2. Comparison of GCV values for the one-knot model

Quantil	Knot Point				GCV
	K1	K2	K3	K4	
0.5	8.015	3.070	10310.5	68.095	12.3413*
0.4	7.564	2.870	9811.0	67.328	12.7217
0.1	6.159	1.976	8738.6	65.962	12.8213
0.2	6.408	2.358	9051.2	66.370	13.8948
0.6	8.340	3.332	10988.0	70.890	14.6792
0.3	6.744	2.694	9407.5	66.855	17.4266
0.9	9.915	5.154	13116.3	72.970	24.6987
0.8	9.356	4.192	11888.0	72.112	24.7862
0.7	8.723	3.591	11425.3	71.671	24.8083

Table 2 shows that among all quantile combinations tested for the one-knot model, combination at quantile 0.5 produces the lowest GCV value of 12.3413, with optimal knot points at K1 = 8.015 (X1), K2 = 3.07 (X2), K3 = 10310.5 (X3), and K4 = 68.095 (X4). A similar evaluation was conducted for models with 2 and 3 knot points. Furthermore, by substituting the estimation of parameters and the location of the node point, a nonparametric regression model of a truncated spline is obtained as follows:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 (x_{1i} - K_1)_+ + \beta_3 x_2 + \beta_4 (x_2 - K_2)_+ + \beta_5 x_3 + \beta_6 (x_3 - K_3)_+ + \beta_7 x_4 + \beta_8 (x_4 - K_4)_+$$

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 (x_{1i} - 8.015)_+ + \beta_3 x_2 + \beta_4 (x_2 - 3.070)_+ + \beta_5 x_3 + \beta_6 (x_3 - 10310.5)_+ + \beta_7 x_4 + \beta_8 (x_{4i} - 68.095)_+$$

Table 3. Comparison of GCV values for the two-knot model

Knot Point								GCV
K11	K12	K21	K22	K31	K32	K41	K42	
6.4080	9.3560	2.3580	4.1920	9051.20	11888.00	66.3700	72.1120	14.6336*
7.5640	8.3400	2.8700	3.3320	9811.00	10988.00	67.3280	70.8900	14.8974
6.6225	8.9775	2.5350	4.0025	9270.75	11622.25	66.5850	71.8800	16.0861
6.7440	8.7230	2.6940	3.5910	9407.50	11425.30	66.8550	71.6710	17.4858
6.9392	8.5744	2.7837	3.5274	9439.29	11326.20	66.9274	71.5348	17.9206

Table 3 shows the knot locations for each predictor variable using a two-knot truncated spline, based on the five smallest GCV values. It can be observed that the minimum GCV value is 14.6336, which corresponds to the optimal knot configuration. For the predictor variable X_1 , the selected knot

points are located at 6.4080 and 9.3560. For X_2 , the knot points are 2.3580 and 4.1920. Furthermore, the knot locations for X_3 are 9051.20 and 11888.00, while for X_4 , they are 66.3700 and 72.1120.

By substituting the estimated parameters and the corresponding optimal knot points into the truncated spline regression model, the nonparametric regression model with two knots is obtained as follows:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 (x_1 - K_{11})_+ + \beta_3 (x_1 - K_{12})_+ + \beta_4 x_2 + \beta_5 (x_2 - K_{21})_+ + \beta_6 (x_2 - K_{22})_+ + \beta_7 x_3 + \beta_8 (x_3 - K_{31})_+ + \beta_9 (x_3 - K_{32})_+ + \beta_{10} x_4 + \beta_{11} (x_4 - K_{41})_+ + \beta_{12} (x_4 - K_{42})_+$$

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 (x_1 - 6.4080)_+ + \beta_3 (x_1 - 9.3560)_+ + \beta_4 x_2 + \beta_5 (x_2 - 2.3580)_+ + \beta_6 (x_2 - 4.1920)_+ + \beta_7 x_3 + \beta_8 (x_3 - 9051.20)_+ + \beta_9 (x_3 - 11888.00)_+ + \beta_{10} x_4 + \beta_{11} (x_4 - 66.3700)_+ + \beta_{12} (x_4 - 72.1120)_+$$

