BAREKENG: Journal of Mathematics and Its Applications
March 2026 Volume 20 Issue 1 Page 0123–0136
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss1pp0123-0136

MODELLING SCHOOL DROPOUT RATES IN WEST JAVA PROVINCE WITH MIXED GEOGRAPHICALLY TEMPORALLY WEIGHTED REGRESSION

Prizka Rismawati Arum[1]*, Endang Tri Wahyuni Maharani[2], Diandra Fatimahthus Zahra[3], Tiani Wahyu Utami[4]

[1,3,4]*Department of Statistics, Universitas Muhammadiyah Semarang*
[2]*Department of Chemistry Education, Universitas Muhammadiyah Semarang*
Jln. Kedungmundu No 18, Tembalang, Semarang, Jawa Tengah, 50273, Indonesia

Corresponding author's e-mail: **prizka.rismawatiarum@unimus.ac.id**

Article Info

Article History:
Received: 17th February 2025
Revised: 26th May 2025
Accepted: 4th August 2025
Available online: 24th November 2025

Keywords:
Dropout rate;
Fixed gaussian kernel;
Mixed geographically and temporally weighted regression;
MGTWR.

ABSTRACT

School dropout is a problem in the education sector that can hinder the progress of the quality of human resources and the competitiveness of the nation. West Java Province has the highest school dropout rates among all provinces in Indonesia. The data on school dropout rates exhibit spatial and temporal variations. Additionally, the potential differences between regions allow for the occurrence of diverse data that can be addressed locally and globally. Mixed Geographically Temporally Weighted Regression (MGTWR) is an extension of the GWR method that can produce parameters that are both local and global for each location and time. So, the objective of this research is to obtain factors that have a local and global influence on the school dropout rate in West Java Province using the Mixed Geographically Temporally Weighted Regression method. In this study, the data used includes school dropout rates in West Java Province from 2018 to 2022. The data used is sourced from the official statistical data website of the Ministry of Education, Culture, Research and Technology, and the official website of the West Java Province Central Statistics Agency. The results of the MGTWR modeling show that globally influential variables include the percentage of the poor population, population density, unemployment rate, and average length of schooling, which have local effects. Based on the MGTWR model, the Fixed Kernel Gaussian weighting function is the best model for modeling school dropout rates in regencies/cities in West Java, with an RMSE value of 0.0755 and R-squares of 92.09%.

How to cite this article:

P. R. Arum, E. T. W. Maharani, D. F. Zahra and T. W. Utami., "MODELLING SCHOOL DROPOUT RATES IN WEST JAVA PROVINCE WITH MIXED GEOGRAPHICALLY TEMPORALLY WEIGHTED REGRESSION," *BAREKENG: J. Math. & App.*, vol. 20, iss. 1, pp. 0123-0136, Mar, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

A crucial supportive factor in enhancing the quality of Human Resources is education. Education serves as the driving force for the formation of quality individuals and provides a path to improve the quality and competitiveness of human resources for the future. In Indonesia, education is a means to achieve one of the goals stated in the preamble of the 1945 Constitution, which is to enlighten the life of the nation. One of the government's efforts in the field of education is implementing the mandatory school attendance program. In this program, the government enforces 12 years of basic education. This is because, in line with the development of technology, it requires education and skills to be increasingly higher [1]. Data from the Statistics Indonesia (BPS), the total number of Indonesia's population of compulsory school age is 66,265,300 people [2]. The data indicates that the population whose age already fulfills the 12-year compulsory education program has a significantly high number. Meanwhile, the total number of students in Indonesia registered in the Basic Education Data (DAPODIK) is 44,103,847 students [3]. From the mentioned data, there is still a discrepancy, and it is estimated that there are still individuals in Indonesia who have experienced school dropout. School dropout in Indonesia continues to be a national issue, especially in the field of education. Based on data from the Statistics Indonesia, it is recorded that in 2022, the dropout rate increased compared to 2021 across all levels of education. The number of school dropouts is distributed across all provinces in Indonesia, especially in the West Java Province. This is because in 2022, West Java tops the list of school dropout rates, with a total of 10,099 students [4]. The position of West Java Province, which occupies the first position in the highest school dropout rates, did not only occur in 2022. In previous years, West Java Province has consistently dominated the highest school dropout rates among all provinces in Indonesia.

