E-ISSN 2586-9027
Homepage : *https://tci-thaijo.org/index.php/SciTechAsia* Science & Technology Asia
Vol.30 No.4 October - December 2025 Page: [112-125]
Original research article

Interval Estimation in Truncated Spline Regression: Analyzing the 2023 Indonesian Democracy Index

Andrea Tri Rian Dani[1,2,*], Fachrian Bimantoro Putra[3], Meirinda Fauziyah[1], Ratna Kusuma[4], Darnah[1], Hadi Koirudin[1], Chiko Zet Oroh[1], Raditya Arya Kosasih[1]

[1]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda 75119, Indonesia*
[2]*Doctoral Study Program of Mathematics and Natural Sciences, Faculty of Science and Technology, Airlangga University, Surabaya 60115, Indonesia*
[3]*BNI Staff Banking, Future Relationship Manager, Jakarta 10220, Indonesia*
[4]*Department of Biology, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda 75119, Indonesia*

Received 29 March 2025; Received in revised form 15 September 2025
Accepted 8 October 2025; Available online 17 December 2025

ABSTRACT

The truncated spline is a piecewise polynomial that maintains continuity across segments and offers a high degree of flexibility in estimating data that varies significantly across different intervals. This study applies the truncated spline method with interval estimation to analyze the 2023 Indonesian Democracy Index (IDI) data. The objective is to identify the key factors that significantly influence IDI and to evaluate the estimated results of the nonparametric spline regression curve for the 2023 IDI data. The analysis findings indicate that the predictor variables used in this study have a substantial impact on IDI, with a determined coefficient of approximately 97.81%. Based on the interval estimation analysis of the regression curve, it was found that out of 34 provinces in Indonesia, 8 provinces were estimated accurately. These provinces include Kepulauan Riau, DKI Jakarta, Bali, Central Kalimantan, North Sulawesi, Central Sulawesi, North Maluku, and Papua. Additionally, the analysis indicates that 16 provinces have experienced a decline in their Democracy Index (IDI) criteria, suggesting a possible transition in their democratic performance. This situation calls for further in-depth studies to improve their democratic achievement levels. On the other hand, 18 provinces are projected to experience an increase in their IDI performance.

Keywords: Democracy; Interval estimation; IDI; Nonparametric regression; Truncated spline

***Corresponding author:** andreatririandani@fmipa.unmul.ac.id **doi:** xx.xxxxx/scitechasia.xxxx.xx

1. Introduction

In this modern era, advancements in technology and the increasing volume of data have added complexity to statistical analysis. The growing diversity of data, which does not always conform to specific distribution assumptions, necessitates more flexible methods in statistical modeling. One such approach is nonparametric regression.

Nonparametric regression is a statistical method used to identify the relationship pattern between predictor and response variables without prior knowledge of the regression curve's shape [1]. The causal relationship between the response variable and the predictor variable is assumed to be smooth, meaning it is contained within a specific function space [2, 3]. The advantage of nonparametric regression lies in its flexibility in capturing complex relationship patterns. This flexibility is evident when the modeled nonparametric regression does not require prior information or knowledge about the curve's shape [4].

Several estimators can be used in nonparametric regression, including Kernel [5], Spline [6], Local Polynomial [7], Fourier Series [8], Wavelets [9] , MARS [10], and others. One of the most used estimators in nonparametric regression modeling is the spline estimator [11, 12].

A spline is a segmented polynomial that maintains continuity across its segments. One of its key advantages is its flexibility, as it incorporates knot points that indicate changes in data patterns [13, 14]. There are several basis functions in the spline estimator, and in this study, the researcher utilizes the truncated spline basis [15]. The advantage of the truncated spline estimator lies in its simple statistical interpretation and effective visual representation. Additionally, the model can accommodate fluctuating data patterns with the help of knot point [16, 17].

Previous studies that have examined the truncated spline estimator were conducted by [18-21]. Previous studies were limited to examining nonparametric regression modeling using the truncated spline estimator based solely on point estimation. A major drawback of point estimation is that it does not indicate the strength of the estimation's accuracy and is highly prone to errors. This limitation can be overcome by applying interval estimation [22, 23].

Interval estimation of the truncated spline nonparametric regression curve can indicate the upper and lower bounds of the resulting regression curve [24]. This method helps ensure the model's accuracy within a specific range while accounting for a predetermined level of error [25, 26]. This undoubtedly affects the probability of model accuracy compared to point estimation of the regression curve [27].

Previous studies that examined interval estimation were conducted by [22, 27-29]. However, this research introduces a new approach by applying truncated spline nonparametric regression modeling with interval estimation to the analysis of the Indonesian Democracy Index (IDI), a novel application in the field of political science. While interval estimation has been used in various domains, its application to the IDI, especially using a nonparametric regression model, is relatively unexplored.

The Indonesian Democracy Index (IDI) is an objective and empirical measurement tool that assesses the state of political democracy in Indonesia across three key aspects: civil liberties, political rights, and democratic institutions. The primary goal of IDI is to quantitatively measure the level of democratic development. It serves as a general check-up tool for democracy at both the national and provincial levels. Additionally, it is important to emphasize that IDI is not solely intended to evaluate government performance. The indicators, variables, and aspects that make up IDI do not only measure the scope of government responsibilities but also capture the broader democratic dynamics emerging within society [27].

