JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e-ISSN: 2614-8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 23, No. 1, September 2026, pp. 203-214
DOI: 10.20956/yee72c02

CO_2 Emission Modeling in Asian Countries using a Truncated Spline Nonparametric Regression Approach on Panel Data

Dita Amelia[1]*, Anisah Nabilah Ghasani[2], Salsabilla Rusydah Putri[3]
[1,2,3] Department of Mathematics, Universitas Airlangga, Surabaya, Indonesia
***Email**: [1]dita.amelia@fst.unair.ac.id, [2]anisah.nabilah.ghasani-2022@fst.unair.ac.id, [3]salsabilla.rusydah.putri-2022@fst.unair.ac.id*
**Corresponding author*
Received: 22 June 2026, revised: 28 August 2026, accepted: 1 September 2026

Abstract

Climate change, driven by increasing carbon dioxide (CO_2) emissions, has become a major global challenge, with Asian countries contributing nearly 50% of total worldwide emissions. This study aims to model the factors affecting CO_2 emissions in Asian countries using a truncated spline nonparametric regression approach with panel data from 2020 to 2023. The predictor variables considered are population density, primary energy consumption, and urban population, while CO_2 emissions serve as the response variable. Optimal knot points were determined using the Generalized Cross Validation (GCV) criterion. The results indicate that the model with three knots provides the best performance, yielding a GCV value of 0.00004, a Mean Square Error (MSE) of 0.005248, and a coefficient of determination (R^2) of 99.99206%. The findings reveal that differences in energy consumption patterns, industrialization levels, and dependence on fossil fuels contribute significantly to variations in CO_2 emissions among countries. The thematic map analysis further shows that high emissions are not always associated with high population density or urban population. Overall, the selected predictor variables explain almost all variations in CO_2 emissions across Asian countries. These findings provide a more comprehensive understanding of emission dynamics and offer valuable insights for the development of sustainable energy and environmental policies in the Asian region.

Keywords: CO_2 Emissions; Spline Truncated; Population Density; Primary Energy Consumption; Urban Population.

1. INTRODUCTION

Climate change is now one of the most pressing global challenges, with widespread impacts on ecosystems and human life. One of the main causes is carbon dioxide (CO_2) emissions, which continue to increase in line with population growth and fossil fuel consumption. In recent years, global CO_2 output has shown an upward trend. Although it fell sharply in 2020 due to the COVID-19 pandemic, emissions rebounded in subsequent years, reaching a record high in 2023. In 2020, global carbon output fell to 31.5 gigatons due to restrictions on economic activity during the COVID-19 pandemic [7] . However, this decline was only temporary. In 2021, carbon output increased to

36.3 gigatons, up about 6% from the previous year, as fossil fuel-based economic activity recovered [8] . This trend continued in 2022 when global emissions reached their highest level of around 36.8 gigatons, an increase of about 1% from 2021 [9] . According to the [10] , in 2023 global CO_2 output again reached a record high of 37.4 gigatons, an increase of 1.1% from the previous year. The distribution of CO_2 emissions around the world is uneven, with Asia contributing around 50% of total global CO_2 output, up from around 40% in 2015 [3] . This dominance means that Asia must play an important role in achieving Sustainable Development Goals (SDGs) 13, Climate Action, which emphasizes collective action to tackle climate change.

Asia's dominant role in global emissions is evident from the contribution of several of the world's largest CO_2-producing countries. According to [14] , the three countries with the largest CO_2 contributions in 2023 were China, the United States, and India, which contributed more than 53% of global energy pollution. China's CO_2 emissions rise by 565 million tons to 12.6 gigatons, an increase of 4.7%, while India's emissions increased by 190 million tons to 2.8 gigatons as a result of post-pandemic economic recovery and a decline in hydropower supply [11] . On the other hand, the United States managed to reduce CO_2 emissions by 158.5 million tons to 4.64 gigatons due to the shift from coal to natural gas and increased use of renewable energy [14] . Not only China and India, several other countries in Asia are also among the top 10 highest CO_2 producers in the world. Japan ranks 5th with emissions of 1.01 gigatons, followed by Indonesia in 6th place with 701.4 million tons, Iran in 7th place with 683.6 million tons, Saudi Arabia in 8th place with 620.4 million tons, and South Korea in 9th place with 571.2 million tons [2] . This data shows Asia's significant contribution to global emissions, which is influenced by differences in energy characteristics and economic structures in each country.