Table 4. Comparison of GCV values for the three-knot model

Knot Point												GCV
K11	K12	K13	K21	K22	K23	K31	K32	K33	K41	K42	K43	
6.307	7.770	8.977	2.21	3.010	4.002	8921.4	9989.55	11622.2	66.163	67.63	71.88	15.7868
0	5	5	8	5	5	0	0	5	5	5	0	*
6.159	7.564	8.723	1.97	2.870	3.591	8738.6	9811.00	11425.3	65.962	67.32	71.67	15.8953
0	0	0	6	0	0	0	0	0	0	8	1	
6.408	8.015	9.356	2.35	3.070	4.192	9051.2	10310.5	11888.0	66.370	68.09	72.11	16.2141
0	0	0	8	0	0	0	0	0	0	5	2	
6.622	8.015	8.977	2.53	3.070	4.002	9270.7	10310.5	11622.2	66.585	68.09	71.88	17.356
5	0	5	5	0	5	5	0	5	0	5	0	
6.744	8.340	9.915	2.69	3.332	5.154	9407.5	10988.0	13116.3	66.855	70.89	72.97	19.0849
0	0	0	4	0	0	0	0	0	0	0	0	

Table 4 shows the knot locations for each predictor variable using a three-knot truncated spline model, based on the five smallest GCV values. It is observed that the minimum GCV value is 15.7868, indicating the optimal combination of knot points.

For the predictor variable X_1 , the selected knot points are located at 6.3070, 7.7705, and 8.9775. For X_2 , the knot points are 2.218, 3.0105, and 4.0025. Furthermore, the knot locations for X_3 are 8921.40, 9989.55, and 11622.25, while for X_4 , the knot points are 66.1635, 67.635, and 71.880. The minimum CGV of the three-knot model is the highest among all three configurations compared, surpassing both the one-knot and two-knot models. The results confirms that progressively adding knot points beyond one does not generalization and leads to model overfitting.

By substituting the estimated parameters and the corresponding optimal knot points into the truncated spline regression model, the nonparametric regression model with three knots is obtained as follows:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 (x_1 - K_{11})_+ + \beta_3 (x_1 - K_{12})_+ + \beta_4 (x_1 - K_{13})_+ + \beta_5 x_2 + \beta_6 (x_2 - K_{21})_+ + \beta_7 (x_2 - K_{22})_+ + \beta_8 (x_2 - K_{23})_+ + \beta_9 x_3 + \beta_{10} (x_3 - K_{31})_+ + \beta_{11} (x_3 - K_{32})_+ + \beta_{12} (x_3 - K_{33})_+ + \beta_{13} x_4 + \beta_{14} (x_4 - K_{41})_+ + \beta_{15} (x_4 - K_{42})_+ + \beta_{16} (x_4 - K_{43})_+$$

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 (x_1 - 6.3070)_+ + \beta_3 (x_1 - 7.7705)_+ + \beta_4 (x_1 - 8.9775)_+ + \beta_5 x_2 + \beta_6 (x_2 - 2.2180)_+ + \beta_7 (x_2 - 3.0105)_+ + \beta_8 (x_2 - 4.0025)_+ + \beta_9 x_3 + \beta_{10} (x_3 - 8921.40)_+ + \beta_{11} (x_3 - 9989.55)_+ + \beta_{12} (x_3 - 11622.25)_+ + \beta_{13} x_4 + \beta_{14} (x_4 - 66.1635)_+ + \beta_{15} (x_4 - 67.6350)_+ + \beta_{16} (x_4 - 71.8800)_+$$

Table 5. Summary of Minimum GCV by Number of Knot Points

Model	Number of Knots	Minimum GCV	R ²	MSE
Model 1	1 Knot	12.3413*	59.38%	9.9964
Model 2	2 Knot	14.6336	55.09%	11.0524
Model 3	3 Knot	15.7868	56.26%	10.7635

Table 5 above demonstrates that the one-knot model yields the smallest GCV value of 12.3413, compared to the two-knot model (GCV = 14.6336) and the three-knot model (GCV = 15.7868). In accordance with the minimum GCV principle, the one-knot model with degree 1 is selected as the optimal truncated spline nonparametric regression model. This result implies that a single behavioral change point in each predictor variable is sufficient to capture the dynamics of poverty in West Nusa Tenggara Province. The coefficient of determination (R^2) of the optimal model is 59.38%, explaining a substantive majority of poverty variation across 90 district-year observations while the residuals 40.6% of unexplained variance is attributed to other factors outside the model.