In each region, the school dropout rate is influenced by different factors. This is due to varying geographical, economic, and socio-cultural conditions among different regions [5], [6]. Spatial data is often found in this field. Spatial data refers to data types containing information in the form of interrelated attributes and locations. The relationship between adjacent observation points will create spatial effects. Spatial effects are closely related to the location and geographic characteristics of each observation point, resulting in variations in each study, known as spatial heterogeneity.To analyze data influenced by regional factors, a weighted regression analysis that considers the geographical aspect is employed, known as Geographically Weighted Regression (GWR). GWR is a method developed from the classical linear regression model that takes into account weights based on geographical location [7], [8], [9]. Space and time constitute the two essential dimensions present in all human activities, social occurrences, and environmental processes [10], [11]. While the GWR method addresses spatial information, it does not account for temporal information, despite time being another available dimension [12]. In data on school dropout rates in West Java, there are variations from year to year and differences from one region to another. Therefore, there is the development of a spatial-temporal (time) method called Geographically Temporally Weighted Regression (GTWR). GTWR is a more representative method because it generates parameters with local characteristics for each location and time [13].

West Java Province exhibits significant diversity among its regencies/cities. This allows for similarities to occur in a regency/city in West Java. By employing the GTWR model, the resulting parameters will be local for each location and time. If there are variables that are not significant in the test, it means that not all explanatory variables in the model have a spatial effect on the response or have a global effect. Therefore, the GTWR model can be developed into a Mixed Geographically and Temporally Weighted Regression (MGTWR) model. Research utilizing the MGTWR method has been conducted by [14], [15]. In their study, the use of the MGTWR method is considered the best approach in the case of poverty in North Sumatra [16]. Research suggests that using the MGTWR method produces more accurate results.

Based on the explanation above, a study on school dropout rates in West Java Province is intended. West Java Province has diverse regencies/cities, leading to varying numbers of school dropouts in the province each year. To analyze this case, modeling will be conducted using the Mixed Geographically Temporally Weighted Regression (MGTWR) method. The research was conducted to find out the general picture of school dropouts in West Java Province and to be able to provide information about the factors that influence locally and globally the school dropout rate in West Java Province to related parties. Thus, it is hoped that these results can help in taking appropriate policies to reduce school dropout rates effectively. [17] [18], [19].

2. RESEARCH METHODS

2.1 Data and Variables

The data used in this study are secondary data obtained from the Basic Education Data, the official website of the Ministry of Education, Culture, Research, and Technology statistics data, and the official website of the Central Statistics Agency of West Java Province. The data pertains to the number of school dropouts in West Java Province from 2018 to 2022. The observation units used were districts/cities in West Java Province, which consists of 18 districts and 9 cities.The response variable in this study is the school dropout rate (y), while the predictor variables are Human Development Index (x_1), Gross Regional Domestic Product (GRDP) per capita (x_2), percentage of population living in poverty (x_3), population density (x_4), mean years of schooling (x_5), unemployment rate (x_6).

2.2 Data Analysis Steps

Steps of data analysis were carried out as follows:

1. Performing global regression modeling.
   a. Testing classical assumptions (Normality Test, Multicollinearity Test, Autocorrelation Test).
   b. Testing spatial heterogeneity with the Breusch-Pagan statistical test. This test uses the following hypothesis.