In an effort to model the nonlinear relationship between factors that influence CO_2 emissions, the truncated spline nonparametric regression approach is the right choice to use. Spline is a nonparametric method that has high flexibility because it is able to divide the observation domain into several sub-intervals through knot points, then form a polynomial basis function on each interval. This approach allows the model to capture complex relationship patterns. To achieve a balance between model complexity and prediction error rates, the Generalized Cross Validation (GCV) criterion is used as the basis for selecting the optimal model. Thus, truncated spline regression is not only capable of producing estimates that are adaptive to data characteristics, but also maintains the validity of inferences through a measurable knot selection mechanism [4] . Panel data analysis was chosen in this study because it is capable of studying a group of subjects by considering both dimensions, namely data and time, simultaneously. Each cross-section unit is observed repeatedly over a certain period of time, thus producing more complete information compared to the separate use of cross-section or time series data [1] . By combining the truncated spline nonparametric regression method and panel data, this study aims to model CO_2 emissions in Asian countries more comprehensively. This approach is expected to provide a clearer picture of the effects of population density, primary energy consumption, and urban population on carbon dioxide emissions, thereby providing a solid basis for the formulation of environmental and sustainable development policies.

2. METHODS

2.1 Data Construction

This study uses secondary data obtained from Our World in Data (https://ourworldindata.org/), covering 48 Asian countries observed annually from 2020 to 2023. The response variable (Y) is CO_2 emissions (metric tons per capita), while the predictor variables consist of population density (X_1), primary energy consumption (X_2), and urban population (X_3). Panel data analysis is applied to accommodate both cross-sectional and time-series dimensions, allowing more efficient estimation and control over unobserved heterogeneity between countries.

2.2 Nonparametric Truncated Spline Regression

To model the nonlinear relationship between CO_2 emissions and its predictors, a nonparametric truncated spline regression is employed. Unlike parametric regression, nonparametric regression does not assume a specific functional form, allowing greater flexibility in capturing data patterns (Eubank, 1999). The truncated spline model divides the data range into several subintervals connected at knot points, which control curve smoothness [12] . The general form of a truncated spline function with one predictor variable is given in Equation (2.1):

$$f(x_i) = \sum_{r=0}^{q} \delta_r x_i^r + \sum_{l=1}^{h} \delta_{q+l}(x_i - \theta_l)_+^q \tag{2.1}$$

The truncated component $(x_i - \theta_l)_+^q$ is defined as shown in Equation (2.2):

$$(x_i - \theta_l)_+^q = \begin{cases} (x_i - \theta_l)^q & , x_i \geq \theta_l \\ 0 & , x_i < \theta_l \end{cases} \tag{2.2}$$

where $f(x_i)$ is the truncated spline regression function, x_i is the predictor variable, q is the degree of the polynomial, θ_l is the l^{th} knot point, and δ_r, δ_{q+l} are the estimated regression parameters.

If there is more than one predictor variable, the truncated spline nonparametric regression model can be written as Equation (2.3):

$$y_i = \delta_0 + \sum_{k=1}^{p}\sum_{r=1}^{q} \delta_{kr} x_{ki}^r + \sum_{k=1}^{p}\sum_{l=1}^{h} \delta_{k(q+l)}(x_{ki} - \theta_{lk})_+^q + \varepsilon_i \tag{2.3}$$

where ε_i is a random error term assumed to have mean zero and constant variance.

2.3 Extension to Panel Data Structure

In panel data, given $i = 1,2,\dots,m$ individuals observed repeatedly over a certain period of time $j = 1,2,\dots,n$, the q^{th}-order spline function with h knots and one predictor variable can be expressed as shown ini Equation (2.4) [13] .

$$y_{ij} = f\left(x_{ij}\right) + \varepsilon_{ij} \tag{2.4}$$

The truncated function for panel data follows the structure shown in Equation (2.5):

$$\left(x_{ij} - \theta_{(q+l)i}\right)_+^q = \begin{cases} \left(x_{ij} - \theta_{(q+l)i}\right)^q & , x_{ij} \geq \theta_{(q+l)i} \\ 0 & , x_{ij} < \theta_{(q+l)i} \end{cases} \tag{2.5}$$

If there is more than one predictor variable, the equation can be extended as represented in Equation (2.6):

$$\begin{aligned} y_{ij} &= \sum_{k=1}^{p} f_k\left(x_{ij}\right) + \varepsilon_{ij} \\ &= \sum_{k=1}^{p}\left(\sum_{r=0}^{q} \delta_{kir}\, x_{kij}^r + \sum_{l=1}^{h} \delta_{(q+l)ki}\left(x_{kij} - \theta_{(q+l)ki}\right)_+^q\right) + \varepsilon_{ij} \end{aligned} \tag{2.6}$$

When written in matrix notation, it can be expressed as shown in Equation (2.7):

$$\boldsymbol{y} = \boldsymbol{X_\theta \delta} + \boldsymbol{\varepsilon}$$

$$\begin{pmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \vdots \\ \boldsymbol{y_m} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X_1(\theta_1)} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X_2(\theta_2)} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{X_m(\theta_m)} \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_m \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \vdots \\ \boldsymbol{\varepsilon_m} \end{pmatrix} \tag{2.7}$$

with

$$\boldsymbol{y_1} = (y_{11}, y_{12}, \dots, y_{1n})'$$
$$\vdots$$
$$\boldsymbol{y_m} = (y_{11}, y_{12}, \dots, y_{mn})'$$