2. Methods And Materials

2.1 Nonparametric regression

Nonparametric regression is an approach used to identify the relationship pattern between a response variable and predictor variables when the shape of the regression curve is unknown and there is no complete prior information about the data pattern [4]. In general, a nonparametric regression model follows the same structure as conventional regression models, which can be expressed by the following Eq. (2.1) [30]

$$y_i = f(x_i) + \varepsilon_i, i = 1, 2, 3, , n, \tag{2.1}$$

where y_i is the response variable for the i-th observation. $f(x_i)$ is the unknown nonparametric function. ε_i is the error term associated with the i-th observation.

2.2 Truncated spline nonparametric regression

A truncated spline is a segmented polynomial that maintains continuity while offering a high degree of flexibility [30]. This capability is demonstrated through the truncated function, where these segments represent knot points [31]. A univariate truncated spline function of order q with knot points $\phi_1, \phi_2, , \phi_r$ can be expressed by the following Eq. (2.2) [32].

$$f(x) = \sum_{u=0}^{q} \theta_u x_i^u + \sum_{k=1}^{r} \theta_{q+k} (x_i - \phi_k)_+^q, \tag{2.2}$$

where x_i is the predictor variable for the i-th observation. θ_u, θ_{q+k} are coefficients of the polynomial and truncated spline components. + is the + sign in the formula $(x_i - \phi_k)_+^q$ refers to the positive part function.

With the truncated function defined as in the following Eq. (2.3).

$$(x_i - \phi_k)_+^q = \begin{cases} (x_i - \phi_k), & x_i \geq \phi_k \\ 0, & x_i < \phi_k \end{cases}, \tag{2.3}$$

where ϕ_k represents the knot points, and $(x_i - \phi_k)_+^q$ is the truncated spline component for each knot.

2.3 Optimal knot point selection

Several methods have been developed by researchers to determine the optimal knot points, one of which is Generalized Cross-Validation (GCV). The GCV is an extension of the traditional Cross-Validation (CV) technique, adapted to accommodate nonparametric regression models. It calculates an estimate of the prediction error by adjusting for the complexity of the model, which is determined by the number and placement of knot points. The GCV minimizes the residual sum of squares while considering the effective degrees of freedom, which are influenced by the model's smoothing parameter.

The best truncated spline nonparametric regression model is obtained by minimizing the GCV value, which is formulated in the following Eq. (2.4) [33].

$$GCV = \frac{n^{-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)}{(n^{-1} trace(I - A(\phi_1, \phi_2, , \phi_r)))}, \tag{2.4}$$

where $\phi_1, \phi_2, , \phi_r$ represents the knot points, I is the identity matrix, and $A(\phi_1, \phi_2, , \phi_r)$ can be expressed by the following Eq. (2.5).

$$A(\phi_1, , \phi_r) = X(\phi)[X(\phi)^T X(\phi)]^{-1} X(\phi)^T, \quad (2.5)$$

where $X(\phi)$ is the design matrix based on the knot points.

2.4 Interval estimation of truncated spline nonparametric regression

The interval estimation of multivariable truncated spline nonparametric regression is obtained by determining the pivotal quantity for $\hat{f}_i(x)$. One approach to determining the pivotal quantity is by transforming the Equation in the following by Eq. (2.6) [27].

$$T_i(x, y) = \frac{\hat{f}_i(x) - E(\hat{f}_i(x))}{\sqrt{(Var(\hat{f}_i(x)))_{ii}}}, \quad (2.6)$$

where $\hat{f}_i(x)$ is the predicted value from the nonparametric regression model. $E(\hat{f}_i(x))$ is the expected value of $\hat{f}_i(x)$. $Var(\hat{f}_i(x)) =$ is the variance value of $\hat{f}_i(x)$.

The pivotal quantity is a concept used in statistical inference, particularly in the construction of confidence intervals and hypothesis testing [34]. It is a function of the observed data and the parameters of interest that has a known probability distribution, independent of the parameters. In the context of your study on interval estimation for truncated spline nonparametric regression, the pivotal quantity is used to calculate the confidence intervals for the estimated values of the response variable [35]. Based on Eq. (2.6), the pivotal quantity is obtained as shown in Eq. (2.7).

$$T_i(x, y) = \frac{\hat{f}_i(x) - f_i(x)}{\sqrt{\frac{y^T(1-A(\phi))y}{n-(1+p+pr)} A(\phi)_{ii}}}. \quad (2.7)$$

The process of constructing the interval estimation for multivariable truncated spline nonparametric regression can be completed using a probability equation of size $(1 - \alpha)$ with $0 \leq \alpha \leq 1$.

In nonparametric regression, such as the truncated spline model used in this analysis, the model is free from parametric assumptions regarding the functional form between the response variable and the predictors. This means that classical assumptions such as normality of errors, independence, and constant variance (homoscedasticity), which are typically required in linear regression, are not necessary in nonparametric regression.