4.4. Simultaneous Significance Test

Simultaneous hypothesis testing was conducted to determine whether the predictor variables jointly have a significant effect on the response variable. The hypotheses for the simultaneous test are defined as follows:

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_8 = 0$$

$$H_1 : \text{at least one } \beta_j \neq 0, j = 0, 1, 2, \dots, 8$$

The results of the simultaneous test based on the generated model are present below:

Table 6. Simultaneous Test Results

Source of Variation	Degrees of Freedom	Sum of Square	Mean Square	F	p-value
Regression	8	1314.993	164.374	14.8	0
Error	81	899.678	11.107		
Total	89	2214.670	2214.670		

Based on Table 6, using a significance level α of 0.05, the simultaneous test results indicate that H_0 is rejected as the p-value is $0.000 < 0.05$. This implies that at least one independent variable significantly affects the dependent variable. Consequently, the single-knot truncated spline model is considered feasible for this research.

4.5. Partial Significance Test

Partial testing is conducted to evaluate the significance of each individual parameter within the single-knot truncated spline model. The hypotheses for this partial test are as follows:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 0, 1, 2, \dots, 8$$

The results of the partial parameter estimation are summarized in the following table:

Table 7. Partial Parameter Test Results

Parameter	Estimator	p-value
β_0	0.030	0.0000
β_1	-1.606	0.0000
β_2	-0.999	0.0000
β_3	-0.436	0.0000
β_4	-0.256	0.0002
β_5	-0.002	0.0175
β_6	0.002	0.0477
β_7	0.679	0.0000
β_8	-0.747	0.0001

Based on the results presented in Table 7, using a significance level α of 0.05, it is observed that the p-values for all parameters are less than 0.05. Consequently, the null hypothesis (H_0) is rejected for every parameter. This indicates that all linear and truncated components within the model exert a statistically significant influence on the dependent variable. These findings confirm that the inclusion of each parameter effectively contributes to the model's ability to explain the variance in the observed data.

4.6. Poverty Rate Prediction in West Nusa Tenggara Based on the Developed Model

Utilizing the estimated single-knot truncated spline nonparametric regression model, the findings of this study can be employed to project the poverty rate in West Nusa Tenggara Province from 2017 to 2023. An example of the estimation process for West Lombok Regency in the year 2017 is as follows.

The Poverty Percentage Model

$$\hat{y}_i = 0.030 - 1.606x_{1i} - 0.999(x_{1i} - 8.015)_+ - 0.436x_{2i} - 0.256(x_{2i} - 3.070)_+ - 0.002x_{3i} + 0.002(x_{3i} - 10310.5)_+ + 0.679x_{4i} - 0.747(x_{4i} - 68.095)_+$$

The prediction for West Lombok Regency in the year 2017 is calculated as

$$\hat{y}_1 = 0.030 - 1.606x_{11} - 0.999(x_{11} - 8.015)_+ - 0.436x_{21} - 0.256(x_{21} - 3.070)_+ - 0.002x_{31} + 0.002(x_{31} - 10310.5)_+ + 0.679x_{41} - 0.747(x_{41} - 68.095)_+$$

The same procedure is carried out iteratively through 2025 for all regencies and cities. Consequently, the plot comparing the actual poverty percentages and their predictions in West Nusa Tenggara is obtained as follows

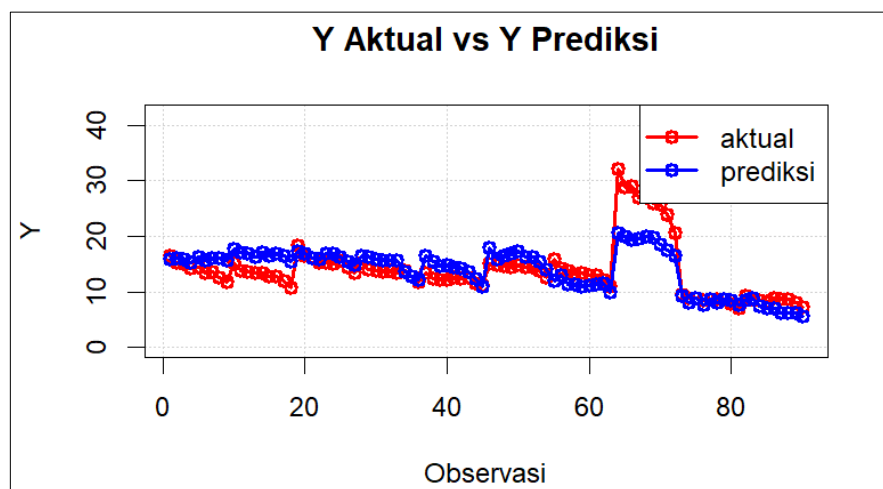


Figure 2. Plot of Actual vs. Predicted Poverty Rates in West Nusa Tenggara

5. Conclusion

The optimal model for the poverty percentage in West Nusa Tenggara Province, developed using a single-knot truncated spline nonparametric regression approach, is formulated as follows:

$$\hat{y}_i = 0.030 - 1.606x_{1i} - 0.999(x_{1i} - 8.015)_+ - 0.436x_{2i} - 0.256(x_{2i} - 3.070)_+ - 0.002x_{3i} + 0.002(x_{3i} - 10310.5)_+ + 0.679x_{4i} - 0.747(x_{4i} - 68.095)_+$$

This model achieves a GCV value of 12.3413 and accounts for 59.38% of the variance in poverty data, as indicated by the R^2 value. All parameters within the model are statistically significant, confirming the effectiveness of the selected optimal knots.

Predictions of poverty rates for all regencies and cities in West Nusa Tenggara for the 2017–2025 period were based on the general model obtained above. According to the comparison plot of actual versus predicted data, the predicted patterns closely align with the fluctuations in the actual data. This is further substantiated by an overall MSE of 9.9964, indicating that the model maintains a relatively low error rate and high precision in projecting poverty figures across the province. Consequently, this model serves as a reliable instrument for evidence-based policy making and strategic planning for poverty reduction in West Nusa Tenggara.

REFERENCE

- [1] Y. Yacoub. "The effect of unemployment rates on poverty rates in districts/cities of West Kalimantan Province." *Jurnal Eksos*. vol. 8. no. 3. pp. 176–185. Oct. 2012.
- [2] World Bank. *World Development Report 2004: Making Services Work for Poor People*. Washington, DC. USA: World Bank and Oxford University Press. 2004.
- [3] BPS-Statistics Indonesia. "Percentage of poor people (P0) by province and region (percent). 2025." *BPS-Statistics Indonesia*. Feb. 5. 2026. [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/MTkylzl=/persentase-penduduk-miskin--september-2025.html>. [Accessed: Apr. 2. 2026].
- [4] BPS-Statistics of West Nusa Tenggara Province. "Percentage of poor people in West Nusa Tenggara Province by period (percent). 2025." *BPS-Statistics of West Nusa Tenggara Province*. Feb. 6. 2026. [Online]. Available: <https://ntb.bps.go.id/id/statistics-table/2/MTYxlzl=/persentase-penduduk-miskin--september-2025.html>. [Accessed: Apr. 2. 2026].
- [5] S. N. Tupen and I. N. Budiantara. "Hypothesis testing in spline nonparametric regression." in *Proc. National Statistics Seminar. Diponegoro University 2011*. Semarang. Indonesia. 2011. pp. 184–199.
- [6] R. L. Eubank. *Spline Smoothing and Nonparametric Regression*. New York, NY. USA: Marcel Dekker. 1988.
- [7] H. Husain, A. F. Dewi, and A. E. Wardani. "Modeling of Indonesian stunting prevalence using spline truncated nonparametric regression." *J. Anal. Res. Stat. Comput.*. vol. 3. no. 1. pp. 1–13. Mar. 2024.
- [8] N. Primadianti and C. Sugiyanto. "Regional inequality, pro-poor economic growth, and poverty in West Nusa Tenggara." *J. Din. Ekon. Pemb.*. vol. 3. no. 1. pp. 1–20. 2020.
- [9] M. Amin. "Impact of the PIJAR program on poverty and unemployment reduction efforts in West Nusa Tenggara Province." *Jurnal Distribusi*. vol. 8. no. 1. pp. 1–12. Mar. 2020.
- [10] I. Budiantara. *Spline in Nonparametric and Semiparametric Regression: A Statistical Approach for Today and the Future*. Surabaya. Indonesia: ITS Press. 2009.
- [11] A. F. D. Rositawati and I. N. Budiantara. "Modeling of provincial happiness index in Indonesia using spline truncated nonparametric regression." *J. Sains Seni ITS*. vol. 8. no. 2. pp. D113–D120. 2019.
- [12] W. Hardle. *Applied Nonparametric Regression*. New York: Cambridge University Press. 1994.
- [13] H. Wu dan J. T. Zhang. *Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-Effects Modeling Approaches*. Hoboken, NJ. USA: John Wiley & Sons. 2006.