In the first-order truncated spline model, the design matrix formed using p predictors and h knots can be written as follows.

$$\boldsymbol{X_1(\theta_1)} = \begin{bmatrix} 1 & X_{111} & X_{211} & \cdots & X_{p11} & (X_{111}-\theta_{111}) & \cdots & (X_{111}-\theta_{11h}) & \cdots & (X_{p11}-\theta_{p11}) & \cdots & (X_{p11}-\theta_{p1h}) \\ 1 & X_{112} & X_{212} & \cdots & X_{p12} & (X_{112}-\theta_{111}) & \cdots & (X_{112}-\theta_{11h}) & \cdots & (X_{p12}-\theta_{p11}) & \cdots & (X_{p12}-\theta_{p1h}) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & X_{11n} & X_{21n} & \cdots & X_{p1n} & (X_{11n}-\theta_{111}) & \cdots & (X_{11n}-\theta_{11h}) & \cdots & (X_{p1n}-\theta_{p11}) & \cdots & (X_{p1n}-\theta_{p1h}) \end{bmatrix}$$
$$\vdots$$
$$\boldsymbol{X_m(\theta_m)} = \begin{bmatrix} 1 & X_{1m1} & X_{2m1} & \cdots & X_{pm1} & (X_{1m1}-\theta_{1m1}) & \cdots & (X_{1m1}-\theta_{1mh}) & \cdots & (X_{pm1}-\theta_{pm1}) & \cdots & (X_{pm1}-\theta_{p1h}) \\ 1 & X_{1m2} & X_{2m2} & \cdots & X_{pm2} & (X_{1m2}-\theta_{1m1}) & \cdots & (X_{1m2}-\theta_{1mh}) & \cdots & (X_{pm2}-\theta_{pm1}) & \cdots & (X_{pm2}-\theta_{p1h}) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & X_{1mn} & X_{2mn} & \cdots & X_{pmn} & (X_{1mn}-\theta_{1m1}) & \cdots & (X_{1mn}-\theta_{1mh}) & \cdots & (X_{pmn}-\theta_{pm1}) & \cdots & (X_{p1n}-\theta_{p1h}) \end{bmatrix}$$

And the vector $\boldsymbol{\delta}$ contains

$$\boldsymbol{\delta_1} = \left(\delta_0, \delta_{11}, \delta_{21}, \dots, \delta_{p1}, \delta_{(p+1)1}, \dots, \delta_{(p+k)1}\right)'$$
$$\vdots$$
$$\boldsymbol{\delta_m} = \left(\delta_0, \delta_{1m}, \delta_{2m}, \dots, \delta_{pm}, \delta_{(p+1)m}, \dots, \delta_{(p+k)m}\right)'$$

To solve the matrix above, the Weighted Least Squares method is used by minimizing the weighted sum of squares error as follows in Equation (2.8):

$$\boldsymbol{G(\delta)} = \boldsymbol{\varepsilon' W \varepsilon} = (\boldsymbol{y} - \boldsymbol{X_\theta \delta})' \boldsymbol{W} (\boldsymbol{y} - \boldsymbol{X_\theta \delta})$$

where $\boldsymbol{G(\delta)}$ is the objective function to be minimized, $\boldsymbol{\varepsilon}$ is the error vector, $\boldsymbol{y}$ is the response vector, $\boldsymbol{X_\theta}$ is the design matrix, $\boldsymbol{\delta}$ is the parameter vector, and $\boldsymbol{W}$ is the weighting matrix.

Expanding $\boldsymbol{G(\delta)}$**,** the objective function can be written as:

$$\boldsymbol{G(\delta)} = \boldsymbol{y'Wy - y'WX_\theta\delta - \delta'X'_\theta Wy + \delta'X'_\theta WX_\theta\delta}$$
$$\frac{\partial\, \boldsymbol{G(\delta)}}{\partial \boldsymbol{\delta}} = \boldsymbol{0}$$
$$\boldsymbol{0} = \boldsymbol{-2X'_\theta Wy + 2X'_\theta WX_\theta\delta}$$
$$\boldsymbol{X'_\theta Wy} = \boldsymbol{X'_\theta WX_\theta\delta}$$
$$\boldsymbol{\hat{\delta}} = (\boldsymbol{X'_\theta WX_\theta})^{-1}\boldsymbol{X'_\theta Wy} \tag{2.8}$$