However, although nonparametric regression does not require assumption checks, the pivotal quantity assumption is still used to construct confidence intervals. The pivotal quantity is a statistical quantity that plays a crucial role in building the estimation interval for predictions or model parameters. In this context, confidence intervals for predicted IDI values are based on the distribution of the pivotal quantity, allowing us to calculate the lower and upper bounds of the estimates at a specified confidence level.

Suppose there is a constant $a_i(x_i) = a_i$ and $b_i(x_i) = b_i$, pivotal quantity $T_i(x, y)$, with α is the error level, then the probability equation can be formulated as shown in Eq. (2.8).

$$P(a_i \leq T_i(x, y) \leq b_i) = 1 - \alpha, 0 < \alpha < 1. \quad (2.8)$$

This is equivalent to:

$$P\left(a_i \leq \frac{\hat{f}_i(x) - f_i(x)}{\sqrt{\frac{y^T(1-A(\phi))y}{n-(1+p+pr)} A(\phi)_{ii}}}\right) = 1 - \alpha. \quad (2.9)$$

Eq. (2.9) can be rewritten as Eq. (2.10).

$$P\left(\hat{f}_i(x) - b_i\sqrt{\frac{y^T(1-A(\phi))y}{n-(1+p+pr)}A(\phi)_{ii}} \leq f_i(x) \leq \hat{f}_i(x) - a_i\sqrt{\frac{y^T(1-A(\phi))y}{n-(1+p+pr)}A(\phi)_{ii}}\right) = 1-\alpha. \quad (2.10)$$

Based on the point estimation, Eq. (2.10) can be constructed into the shortest interval estimation for the multivariable truncated spline nonparametric regression curve, as shown in Eq. (2.11).

$$P\left(\hat{f}_i(x) - t_{(\frac{\alpha}{2},(n-(1+p+pr)))}\sqrt{\frac{y^T(1-A(\phi))y}{n-(1+p+pr)}A(\phi)_{ii}} \leq f_i(x) \leq \hat{f}_i(x) - t_{(\frac{\alpha}{2},(n-(1+p+pr)))}\sqrt{\frac{y^T(1-A(\phi))y}{n-(1+p+pr)}A(\phi)_{ii}}\right) = 1-\alpha, \quad (2.11)$$

where $t_{(\frac{\alpha}{2},(n-(1+p+pr)))}$ is the critical value from the t-distribution at the specified confidence level. a_i, b_i are Constants that define the interval boundaries for the response variable.

2.5 Indonesian democracy index

The global wave of democratization can be likened to a tidal surge, displacing authoritarian regimes and replacing them with democratic systems. In the wake of this widespread transformation, there arises a critical need to evaluate the extent of democratic progress, including within the context of Indonesia. For Indonesia, it is essential to assess the development of democracy at the regional level, as the overall success of its democratic governance is fundamentally linked to the degree to which democratic principles are cultivated and enacted across its provinces.

The Indonesian Democracy Index (IDI) is a comprehensive and empirical measurement tool designed to assess the state of political democracy in Indonesia. It evaluates the nation's democratic development across three core dimensions: civil liberties, political rights, and democratic institutions. The primary aim of the IDI is to provide a quantitative assessment of democratic progress, offering a systematic approach to measure the advancement of democracy. As such, the IDI serves as a vital diagnostic tool for evaluating democratic performance, both at the national level and within each province, providing insights into areas of strength and those requiring further attention for democratic consolidation [27].

3. Result and Discussions

3.1 Description of data

The data used in this study is secondary data, specifically the Indonesian Democracy Index (IDI). The population in this study consists of all provinces in Indonesia, while the sample includes IDI data from 34 provinces in Indonesia for the year 2023. The Indonesian Democracy Index serves as the response variable, with seven predictor variables utilized in this study. A detailed presentation of the variables used is provided in Table 1.

3.2 Data exploration

Based on the descriptive statistics analysis, it can be observed that no province in Indonesia has a low Indonesian Democracy Index (IDI) score. All provinces fall within the medium and high categories, indicating an overall improvement in the quality of democracy across the country. This descriptive statistics analysis provides

Table 1. Research Variables.

Notation	Variable	Description	Operational Definition
Y	IDI	Indonesian Democracy Index	Index measuring the level of democracy in Indonesia across civil liberties, political rights, and democratic institutions.
X_1	GDI	Gender Development Index	Index measuring gender disparity in development, focusing on gender-related aspects of human development.
X_2	RGDP	Regional Gross Domestic Product Growth Rate	The annual percentage change in the Gross Domestic Product of a region, reflecting economic growth.
X_3	PPP	Percentage of Poor Population	The proportion of the population living below the poverty line in each region.
X_4	ICTDI	Information and Communication Technology Development Index	Index measuring the development of information and communication technology infrastructure and usage.
X_5	GR	Gini Ratio	A measure of income inequality in a region.
X_6	GEI	Gender Empowerment Index	Index measuring the empowerment of women based on economic, political, and social participation.
X_7	WPP	Women's Participation in Parliament	The percentage of seats held by women in the parliament of a region.