   $H_0 : \sigma_i^2 = 0$ (there is no spatial heterogeneity)
   $H_1 : \sigma_i^2 \neq 0$ (there is spatial heterogeneity)

   Breusch-Pagan Test Statistics:

   $$BP = \frac{1}{2}\boldsymbol{f^T Z}\left(\boldsymbol{Z^T Z}\right)^{-\mathbf{1}}\boldsymbol{Z^T f} \tag{1}$$

   with $f_i = \frac{e_i^2}{\sigma^2} - 1$
   Test criteria used are rejected H_0 if the test statistic value or if the p-value < α with k is the number of parameters BP > $x^2_{\alpha;(p)}$, with p is the number of predictors.
2. Determine the longitude (u_i) and latitude (v_i) coordinates for each district/city in West Java Province
3. Finding the values of τ, μ, and λ.
   a. Calculating the Euclidean distance at coordinates ($u_i v_i t_i$)
   b. Iterating to obtain the optimal parameter estimation for τ by comparing the R^2 values.
   c. Obtaining parameter estimates μ and λ from the results of the optimal parameter estimation for τ.
   d. Determining the spatial-temporal bandwidth

   $$\left(d_{ij}^{ST}\right)^2 = \lambda\left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \mu\left(t_i - t_j\right)^2 \tag{2}$$
   $$h_S^2 = {h_{ST}^2}/{\lambda}$$

   e. Calculating the weighted values using the Fixed Kernel Gaussian and Adaptive Kernel Gaussian weighting functions.

   $$W_j(u_i v_i) = exp\left(-\frac{1}{2}\left(\frac{d_{ij}^{ST}}{h_{ST}^2}\right)^2\right) \tag{3}$$

4. Estimating the parameters of the GTWR model.

   $$\widehat{\boldsymbol{\beta}}_{(u_i v_i t_i)} = \left[\boldsymbol{X}^T \boldsymbol{W}_{(u_i v_i t_i)} \boldsymbol{X}\right]^{-1} \boldsymbol{X}^T \boldsymbol{W}_{(u_i v_i t_i)} \boldsymbol{Y} \tag{4}$$

5. Testing the significance of model parameters Geographically Temporally Weighted Regression (GTWR)
   a. Model fitness testing (F-Test)
   The test hypothesis as follows:

   $H_0 = \beta_k(u_i v_i t_i) = \beta_k, , k = 0, 1, 2, \cdots$

$H_1 = \beta_k(u_i v_i t_i) \neq \beta_k$

The test statistic used is

$$F = \frac{\frac{SSE(H_0)}{df_1}}{\frac{SSE(H_1)}{df_2}} \tag{5}$$

with:

$SSE(H_0) = \boldsymbol{Y^T}(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}$
$SSE(H_1) = \boldsymbol{Y^T}(\boldsymbol{I}-\boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{Y}$
$\boldsymbol{I}(n \times n)$ = identity matrix

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{X}_1^T[\boldsymbol{X}^T W(u_1 v_1 t_1)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_1 v_1 t_1) \\ \boldsymbol{X}_2^T[\boldsymbol{X}^T \boldsymbol{W}(u_2 v_2 t_2)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_2 v_2 t_2) \\ \vdots \\ \boldsymbol{X}_n^T[\boldsymbol{X}^T \boldsymbol{W}(u_n v_n t_n)X]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_n, v_n, t_n) \end{bmatrix}, \boldsymbol{X}_i^T = (1, x_{i1}, x_{i2}, \cdots, x_{ip})$$

$\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$
$df_1 = n - p - 1$
$df_2 = (n - 2tr(\boldsymbol{L}) + tr(\boldsymbol{L}^T\boldsymbol{L}))$

Reject H_0 if F > F_{α,df_1,df_2} indicating that the presence of significant differences between the classical regression model and the GTWR model.

b. Testing of model parameters (T-Test)
The test hypothesis as follows:

$H_0 = \beta_k(u_i v_i t_i) = \beta_k, k = 0, 1, 2, \cdots$
$H_1 = \beta_k(u_i v_i t_i) \neq \beta_k$

The test statistic used is

$$t = \frac{\hat{\beta}_k(u_i, v_i, t_i)}{\hat{\sigma}\sqrt{c_{kk}}} \tag{6}$$

Reject H_0 if $|t| > t\alpha_{/2,df}$ indicating there is a significant difference in the influence between the response variable with location and time.

6. Determine global variables and local variables using spatial variability testing
This testing is conducted through the calculation of the F-test with the hypothesis as follows.