Based on the estimation results of $\boldsymbol{\hat{\delta}}$ above, the following equation is obtained through Equation (2.9):

$$\boldsymbol{\hat{y}} = \boldsymbol{X_\theta\hat{\delta}}$$
$$= \boldsymbol{X_\theta}(\boldsymbol{X'_\theta WX_\theta})^{-1}\boldsymbol{X'_\theta Wy} = \boldsymbol{A(h)y}$$

with

$$\boldsymbol{A(h)} = \boldsymbol{X_\theta}(\boldsymbol{X'_\theta WX_\theta})^{-1}\boldsymbol{X'_\theta W} \tag{2.9}$$

To obtain the properties of the estimator $\boldsymbol{\hat{\delta}}$, the estimated value and its covariance are reduced as follows.

$$\boldsymbol{\hat{\delta}} = (\boldsymbol{X'_\theta WX_\theta})^{-1}\boldsymbol{X'_\theta Wy}$$

The expected value of the estimator $\widehat{\boldsymbol{\delta}}$ can be obtained by taking the expectation of its estimation form to show that it is an unbiased estimator, resulting in:

$$\begin{aligned}
\boldsymbol{E}(\widehat{\boldsymbol{\delta}}) &= \boldsymbol{E}\left[(\boldsymbol{X}_{\boldsymbol{\theta}}'\boldsymbol{W}\boldsymbol{X}_{\boldsymbol{\theta}})^{-1}\boldsymbol{X}_{\boldsymbol{\theta}}'\boldsymbol{W}\boldsymbol{y}\right] \\
&= (\boldsymbol{X}_{\boldsymbol{\theta}}'\boldsymbol{W}\boldsymbol{X}_{\boldsymbol{\theta}})^{-1}\boldsymbol{X}_{\boldsymbol{\theta}}'\boldsymbol{W}\,\boldsymbol{E}(\boldsymbol{y}) \\
&= (\boldsymbol{X}_{\boldsymbol{\theta}}'\boldsymbol{W}\boldsymbol{X}_{\boldsymbol{\theta}})^{-1}\boldsymbol{X}_{\boldsymbol{\theta}}'\boldsymbol{W}\boldsymbol{X}_{\boldsymbol{\theta}}\boldsymbol{\delta} \\
&= \boldsymbol{I}\boldsymbol{\delta} \\
\boldsymbol{E}(\widehat{\boldsymbol{\delta}}) &= \boldsymbol{\delta}
\end{aligned}$$

2.4 Model Selection Using Generalized Cross Validation (GCV)

To determine the optimal number and location of knots, the Generalized Cross Validation (GCV) criterion is applied [15] .
The GCV method can be calculated using the following formula in Equation (2.10):

$$GCV = \frac{MSE(h)}{\left[\frac{1}{mn}tr\left(\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{h})\right)\right]^2} \tag{2.10}$$

where the Mean Square Error (MSE) is calculated using Equation (2.11):

$$MSE(h) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(y_{ij}-\hat{y}_{ij}\right)^2 \tag{2.11}$$

Where m is the number of individuals and n is the number of observation periods. Then the optimal knot value is the smallest GCV value.

2.5 Model Evaluation

The coefficient of determination or R-Square is a coefficient that shows how much the variation in the independent variable explains the dependent variable [5]. The value of the coefficient of determination ranges from 0 to 1. If the R^2 value is close to 1, then the independent variable provides almost all the information needed to predict the dependent variable. However, if the R^2 value is smaller, it means that the ability of the independent variable to explain the dependent variable is quite limited [6]. The R^2 value can be formulated based on Equation (2.12):

$$R^2 = 1 - \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(y_{ij}-\hat{y}_{ij}\right)^2}{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{i}\right)^2} \tag{2.12}$$

where y_{ij} is the actual value of the response variable for unit i in time period j, $\hat{y}_{ij}$ is the predicted value from the model, and $\bar{y}_i$ is the average value of the response variable for unit i.

3. RESULT AND DISCUSSION

3.1 Descriptive Statistic

Statistical descriptive analysis was performed on all research variables, the results of which are presented in Table 1 to provide an overview of the data.

Table 1. Descriptive Statistics

Variable	Mean	StDev	Min.	Max.
CO_2 Emission (Y)	6.381	8.153	0.250	38.840
Population Density (X_1)	431.9	1159.6	2.11	8062.7
Primary Energy Consumption (X_2)	36395	47576	863	208042

Variable	Mean	StDev	Min.	Max.
Urban Population (X_3)	60.50	23.59	18.71	100

Based on Table 1, the average CO_2 emissions from 48 countries is 6.381 metric tons, which falls into the moderate category (5-10 tons per capita). The lowest value is in Yemen in 2023 (0.25 metric tons), a condition caused by prolonged conflict that has reduced industrial activity and energy consumption. Conversely, the highest value was found in Qatar in 2023 (38.84 metric tons), which is in line with high fossil energy consumption and economic dependence on the oil and gas sector. The average population density is 431.9 people/km2, which is classified as low (<500 people/km2). The minimum value was recorded in Mongolia in 2020 (2.11 people/km2), which has a very large area but a sparse population, while the maximum value was in Singapore in 2023 (8,062.7 people/km2) as a densely populated country. The average primary energy consumption is 36,395 terawatt-hours, with the lowest value in Yemen in 2021 (863 terawatt-hours) due to damage to energy infrastructure, while the highest value is in Qatar in 2023 (208,042 terawatt-hours) due to the dominance of the fossil fuel sector. Meanwhile, the proportion of urban population in Sri Lanka in 2020 (18.71%) is still dominated by rural residents, while the maximum value of 100% is found in Kuwait and Singapore in 2020-2023 as fully urbanized countries.