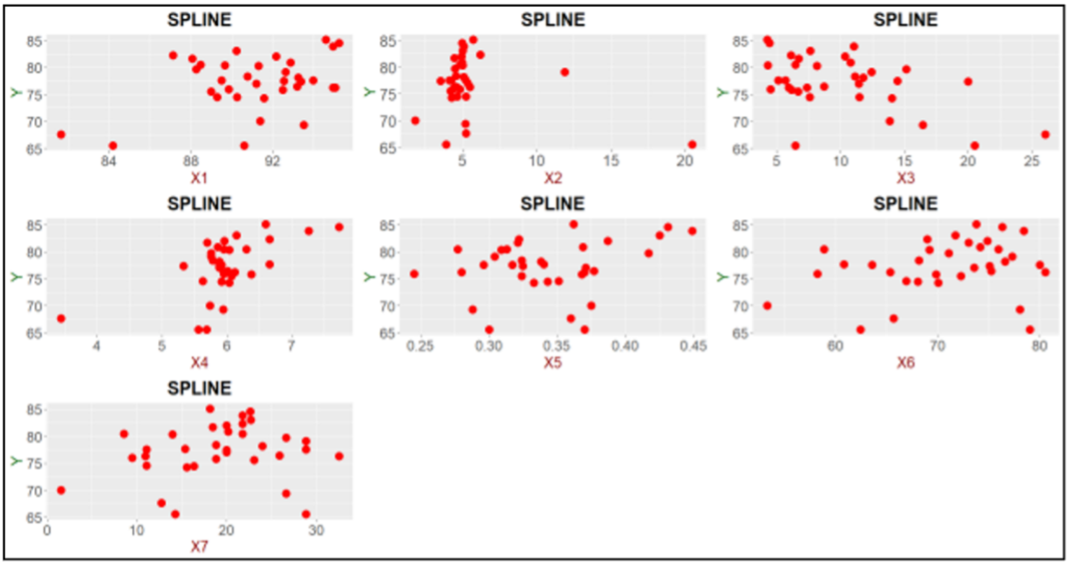

Fig. 1. Scatter plot of predictor variables.

an overview of the distribution of IDI scores at the provincial level, suggesting that while there are differences between provinces, the majority show positive progress in democratic development.

Next, a scatter plot for each predictor variable is presented in Fig. 1. This plot will provide a visual representation of the relationship between the variables influencing IDI across provinces.

Based on Fig. 1, it can be observed that the data pattern between the response variable (IDI) and all predictor variables does not follow a specific trend. The data points are scattered, with some observations positioned far from the primary data distribution. This suggests that a nonparametric regression approach is a suitable choice, given its flexibility in handling nonlinear relationships and unstructured data patterns, as illustrated by the scatter plot.

Based on Fig. 2, the IDI displays a fluctuating trajectory over the years, highlighting the dynamic and evolving nature of democracy in Indonesia. Although the IDI has shown improvements in certain years, there have been notable periods of decline. In 2023, the index recorded a decrease relative to the previous year. This trend warrants a more comprehensive analysis to identify the underlying factors contributing to these fluctuations and to determine the necessary actions for strengthening the quality of democracy moving forward.

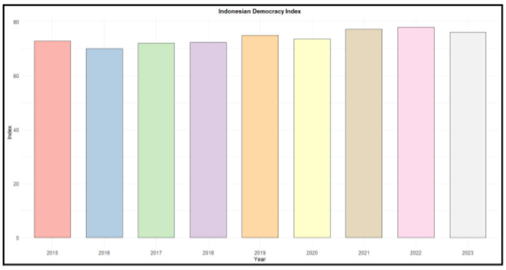

Fig. 2. The IDI from 2015 to 2023.

3.3 Multicollinearity check

Multicollinearity occurs when the Variance Inflation Factor (VIF) exceeds 10. Table 2 presents the VIF values for each predictor variable, allowing us to assess the degree of correlation among them. If any variable has a VIF greater than 10, it may need to be addressed through variable selection or transformation to improve model reliability.

Table 2. Multicollinearity Check Using VIF.

Predictor Variables	VIF
GDI	1.94
RGDP	1.38
PPP	2.26
ICTDI	3.16
GR	1.50
GEI	6.01
WPP	5.72

Based on Table 2, all predictor variables have VIF values below 10, indicating that there is no multicollinearity among the predictor variables. This ensures that the regression model can provide reliable estimates without significant distortion due to highly correlated independent variables.

3.4 Selection of optimal knot points

The optimal knot points are selected by identifying the model that minimizes the Generalized Cross-Validation (GCV) value. Table 3 presents the GCV values for models with one, two, and three knot points. The model that achieves the lowest GCV value is deemed the most optimal, as it strikes the best balance between model complexity and goodness of fit, thereby ensuring that the model is neither overfitting nor underfitting the data.

Table 3. Comparison of Minimum GCV Values.