$H_0 = \beta_k(u_i v_i t_i) = \beta_k, i = 1, 2, \cdots, n\,; k$ = the coefficient index is assumed to be global.
$H_1 = \beta_k(u_i v_i t_i) \neq \beta_k, i = 1, 2, \cdots, n\,; k$

Test statistic:

$$F = \frac{\frac{V_k^2}{tr\left(\frac{1}{n}\boldsymbol{B}_k^T\left[\boldsymbol{I}-\frac{\boldsymbol{1}}{\boldsymbol{n}}\boldsymbol{J}\right]\right)\boldsymbol{B}_k}}{\frac{SSE(H_1)}{b_1}} \tag{7}$$

with

$\boldsymbol{B}_k = \boldsymbol{e}_1^T[\boldsymbol{X}^T\boldsymbol{W}(u_1 v_1 t_1)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_1 v_1 t_1),$

$$\boldsymbol{\beta}_k(u_i v_i t_i) = \begin{bmatrix} \beta_0(u_i v_i t_i) \\ \beta_1(u_i v_i t_i) \\ \vdots \\ \beta_k(u_i v_i t_i) \end{bmatrix},$$

$V_k^2 = \frac{1}{n}\sum_{i=1}^n\left(\hat{\beta}_k(u_i v_i t_i) - \frac{1}{n}\sum_{i=1}^n\hat{\beta}_k(u_i v_i t_i)\right)^2 = \frac{1}{n}\boldsymbol{\beta'}_k\left[\mathbf{I}-\frac{1}{n}\boldsymbol{J}\right]\boldsymbol{\beta}_k,$
$b_i = tr\left([(I-L)^T(I-L)]^i\right), i = 1,2$

$$c_i = tr\left(\left(\frac{1}{n}\boldsymbol{B}_k^T\left[\mathbf{I} - \frac{\mathbf{1}}{\boldsymbol{n}}\boldsymbol{J}\right]\right)\boldsymbol{B}_k\right)^i, i = 1,2.$$

$\mathbf{J}$ is a matrix of size $(n \times n)$ consisting of all 1 values, and $\boldsymbol{e}_{\boldsymbol{k}}^{\boldsymbol{T}}$ is a column vector of size $(p+1)$with a value of one for the k-th element and zero for others.
The testing criterion is to reject H_0 if the F-test statistic > $F_{(\alpha;df_1;df_2)}$, which means the predictor variable is locally specific. where, $df_1 = \frac{c_1^2}{c_2}$ and $df_2 = \frac{b_1^2}{b_2}$.

6. Estimating the parameters of the MGTWR model
   a. Calculating the global variable estimator

$$\widehat{\boldsymbol{\beta}_g} = \left[\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g\right]^{-1}\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{Y}$$

$$\boldsymbol{S}_l = \begin{bmatrix} \boldsymbol{X}_{l1}\left[\boldsymbol{X}_l^T\boldsymbol{W}\,(u_1v_1t_1))\boldsymbol{X}_l\right]^{-1}\boldsymbol{X}_l^T\boldsymbol{W}\,(u_1v_1t_1) \\ \boldsymbol{X}_{l2}\left[X_l^T\boldsymbol{W}\,(u_2v_2t_2)\boldsymbol{X}_l\right]^{-1}\boldsymbol{X}_l^T\boldsymbol{W}\,(u_2v_2t_2) \\ \vdots \\ \boldsymbol{X}_{ln}\left[\boldsymbol{X}_l^T\boldsymbol{W}\,(u_nv_nt_n)\boldsymbol{X}_l\right]^{-1}\boldsymbol{X}_l^T\boldsymbol{W}\,(u_nv_nt_n) \end{bmatrix}$$

   b. Calculating the local variable estimator

$$\widehat{\boldsymbol{\beta}_l}\,(u_iv_it_i) = \left[\boldsymbol{X}_l^T\boldsymbol{W}\,(u_iv_it_i)\boldsymbol{X}_l\right]^{-1}\boldsymbol{X}_l^T\boldsymbol{W}\,(u_iv_it_i)(\boldsymbol{Y} - \boldsymbol{X}_g\widehat{\boldsymbol{\beta}_g})$$

   c. Determine the MGTWR model response estimator

$$\widehat{\boldsymbol{Y}} = \boldsymbol{SY}$$
$$\boldsymbol{S} = \boldsymbol{S}_l + (\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g\left[\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g\right]^{-1}\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)$$