3.2 Determination of the Optimum Knot Point

The best truncated spline nonparametric regression model was determined based on the minimum GCV value. This study compared models with one and two first knot points. The GCV value was analyzed using X_1 (Population Density), X_2 (Primary Energy Consumption), and X_3 (Urban Population) as predictors of CO_2 emissions.

Table 2. Output Based on Minimum GCV 1 Knot

Country	Knot			GCV
	X_1	X_2	X_3	
Afghanistan	62.92897	1164.487	26.77662	0.00671
Armenia	102.9386	17335.18	63.66555	
Azerbaijan	124.5555	20571.13	57.37355	
Bahrain	1967.97	162722	89.80641	
Bangladesh	1310.403	2976.668	40.07714	
Bhutan	20.54586	18712.53	43.99931	
⋮	⋮	⋮	⋮	
Yemen	73.54621	909.1252	39.49945	

Table 3. Output Based on Minimum GCV 2 Knot

Country	Knot			GCV
	X_1	X_2	X_3	
Afghanistan	60.2786	1012.22	26.1198	0.00159968
	63.4338	1193.49	26.9017	
Armenia	101.085	15708	63.3571	
	103.292	17645.1	63.7243	
Azerbaijan	123.361	18463.9	56.5191	

Country	Knot X_1	Knot X_2	Knot X_3	GCV
	124.783	20972.5	57.5363	
Bahrain	1888.59	153416	89.5436	
	1983.09	164495	89.8565	
Bangladesh	1281.65	2779.01	38.4145	
	1315.88	3014.32	40.3938	
Bhutan	20.2345	12663.3	42.5264	
	20.6052	19864.8	44.2799	
⋮	⋮	⋮	⋮	
Yemen	69.0783	869.037	38.1069	
	74.3972	916.761	39.7647	

Table 4. Output Based on Minimum GCV 3 Knot

Country	Knot X_1	Knot X_2	Knot X_3	GCV
Afghanistan	60.0262	997.721	26.0573	0,00004
	60.7834	1041.22	26.2449	
	63.1814	1178.99	26.8392	
Armenia	100.908	15553	63.3277	
	101.438	16017.9	63.4158	
	103.115	17490.2	63.6949	
Azerbaijan	123.247	18263.2	56.4377	
	123.588	18865.3	56.6818	
	124.669	20771.8	57.4549	
Bahrain	1881.03	152530	89.5185	
	1903.71	155189	89.5936	
	1975.53	163608	89.8314	
Bangladesh	1278.91	2760.18	38.2562	
	1287.13	2816.66	38.7312	
	1313.14	2995.49	40.2355	
Bhutan	20.2048	12087.1	42.3861	
	20.2938	13815.5	42.807	
	20.5755	19288.6	44.1396	
⋮	⋮	⋮	⋮	
Yemen	68.6528	865.219	37.9743	

Country	Knot			GCV
	X_1	X_2	X_3	
	69.9293	876.673	38.3722	
	73.9717	912.943	39.6321	

Based on the results in Table 2, Table 3, and Table 4, a summary comparison is shown to determine the best knot model by considering the largest R^2 and minimum GCV, which are presented in Table 5.

Table 5. Comparison of the Best Knot Models

Knot	Trial	GCV	MSE	R^2
1	28	0.006697127	0.02678851	99.95949
2	378	0.001597693	0.00933097	99.98589
3	3276	0.00004	0.005248153	99.99206

Table 5 shows that the model with three knots provides the best results. Of the 3276 alternatives in the three-knot model, the minimum GCV obtained was 0.00004 with an MSE of 0.005248 and the highest R^2 value of 99.99206. This indicates that the model with three knots has the best prediction accuracy compared to models with one or two knots.

3.3 Parameter Estimation

As an illustration, Bhutan was used in the interpretation of the truncated spline regression model with three knots, due to the pattern of CO_2 emissions changes during the period 2020 to 2023. The parameter estimates at the best knot points can be seen in Table 6.