Number of Knot Points	Minimum GCV Values
1	15.63
2	7.19
3	**6.16**

Based on Table 3, the optimal knot points are determined to be three knots. With these knots, the nonparametric truncated spline regression model is formulated as follows:

$\hat{y}_i = 39.35 + 0.41x_{1i} - 2.89x_{2i} + 18.02_x3i - 6.40x_{4i} + 98.24x_{5i} + 57.09x_{6i} - 153.65x_{7i} + 1.89(x_{1i} - 83.34)_+ + 0.03(x_{2i} - 85.04)_+ - 39.61(x_{3i} - 89.57)_+ - 1.10(x_{4i} - 4.14)_+ - 82.82(x_{5i} - 6.47)_+ + 28.54(x_{6i} - 12.70)_+ - 62.49(x_{7i} - 6.97) - 0.34(x_{1i} - 9.70)_+ + 0.13(x_{2i} - 16.96)_+ - 19.20(x_{3i} - 3.98)_+ - 0.73(x_{4i} - 4.51)_+ - 41.41(x_{5i} - 5.94)_+ - 10.83(x_{6i} - 0.27)_+ + 87.14(x_{7i} - 0.30)_+ - 1.33(x_{1i} - 0.36)_+ - 5.93(x_{2i} - 56.69)_+ - 12.80(x_{3i} - 60.10)_+ - 0.37(x_{4i} - 69.19)_+ + 81.73(x_{5i} - 5.42) + 75.41(x_{6i} - 9.30)_+ + 43.57(x_{7i} - 19.64)_+$

The truncated spline nonparametric regression model for the IDI data, using three knot points, has a coefficient of determination of 97.81%. This means that the predictor variables used in this study, is GDI (x_1), RGDP (x_2), PPP(x_3), ICTDI (x_4), GR (x_5), GEI(x_6), dan WPP(x_7) can explain 97.81% of the variation in IDI. While this high value may raise concerns about overfitting, the best model was selected based on the Generalized Cross-Validation (GCV) value, which is an effective method for preventing overfitting in nonparametric models. Thus, a high coefficient of determination reflects the model's ability to fit the data more accurately without indicating overfitting.

To further assess the predictive accuracy of the model, several error metrics were calculated: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The MSE value was found to be 4.998, indicating that, on average, the squared differences between the actual and predicted IDI values are relatively small, which suggests the model is effective at minimizing these errors. The RMSE, which is the square root of the MSE, was 2.236, providing an error measure in the same units as the IDI. This makes it more interpretable, showing that the model's predictions deviate by approximately 2.236 points from the actual IDI values. Additionally, the MAPE was calculated to be 2.179%, meaning that, on average, the model's predictions are 2.179% off from the true values. These error metrics collectively highlight that the model is performing well, with predictions being relatively close to the actual IDI scores, and demonstrate that the model is both accurate and reliable in capturing the relationship between the predictor variables and the IDI.

3.5 Model results with interval estimation

The interval estimation results for the multivariable truncated spline nonparametric regression curve are presented in Table 4.

Based on Table 4, it is observed that out of the 34 actual IDI data points, 8 provinces were accurately estimated. These provinces are Kepulauan Riau, DKI Jakarta, Bali, Central Kalimantan, North Sulawesi, Central Sulawesi, North Maluku, and Papua. Furthermore, Table 8 can be visualized as a comparison graph between the actual data and predicted values, as shown in Fig. 3.

Based on Fig. 3, it can be observed that the predicted data tends to follow the pattern of the actual data. Furthermore, a visualization will be created to illustrate this relationship. However, the point estimation results obtained fall within the interval range defined by the upper and lower bounds of the model. The IDI score interval analysis indicates the possibility that several provinces may experience significant improvements in their democratic performance, shifting from the moderate category to the high category, or vice versa. According to the Badan Pusat Statistik (BPS), an IDI score is considered high if it is above 80%, moderate if it falls between 60 and 80, and low if it is below 60. The IDI score interval analysis is presented in Table 5.

Based on Table 5, it is evident that 16 provinces have experienced a decline in their IDI classification. These provinces may undergo a transition in their democratic performance. This situation calls for a more in-depth analysis to identify strategies for improving their democracy performance category. Additionally, 18 provinces are estimated to see an improvement in their IDI scores. The primary objective of this

Table 4. Interval estimation results for the regression curve.