7. Testing the significance of model parameters Mixed Geographically Temporally Weighted Regression (**MGTWR**)
   a. Model fitness testing (F-Test)
      The test hypothesis as follows:

$$H_0 = \beta_k(u_iv_it_i) = \beta_k,, k = 0, 1, 2, \cdots$$
$$H_1 = \beta_k(u_iv_it_i) \neq \beta_k$$

      The test statistic used is

$$F = \frac{(JKG(H_0) - \frac{(JKG(H_1)}{v}}{\frac{(JKG(H_1)}{\delta}} = \left[\frac{(\boldsymbol{Y}^T\boldsymbol{R}_0\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{R}_4\boldsymbol{Y})}{v}\right]\left[\frac{\boldsymbol{Y}^T\boldsymbol{R}_4\boldsymbol{Y})}{\delta}\right]^{-1} \quad (8)$$

      With $v = tr(\boldsymbol{R}_0 - \boldsymbol{R}_4)$ and $\delta = tr(\boldsymbol{R}_4)$. The testing criterion is to reject H_0 if the F-test > $F_{\left(\alpha;\frac{v^2}{v'};\frac{\delta^2}{\delta'}\right)}$ indicating that the presence of significant differences between the classical regression model and the MGTWR model.
   b. Testing of model parameters (T-Test)
      i. Test the significance of global variables
         The test hypothesis as follows:

$$H_0 = \beta_g = \beta_g \text{ (the global variable } x_k \text{ is not significant)}$$
$$H_1 = \beta_g \neq \beta_g \text{ (the global variable } x_k \text{ is significant)}$$

         With the formula

$$T_g = \frac{\hat{\beta}_g}{\hat{\sigma}\sqrt{g_{kk}}}$$

         With g_{kk} is the -th diagonal element of the matrix $\boldsymbol{G}\boldsymbol{G}^T$ and $\boldsymbol{G} = \left[\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g\right]^{-1}\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)$, $\hat{\sigma}^2 = \frac{\boldsymbol{Y}^T[(\boldsymbol{I}-\boldsymbol{S})^T(\boldsymbol{I}-\boldsymbol{S})]\boldsymbol{Y}}{tr[(\boldsymbol{I}-\boldsymbol{S})^T(\boldsymbol{I}-\boldsymbol{S})]}$. Reject H_0 if $|t| > t_{\alpha/2,df}$.
      ii. Test the significance of local variables

$H_0 = \beta_k(u_i v_i t_i) = \beta_k$ (local variable x_l at location-i is not significant)
$H_1 = \beta_k(u_i v_i t_i) \neq \beta_k$ (local variable x_l at location-i is significant)

With the formula

$$T_l = \frac{\hat{\beta}_k(u_i v_i t_i)}{\hat{\sigma}\sqrt{m_{kk}}} \tag{9}$$

With m_{kk} is the -th diagonal element of the matrix $\boldsymbol{MM}^T$ and $\boldsymbol{M} = \left[\boldsymbol{X}_l^T \boldsymbol{W}\,(u_i v_i t_i) X_l\right]^{-1} \boldsymbol{X}_l^T \boldsymbol{W}\,(u_i v_i t_i)(1 - \boldsymbol{X}_g \boldsymbol{G})$. Reject H_0 if $|t| > t_{\alpha/2, df}$.