Table 6. Bhutan Country Parameter Model Estimates

δ_{060}	δ_{161}	δ_{261}	δ_{361}	δ_{162}	δ_{163}	δ_{164}
1.31×10^{-8}	2.71×10^{-7}	0.000137	5.83×10^{-7}	6.72×10^{-9}	4.73×10^{-9}	3.44×10^{-9}

δ_{262}	δ_{263}	δ_{264}	δ_{362}	δ_{363}	δ_{364}
-2.33×10^{-5}	-6.19×10^{-5}	-0.00011	2.80×10^{-8}	1.86×10^{-8}	1.63×10^{-8}

Based on Table 6, the results of panel analysis using truncated splines for Bhutan produced the following equation.

a. **Variable X_1**

$$\hat{y}_{6j} = 1.31 \times 10^{-8} + 2.71 \times 10^{-7} X_{16j} + 6.72 \times 10^{-9} (X_{16j} - 20.2048)_+ + 4.73 \times 10^{-9} (X_{16j} - 20.2938)_+ + 3.44 \times 10^{-9} (X_{16j} - 20.5755)_+$$

or

$$\hat{y}_j = \begin{cases} 1.31 \times 10^{-8} + 2.71 \times 10^{-7} X_{16j} & , \quad X_{16j} < 20.2048 \\ -1.27 \times 10^{-7} + 2.78 \times 10^{-7} X_{16j} & , 20.2048 \le X_{16j} < 20.2938 \\ -2.19 \times 10^{-7} + 2.82 \times 10^{-7} X_{16j} & , 20.2938 \le X_{16j} < 20.5755 \\ -2.89 \times 10^{-7} + 2.86 \times 10^{-7} X_{16j} & , \quad X_{16j} \ge 20.5755 \end{cases}$$

b. Variable X_2

$$\hat{y}_{6j} = 1.31 \times 10^{-8} + 0.000137X_{26j} - 2.33 \times 10^{-5}(X_{26j} - 12087.1)_+ - 6.19 \times 10^{-5}(X_{26j} - 13815.5)_+ - 0.00011(X_{26j} - 19288.6)_+$$

or

$$\hat{y}_j = \begin{cases} 1.31 \times 10^{-8} + 0.000137X_{26j} & , \quad X_{26j} < 12087.1 \\ 0.28163964 + 0.0001137X_{26j} & , 12087.1 \le X_{26j} < 13815.5 \\ 1.13681908 + 0.0000518X_{26j} & , 13815.5 \le X_{26j} < 19288.6 \\ 3.25856508 - 0.0000582X_{26j} & , \quad X_{26j} \ge 19288.6 \end{cases}$$

c. Variable X_3

$$\hat{y}_{6j} = 1.31 \times 10^{-8} + 5.83 \times 10^{-7}X_{36j} + 2.80 \times 10^{-8}(X_{36j} - 42.3861)_+ + 1.86 \times 10^{-8}(X_{36j} - 42.807)_+ + 1.63 \times 10^{-8}(X_{36j} - 44.1396)_+$$

or

$$\hat{y}_j = \begin{cases} 1.31 \times 10^{-8} + 5.83 \times 10^{-7}X_{36j} & , \quad X_{36j} < 42.3861 \\ -1.18 \times 10^{-6} + 6.11 \times 10^{-7}X_{36j} & , 42.3861 \le X_{36j} < 42.807 \\ -1.97 \times 10^{-6} + 6.3 \times 10^{-7}X_{36j} & , 42.807 \le X_{36j} < 44.1396 \\ -2.69 \times 10^{-6} + 6.46 \times 10^{-7}X_{36j} & , \quad X_{36j} \ge 44.1396 \end{cases}$$

3.4 Model Interpretation

Based on the truncated spline regression model that has been constructed and estimated, the following is a description of how each predictor variable contributes to the variation in carbon dioxide (CO_2) emissions in Bhutan.

a. The Effect of Population Density (X_1) on CO_2 Emissions

Population density and carbon dioxide (CO_2) emissions show a non-linear pattern. If the population density is below 20.2048 people/km², each additional 1 person/km² will increase CO_2 emissions by 2.71×10^{-7} metric tons. At densities between 20.2048 and 20.2938 people/km², the effect increases to 2.78×10^{-7} metric tons per additional 1 person/km². Furthermore, in the range of 20.2938–20.5755 people/km², an additional 1 person/km² increases CO_2 emissions by 2.82×10^{-7} metric tons. Even after density exceeds 20.5755 people/km², an increase of 1 person/km² results in additional emissions of approximately 2.86 × 10-7 metric tons, indicating that the more densely populated an area is, the more significant its contribution to emissions.