Province	Actual Curve	Point Estimation of the Curve	Interval Estimation of the Curve	
			Upper Bound	Lower Bound
Aceh	77.53	75.48	83.50	67.46
North Sumatera	80.34	78.86	84.74	72.98
West Sumatera	76.27	76.25	91.23	61.27
Riau	75.55	78.08	86.56	69.60
Jambi	74.47	77.89	82.77	73.01
South Sumatera	78.17	80.25	87.00	73.49
Bengkulu	74.26	75.73	84.12	67.35
Lampung	78.37	77.67	84.08	71.26
Kep. Bangka Belitung	75.95	77.66	88.91	66.41
Kep. Riau	77.66	77.66	92.64	62.68
DKI Jakarta	84.57	84.57	99.55	69.58
West Java	83.04	79.43	87.70	71.16
Central Java	80.87	79.60	85.26	73.94
DI Yogyakarta	83.88	83.90	98.89	68.92
East Java	82.01	79.20	85.75	72.65
Banten	75.83	79.66	87.44	71.89
Bali	85.13	85.13	100.11	70.14
West Nusa Tenggara	70.03	68.99	82.30	55.69
East Nusa Tenggara	77.39	74.80	85.52	64.08
West Kalimantan	81.69	78.08	87.76	68.41
Central Kalimantan	77.57	77.57	92.55	62.58
South Kalimantan	80.44	80.49	89.14	71.85
East Kalimantan	82.28	80.26	91.03	69.49
North Kalimantan	80.47	76.84	85.04	68.63
North Sulawesi	76.27	76.27	91.25	61.29
Central Sulawesi	79.13	79.13	94.11	64.15
South Sulawesi	76.43	80.10	87.92	72.27
Southeast Sulawesi	77.03	79.31	85.34	73.29
Gorontalo	79.71	76.06	86.62	65.51
West Sulawesi	74.55	77.73	88.38	67.08
Maluku	69.35	69.32	84.31	54.34
North Maluku	65.57	65.57	80.55	50.59
West Papua	65.55	69.81	81.99	57.63
Papua	67.64	67.64	82.62	52.66

method is to provide more realistic and cautious predictions regarding the development of democracy in each province. Interval estimates provide not only a single predicted value (point), but also a range of values (lower and upper bounds) that reflect the potential IDI score for a region. This range demonstrates uncertainty, social variability, and possible changes resulting from political, economic, and social dynamics in the area.

These findings demonstrate that democratic dynamics at the provincial level are not linear and are always influenced by various local social, economic, and policy factors. Any decline in the IDI classification is not merely a statistical figure, but rather a reflection of fundamental challenges faced by society, such as reduced space for public participation, weakened human rights protections, or the suppression of the voices of vulnerable groups in decision-making. Conversely, for provinces predicted to experience an

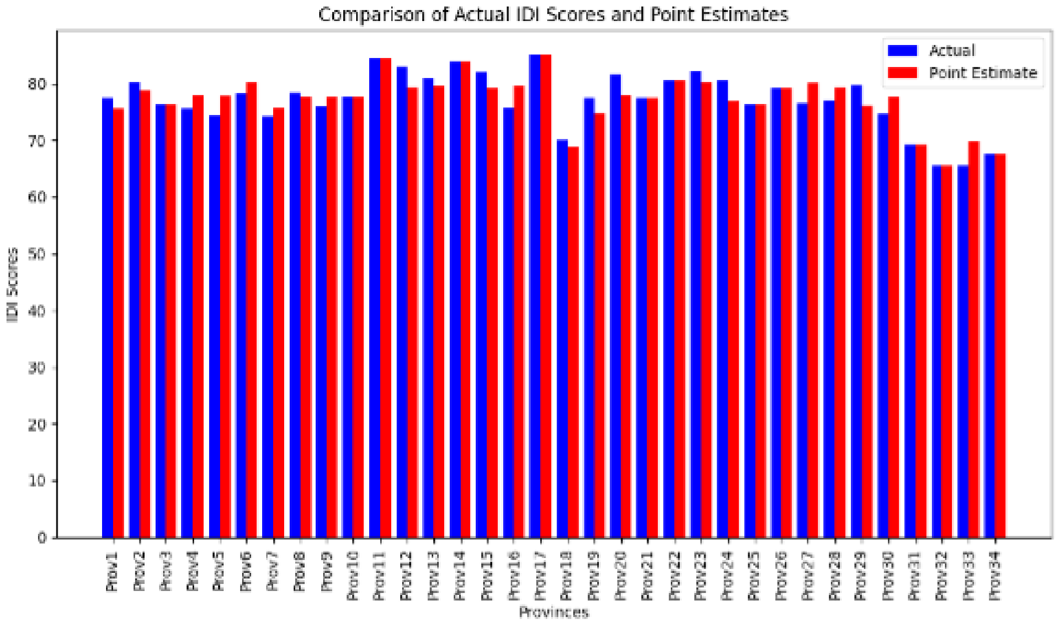

Fig. 3. Graph of actual vs. predicted data.

increase in their IDI scores, opportunities to build more inclusive and equitable democratic practices are greater. The interval data generated from this analysis provides a starting point for stakeholders to design interventions based on real needs: democracy literacy programs, strengthening spaces for citizen dialogue, and ensuring government transparency. Thus, the benefits of democracy are not only measured in terms of numerical achievements but are truly felt in the form of access, justice, and respect for the dignity of every individual in community life.

This analysis underscores the need for targeted interventions and strategies tailored to the specific needs of each province. For provinces with potential for improvement, such as South Sumatera or East Kalimantan, policy measures aimed at enhancing public participation, strengthening democratic institutions, and improving civil liberties could help shift the IDI score into the high category. Conversely, provinces showing potential decline, like West Nusa Tenggara and Papua, may require focused attention to address challenges such as economic inequality, political instability, or weak democratic institutions to prevent their democratic scores from declining.