8. Determining the best model based on R-Square and RMSE values.
   a. The formula for the R-Square is as follows

$$R^2 = 1 - \frac{SSE}{SST} \tag{10}$$

   With SSE is Sum Square Error and SST is Sum Square Total
   b. The Root Mean Square Error (RMSE) is the square root of the sum of the squares of the differences between actual and predicted values. The formula for RMSE is:

$$RMSE = \sqrt{\sum \frac{(y_i - \hat{y}_i)^2}{n}} \tag{11}$$

3. RESULTS AND DISCUSSION

3.1 Descriptive Analysis

Descriptive analysis is used to understand the overview of data and characteristics of the variables used. The results of the descriptive analysis of the data on School Dropout Rates in Regencies/Cities in West Java Province for the years 2018-2022 are obtained as shown in the Table 1 below:

Table 1. Descriptive Analysis

Variable	N	Minimum	Maximum	Mean	Std. Deviation
y	135	13.00	7710.00	934.04	1110.85
x_1	135	64.62	82.50	71.76	4.72
x_2	135	12869.00	85820.00	29201.50	18753.92
x_3	135	2.07	13.13	8.28	2.77
x_4	135	383.00	15643.00	3924.32	4643.89
x_5	135	5.98	11.75	8.56	1.45
x_6	135	1.56	14.29	8.59	2.32

Table 1 shows that the average number of school dropouts in regencies/cities in West Java Province for the years 2018-2022 is 934.04 students. This is a considerably high figure and makes West Java Province the province with the highest number of school dropouts. The highest number of school dropouts from 2018-2022 is 7710 students, found in Bogor Regency in 2019. Meanwhile, the lowest number is 13 students in Cimahi City in 2022.

3.2 Classic Assumption Test

1. Normality Test

Table 2. Normality Test Result

Kolmogorov-Smirnov	P-value
0.073061	0.07427 > 0.05

Based on Table 2, the obtained values are D = 0.073061 and P-value = 0.07427. The hypothesis rule is that if the value of D < D_{table} and p-value > α, then accept H_0 or the data follows a normal distribution. Therefore, it can be said that the value of D (0.073061) < D_{table} (0.11705) and P-value (0.07427) > α (0.05), which means that the residuals have a normal distribution.

2. Multicollinearity Test

Table 3. Multicollinearity Test Result

Variable	VIF
x_1	4.9232
x_2	1.6944
x_3	2.0480
x_4	5.5277
x_5	7.1200
x_6	1.2649

Multicollinearity examination is checked from the VIF value. Multicollinearity occurs if the VIF value is greater than ten (VIF > 10). In Table 3, it is found that all explanatory variables have a VIF value of less than ten. Therefore, it can be concluded that there is no multicollinearity between variables.

3. Autocorrelation Test

Table 4. Autocorrelation Test Result

Durbin-Watson Value
1.7158

Based on Table 4, the obtained value of DW (Durbin Watson) is 1.7158. The testing is conducted with the criterion that if the Durbin Watson result falls between -2 and +2, then there is no autocorrelation. Based on the Durbin Watson result, the value falls within the range of -2 to +2. Therefore, it can be interpreted that there is no autocorrelation in the model.

3.3 Spatial and Temporal Heterogeneity Test

Table 5. Spatial Heterogeneity Test

Breusch-Pagan	P-value
19.1200	0.0039 < 0.05

Based on the Table 5, the obtained values are BP (Breusch-Pagan) = 19.1200 and p-value = 0.0039. The value of $\chi^2_{(6;0,05)}$ is 12.59158. Based on the two criteria where BP > $\chi^2_{(6;0,05)}$) and p-value < α, it can be concluded that there is spatial heterogeneity.