b. The Effect of Primary Energy Consumption (X_2) on CO_2 Emissions

Primary energy consumption also has a significant effect on CO_2 emissions. At consumption levels below 12,087.1 TWh, each additional 1 TWh increases emissions by 1.37×10^{-4} metric tons. However, after exceeding 12,087.1 TWh, the effect reverses, where each additional 1 TWh actually increases emissions by 0.0001137 3 metric tons. This increase also occurs in the range of 13,815.5–19,288.6 TWh, amounting to 0.0000518 metric tons per additional 1 TWh. After consumption exceeds 19,288.6 TWh, the downward effect becomes most significant, with an additional 1 TWh associated with a reduction in emissions of around 0.0000582 metric tons. This is in line with a report by the Intergovernmental Panel on Climate Change, which explains that when energy consumption is high, the role of renewable energy and energy efficiency becomes more important in reducing CO_2 emissions.

c. The Effect of Urban Population (X_3) on CO_2 Emissions

The urban population shows an increasingly strong positive correlation with CO_2 emissions. At urban population levels below 42.3861%, each 1% increase in the urban population adds 5.83 $\times 10^{-7}$ metric tons to emissions. In the range of 42.3861–42.807%, A 1% increase is associated with an increase in emissions of 6.11×10^{-7} metric tons. Furthermore, in the range of 42.807–44.1396%, an additional 1% of urban population increases emissions by 6.3×10^{-7} metric tons. After exceeding 44.1396%, the effect of urban population becomes even greater, where a 1% increase in the urban population results in an additional 6.46×10^{-7} metric tons of emissions.

3.5 Thematic Map

Next, thematic maps are displayed to illustrate the grouping of countries in Asia for each predictor, based on the coefficient signs from the estimation results of the three-knot spline truncated panel data model.

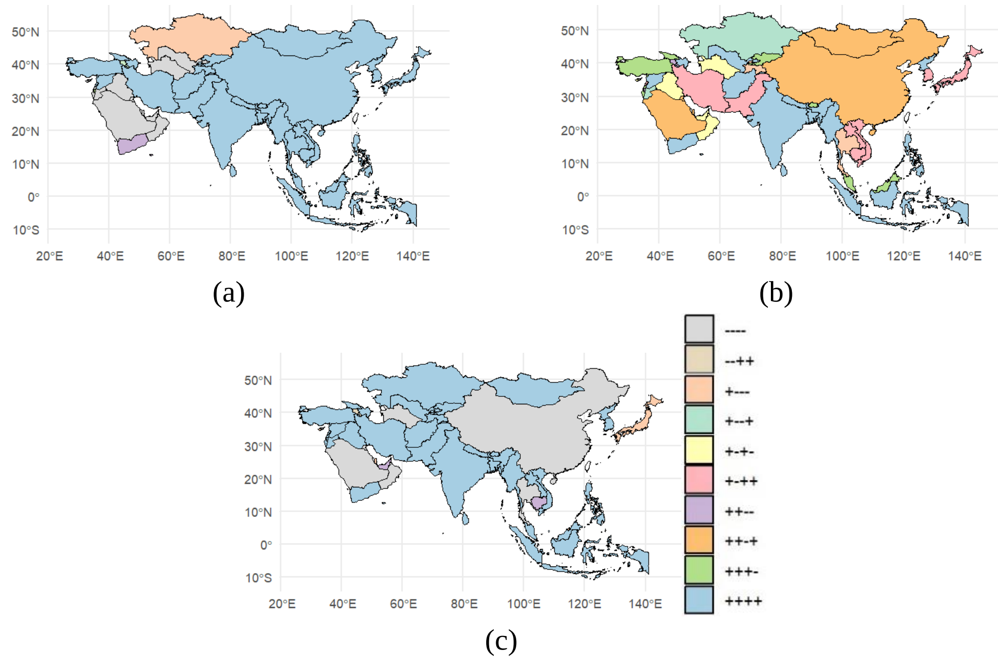

Figure 1. The effect of Predictor Variables on CO_2 Emissions: (a) Population Density, (b) Primary Energy Consumption, and (c) Urban Population

Based on the thematic map above, each predictor variable shows a different effect on CO_2 emissions in Asian countries. In terms of population density, Kazakhstan is colored orange, which means that the country's density is relatively low, but its CO_2 emissions remain high because the country's economy is still heavily dependent on fossil fuels. For the primary consumption variable, the colors vary between countries, indicating differences in energy use influenced by industrialization, energy availability, and the socioeconomic conditions of each country. In terms of urban population, several countries such as China, Kuwait, Oman, Saudi Arabia, Thailand, and Turkmenistan are gray, which means that the level of urban population in these countries is not always in line with high CO_2 emissions. This is due to several other factors such as the energy transition in China or dependence on oil and gas in the Gulf countries.