4. Conclusion

From the analysis and discussion, this nonparametric regression with a truncated spline model of 3 knots has been sufficiently precise for capturing the variation in IDI performance across provinces, represented by the high coefficient of determination at 97.81%. The interval estimation results indicate that 16 provinces are estimated to have their IDI performance decrease, but the estimates suggest an improvement in IDI performance of 18 provinces. These results highlight the importance of public participation in decision-

Table 5. IDI criteria for point and interval estimation results.

Provinces	IDI Actual	IDI Criteria Estimation Point Results	IDI Criteria Estimation Interval Results	
			Upper Bound	Lower Bound
Aceh	Medium	Medium	High	Medium
North Sumatera	High	Medium	High	Medium
West Sumatera	Medium	Medium	High	Medium
Riau	Medium	Medium	High	Medium
Jambi	Medium	Medium	High	Medium
South Sumatera	Medium	High	High	Medium
Bengkulu	Medium	Medium	High	Medium
Lampung	Medium	Medium	High	Medium
Kep. Bangka Belitung	Medium	Medium	High	Medium
Kep. Riau	Medium	Medium	High	Medium
DKI Jakarta	High	High	High	Medium
West Java	High	Medium	High	Medium
Central Java	High	Medium	High	Medium
DI Yogyakarta	High	High	High	Medium
East Java	High	Medium	High	Medium
Banten	Medium	Medium	High	Medium
Bali	High	High	High	Medium
West Nusa Tenggara	Medium	Medium	High	Low
East Nusa Tenggara	Medium	Medium	High	Medium
West Kalimantan	High	Medium	High	Medium
Central Kalimantan	Medium	Medium	High	Medium
South Kalimantan	High	High	High	Medium
East Kalimantan	High	High	High	Medium
North Kalimantan	High	Medium	High	Medium
North Sulawesi	Medium	Medium	High	Medium
Central Sulawesi	Medium	Medium	High	Medium
South Sulawesi	Medium	High	High	Medium
Southeast Sulawesi	Medium	Medium	High	Medium
Gorontalo	Medium	Medium	High	Medium
West Sulawesi	Medium	Medium	High	Medium
Maluku	Medium	Medium	High	Low
North Maluku	Medium	Medium	High	Low
West Papua	Medium	Medium	High	Low
Papua	Medium	Medium	High	Low

making for improving the quality of democracy in Indonesia.

For future research, it is worth extending the methodology with other non-parametric models that could better fit the complex relationships between the variables. One avenue worth exploring would be to take the approach of using estimators based on Fourier series or kernels in non-parametric regression, which may be more flexible at capturing the nonlinear structure present in our data.

Moreover, to see whether one can gain any additional insights about the strengths/weaknesses of each approach in modeling democratic performance, we could compare our results against those with more typical ML models like random forests or SVMs. The methods could further enrich our understanding of what drives the IDI at the provincial level and make an essential theoretical and practical contribution to the development of democracy in Indonesia.

References

[1] Dani ATR, Ratnasari V, Budiantara IN. Optimal Knots Point and Bandwidth Selection in Modeling Mixed Estimator Nonparametric Regression. IOP Conf Ser Mater Sci Eng. 2021;1115(1):012020.

[2] Nurcahyani H, Budiantara IN, Zain I. Nonparametric Truncated Spline Regression on Modelling. AIP Conf Proc. 2023;020073(May).

[3] Sriliana I, Budiantara IN, Ratnasari V. A Truncated Spline and Local Linear Mixed Estimator in Nonparametric Regression for Longitudinal Data and Its Application. Symmetry. 2022;14(12).

[4] Dani ATR, Putra FB, Ni'matuzzaroh L, Ratnasari V, Budiantara IN. BTS Buku Tentang Statistika Regresi parametrik dan non Parametrik Teori dan Aplikasi dengan Software R. Jawa Barat: Perkumpulan Rumah Cemerlang Indonesia; 2024.

[5] Chu CK, Marron JS. Choosing a Kernel Regression Estimator. Statistical Science. 1991;6(4):404–436.

[6] Mariati NPAM, Budiantara IN, Ratnasari V. Combination Estimation of Smoothing Spline and Fourier Series in Nonparametric Regression. J Math. 2020;2020.

[7] Cattaneo MD, Jansson M, Ma X. Simple Local Polynomial Density Estimators. J Am Stat Assoc. 2020;115(531):1449–55.

[8] Bilodeau M. Fourier smoother and additive models. Can J Stat. 1992;20(3):257–69.

[9] Amato U, Antoniadis A, De Feis I, Gijbels I. Wavelet-based robust estimation and variable selection in nonparametric additive models. Stat Comput. 2022;32(1).

[10] Ampulembang AP, Otok BW, Rumiati AT, Budiasih. Bi-responses nonparametric regression model using MARS and its properties. Appl Math Sci. 2015;9(29–32):1417–27.

[11] Tripena A, Prabowo A, Lianawati Y, Bon AT. Estimated spline in nonparametric regression with a generalized cross validation and unbiased risk approach. Proc Int Conf Ind Eng Oper Manag. 2021:3788–98.

[12] Chamidah N, Zaman B, Muniroh L, Lestari B. Designing local standard growth charts of children in East Java province using a local linear estimator. Int J Innov Creat Change. 2020;13(1):45–67.