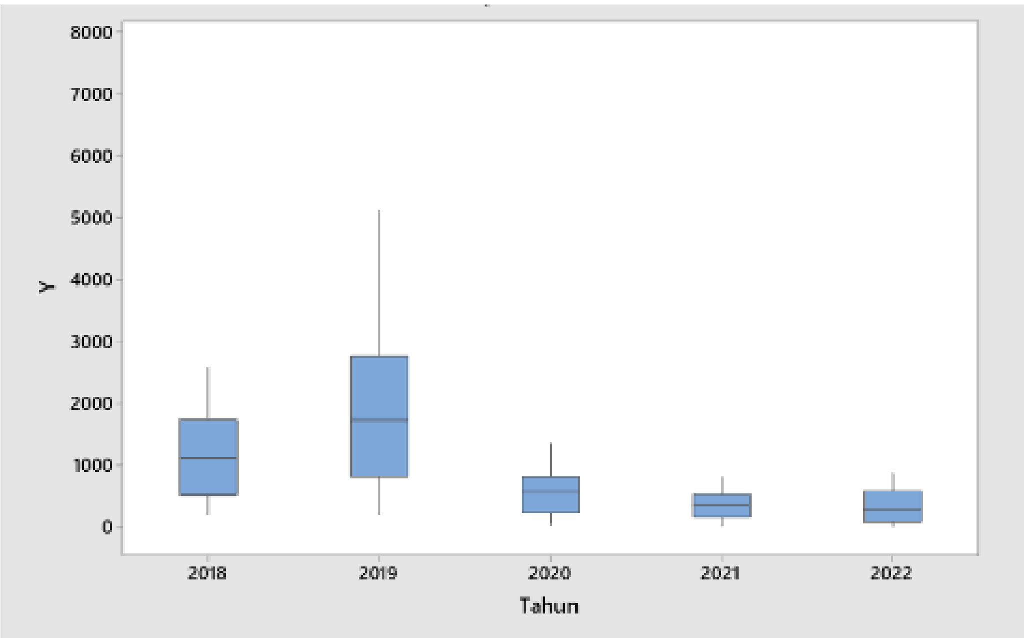

Figure 1. School Dropout Boxplot

Based on Fig. 1, diversity indicates the presence of non-stationarity over time or temporal heterogeneity. Fig. 1 shows the results showing significant differences and diversity in the number of school dropouts between observation periods in the 2018-2022 period. Between 2019 and 2022, a downward trend in this measurement was observed.

3.4 Geographically Temporally Weighted Regression

1. Model fitness testing (F-Test)

   The MGTWR model fitness test was conducted to determine whether the GTWR model better than the classical regression model. This test uses the following hypotheses:

   $H_0 = \beta_k(u_i v_i t_i) = \beta_k, k = 0, 1, 2, \cdots, q$ and $i = 1, 2, \cdots, n$ (GTWR model is no different from classical regression model)

   $H_1 =$ There is at least one $\beta_k(u_i v_i t_i)$ that is different (GTWR model is different from classical regression model)

Table 6. F Test Result

Weight	F	F_{table}	Decision
Fixed Kernel Gaussian	188.1410	8.5479	Reject H_0
Adaptive Kernel Gaussian	36.2768	1.3398	Reject H_0

   Based on Table 6, it is known that the F values for each weight are greater than the critical F values. Thus, H_0 is rejected, indicating a significant difference between the classical regression model and the GTWR model. It is concluded that the GTWR model is more suitable than classical regression model for analyzing school dropout rates in West Java.

2. Testing of model parameters (T-Test)

Table 7. T-Test Result

Weight	Parameter	t	Decision
Fixed Kernel Gaussian	$\hat{\beta}_0(u_i v_i t_i)$	2.0757	Reject H_0
	$\hat{\beta}_1(u_i v_i t_i)$	0.9902	Accept H_0
	$\hat{\beta}_2(u_i v_i t_i)$	0.5809	Accept H_0
	$\hat{\beta}_3(u_i v_i t_i)$	1.5336	Accept H_0
	$\hat{\beta}_4(u_i v_i t_i)$	2.4728	Reject H_0
	$\hat{\beta}_5(u_i v_i t_i)$	2.1134	Reject H_0
	$\hat{\beta}_6(u_i v_i t_i)$	3.1732	Reject H_0
Adaptive Kernel Gaussian	$\hat{\beta}_0(u_i v_i t_i)$	2.8125	Reject H_0
	$\hat{\beta}_1(u_i v_i t_i)$	0.0701	Accept H_0
	$\hat{\beta}_2(u_i v_i t_i)$	0.3204	Accept H_0
	$\hat{\beta}_3(u_i v_i t_i)$	4.8994	Reject H_0
	$\hat{\beta}_4(u_i v_i t_i)$	0.5351	Accept H_0
	$\hat{\beta}_5(u_i v_i t_i)$	3.8395	Reject H_0
	$\hat{\beta}_6(u_i v_i t_i)$	1.9605	Accept H_0

   Based on Table 7 in the fixed kernel Gaussian weighting function, the variables of population density(x_4), average years of schooling(x_5), and unemployment rate (x_6) have $t_{statistic} > t_{table}$ (1.97867), so it can be concluded that H_0 is rejected, meaning that the variables significantly influence on the school dropout rate. Meanwhile, for the adaptive kernel Gaussian weighting function, the variables of the percentage of poor population (x_3), and average years of schooling (x_5), have a significant impact on the school dropout rate.