4. CONCLUSION

Based on the analysis, the best model was obtained from truncated spline nonparametric regression with three knot points, resulting in a minimum GCV value of 0.00004, MSE of 0.005248, and R^2 of 99.99206%. These results indicate that almost all of the variation in CO_2 emissions in Asian countries can be explained by population density, primary energy consumption, and urban population, and confirm the effectiveness of truncated splines in capturing nonlinear relationship patterns. In addition, the interpretation of the thematic map shows that each predictor variable has a different effect between countries. In the population density variable, for example, Kazakhstan shows high emission values despite its low population density, indicating a high dependence on fossil fuels. Meanwhile, in the primary energy consumption variable, there are significant differences between countries due to variations in industrialization levels and energy sources used. As for the urban population variable, countries such as China, Kuwait, Oman, Saudi Arabia, Thailand, and Turkmenistan show that high urban population does not always correlate directly with increased CO_2 emissions. This condition can occur due to other factors, such as the use of renewable energy and increased energy efficiency, which contribute to reducing emission rates in urban areas. For further research, the analysis can be expanded with a longer time period, wider geographical coverage, or the addition of other predictor variables, as well as considering variations in the number of knot points to obtain more comprehensive results in support of sustainable energy and environmental policies.

ACKNOWLEDGEMENT

The author would like to express sincere gratitude to Universitas Airlangga for the guidance and support provided throughout the completion of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest

REFERENCES

[1] Ahmad, N. A., & Raupong, R., 2023. Estimation Of Parameter Regression Panel Data Model Using Least Square Dummy Variable Method. *Jurnal Matematika, Statistika Dan Komputasi*, *20*(1), 221–228. https://doi.org/10.20956/j.v20i1.27530

[2] Energy Institute, 2023. Statistical Review of World Energy 2024 (73rd edition). *Energy Institute* (Vol. 73). https://www.energyinst.org/__data/assets/pdf_file/0006/1542714/684_EI_Stat_Review_V16_DIGITAL.pdf

[3] Energy Tracker Asia, 2024. *Are Carbon Emissions Decreasing?* Energy Tracker Asia. https://energytracker.asia/are-carbon-emissions-decreasing/

[4] Fan, J., & Yao, Q., 2003. *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer. https://doi.org/10.1198/tech.2004.s746

[5] Gao, J., 2024. R-Squared (R2) – How Much Variation is Explained? *Research Methods in Medicine & Health Sciences*, *5*(4), 104–109. https://doi.org/10.1177/26320843231186398

[6] Gunawan, I., Kusnawan, A., & Hernawan, E., 2021. Impact of Work from Home Policy Implementation on Work Effectiveness and Productivity in Tangerang City. *Primanomics :*

Jurnal Ekonomi & Bisnis, *19*(1), 99–107. https://doi.org/10.31253/pe.v19i1.508

[7] International Energy Agency, 2021. Global Energy Review 2021. *International Energy Agency*. https://iea.blob.core.windows.net/assets/d0031107-401d-4a2f-a48b-9eed19457335/GlobalEnergyReview2021.pdf

[8] International Energy Agency, 2022. Global Energy Review: CO2 Emissions in 2021. *International Energy Agency*. https://www.iea.org/news/global-co2-emissions-rebounded-to-their-highest-level-in-history-in-2021

[9] International Energy Agency, 2023. CO2 Emissions in 2022. *International Energy Agency*. https://www.iea.org/reports/co2-emissions-in-2022

[10] International Energy Agency, 2024a. CO2 Emissions in 2023. *International Energy Agency*. https://www.iea.org/reports/co2-emissions-in-2023

[11] International Energy Agency, 2024b. *Energy-Intensive Economic Growth, Compounded by Unfavourable Weather, Pushed Emissions Up in China and India*. International Energy Agency. https://www.iea.org/reports/co2-emissions-in-2023/energy-intensive-economic-growth-compounded-by-unfavourable-weather-pushed-emissions-up-in-china-and-india

[12] Nurhuda, G. N., Wasono, W., & Nohe, D. A., 2022. Nonparametric Regression Modeling Based on Spline Truncated Estimator on Simulation Data. *Jurnal Matematika, Statistika Dan Komputasi*, *19*(1), 172–182. https://doi.org/10.20956/j.v19i1.21534

[13] Ramli, M., Ratnasari, V., & Nyoman Budiantara, I., 2020. Estimation of Matrix Variance-Covariance on Nonparametric Regression Spline Truncated for Longitudinal Data. *Journal of Physics: Conference Series*, *1562*(1), 012014. https://doi.org/10.1088/1742-6596/1562/1/012014

[14] Reuters, 2024. *Top 10 Country Emitters Discharged Record Amount of CO2 in 2023*. Reuters. https://www.reuters.com/markets/commodities/top-10-country-emitters-discharged-record-amount-co2-2023-2024-06-21/

[15] Salhuteru, R., & Loklomin, S. B., 2024. Application Of Spline Truncated Nonparametric Regression In Modeling Factors Affecting Human Development Index In Maluku And North Maluku Provinces. *Motekar:Journal of Education and Science*, *1*(2), 89–102.