[13] Eslami MM. Spline and Wavelet Smoothing Techniques for Functional Data. 2023:36–41.

[14] Adrianingsih NY, Budiantara IN, Purnomo JDT. Modeling with Mixed Kernel, Spline Truncated and Fourier Series on Human Development Index in East Java. IOP Conf Ser Mater Sci Eng. 2021;1115(1):012024.

[15] Widyastuti DA, Fernandes AAR, Pramoedyo H. Spline estimation method in nonparametric regression using truncated spline approach. J Phys Conf Ser. 2021;1872(1):012027.

[16] Devi AR, Pratama RFW, Suparti. Comparison of generalized cross validation and unbiased risk method for selecting optimal knot in spline truncated. J Phys Conf Ser. 2019;1217(1):012094.

[17] Maharani M, Saputro DRS. Generalized Cross Validation (GCV) in Smoothing Spline Nonparametric Regression Models. IOP Conf Ser Earth Environ Sci. 2021;1808(1):012053.

[18] Saputro DRS, Demu KR, Widyaningsih P. Nonparametric truncated spline regression model on the data of human development index (HDI) in Indonesia. J Phys Conf Ser. 2018;1028(1):012219.

[19] Chamidah N, Lestari B, Massaid A, Saifudin T. Estimating mean arterial pressure affected by stress scores using spline nonparametric regression model

approach. Commun Math Biol Neurosci. 2020;2020:1–12.

[20] Sriliana I, Budiantara IN, Ratnasari V. The performance of mixed truncated spline-local linear nonparametric regression model for longitudinal data. MethodsX. 2024;12(March).

[21] Budiantara IN, Ratna M, Zain I, Wibowo W. Modeling the Percentage of Poor People in Indonesia Using Spline Nonparametric Regression Approach. Int J Basic Appl Sci IJBAS-IJENS. 2012;12(6):119–24.

[22] Setiawan RNS, Budiantara IN, Ratnasari V. Application of Confidence Intervals for Parameters of Nonparametric Spline Truncated Regression on Index Development Gender in East Java. IPTEK J Sci. 2017;2(3):49–55.

[23] Nidhomuddin, Chamidah N, Kurniawan A. Confidence Interval of the Parameter on Multipredictor Biresponse Longitudinal Data Analysis Using Local Linear Estimator for Modeling of Case Increase and Case Fatality Rates Covid-19 in Indonesia: a Theoretical Discussion. Commun Math Biol Neurosci. 2022;2022:1–12.

[24] Lestari B, Chamidah N, Budiantara IN, Aydin D. Determining confidence interval and asymptotic distribution for parameters of multiresponse semiparametric regression model using smoothing spline estimator. J King Saud Univ Sci. 2023;35(5):1–7.

[25] Budiantara IN, Ratnasari V, Permatasari EO, Prawanti DD. Shortest confidence interval of parameter semi parametric regression model using spline truncated for longitudinal data. AIP Conf Proc. 2019;2194.

[26] Mao W, Zhao LH. Free-knot polynomial splines with confidence intervals. J R Stat Soc B Stat Methodol. 2003;65(4):901–19.

[27] Suprapto. Interval Estimation for Indonesia Democracy Index (IDI) Model Using Multivariables Spline Truncated. Int J Sci Eng Res. 2018;9(9):1740–50.

[28] Setyawati M, Chamidah N, Kurniawan A. Confidence Interval of Parameters in Multiresponse Multipredictor Semiparametric Regression Model for Longitudinal Data Based on Truncated Spline Estimator. Commun Math Biol Neurosci. 2022;2022:1–18.

[29] Lestari B, Chamidah N, Budiantara IN, Aydin D. Determining confidence interval and asymptotic distribution for parameters of multiresponse semiparametric regression model using smoothing spline estimator. J King Saud Univ Sci. 2023;35(5):102664.

[30] Eubank RL. Nonparametric Regression and Spline Smoothing. New York: CRC Press; 1999.

[31] Hardle W. Applied Nonparametric Regression. 1994.

[32] Wahba G. Spline Models for Observational Data. Pennsylvania: SIAM; 1990.

[33] Ratnasari V, Budiantara IN, Andrea TRD. Nonparametric Regression Mixed Estimators of Truncated Spline and Gaussian Kernel based on Cross-Validation (CV), Generalized Cross-Validation (GCV), and Unbiased Risk (UBR) Methods. Int J Adv Sci Eng Inf Technol. 2021;11(6):2400–6.

[34] Nidhomuddin, Chamidah N, Kurniawan A. Confidence Interval Of The Parameter On Multipredictor Biresponse Longitudinal Data Analysis Using Local Linear Estimator For Modeling Of Case Increase And Case Fatality Rates Covid-19 In Indonesia: A Theoretical Discussion. Commun Math Biol Neurosci. 2022;2022(23):1–12.

[35] Wang L, Politis DN. Asymptotic validity of bootstrap confidence intervals in nonparametric regression without an additive model. Electron J Stat. 2021;15(1):392–426.