3. Spatial Variability Test

   The spatial variability test is conducted to identify global and local coefficients in the Mixed Geographically Temporally Weighted Regression model.

Table 8. Variability Test Result

Weight	Parameter	F	F_{table}	Decision
Fixed Kernel Gaussian	$\hat{\beta}_0(u_iv_it_i)$	3132.8233	10.1279	Reject H_0
	$\hat{\beta}_1(u_iv_it_i)$	8.3134		Accept H_0
	$\hat{\beta}_2(u_iv_it_i)$	8.1191		Accept H_0
	$\hat{\beta}_3(u_iv_it_i)$	-0.0135		Accept H_0
	$\hat{\beta}_4(u_iv_it_i)$	0.0043		Accept H_0
	$\hat{\beta}_5(u_iv_it_i)$	37.7365		Reject H_0
	$\hat{\beta}_6(u_iv_it_i)$	0.2130		Accept H_0
Adaptive Kernel Bisquare	$\hat{\beta}_0(u_iv_it_i)$	9644.2420	3.9157	Reject H_0
	$\hat{\beta}_1(u_iv_it_i)$	7.2961		Reject H_0
	$\hat{\beta}_2(u_iv_it_i)$	-0.8490		Accept H_0
	$\hat{\beta}_3(u_iv_it_i)$	0.8478		Accept H_0
	$\hat{\beta}_4(u_iv_it_i)$	-0.0129		Accept H_0
	$\hat{\beta}_5(u_iv_it_i)$	47.9155		Reject H_0
	$\hat{\beta}_6(u_iv_it_i)$	-0.3189		Accept H_0

Based on Table 8, the results indicate that variable Average Years of Schooling (x_5) in the fixed kernel Gaussian weighting function has an F-value > F table, indicating a significant spatial influence. Meanwhile, for the variables Human Development Index (x_1), Gross Regional Domestic Product per capita (x_2), percentage of poor population (x_3), population density (x_4), and unemployment rate (x_6), the F-value < F table, indicating that these variables do not have significant spatial influence or can be assumed as global factors.

In the adaptive kernel Gaussian weighting function, the results show that the coefficients for the intercept and IPM (x_1) and Average Years of Schooling (x_5), have an F-value > F table, meaning these variables exhibit significant spatial influence. On the other hand, the coefficients for Gross Regional Domestic Product per capita (x_2), percentage of poor population (x_3), population density (x_4), and unemployment rate (x_6), have an F-value < F table, indicating that these variables do not have spatial influence or are considered as global factors.

3.5 Mixed Geographically Temporally Weighted Regression

1. Model Fitness Testing (F-Test)

   The MGTWR model fitness test was conducted to determine whether the MGTWR model better than the classical regression model. This test uses the following hypotheses:

   $H_0 = \beta_k(u_iv_it_i) = \beta_k, k = 0, 1, 2, \cdots, q$ and $i = 1, 2, \cdots, n$ (MGTWR model is no different from classical regression model)

   $H_1 =$ There is at least one $\beta_k(u_iv_it_i)$ that is different (MGTWR model is different from classical regression model)

   Table 9. F Test Result

Weight	F	F_{table}	Decision
Fixed Kernel Gaussian	7.6834	1.5241	Reject H_0
Adaptive Kernel Gaussian	9.7147	2.6721	Reject H_0

   Based on the table above with a significance level of 5%, it is found that each weight (Fixed Kernel Gaussian and Adaptive Kernel Gaussian) has an F value >F table. It can be concluded that the null hypothesis (H_0) is rejected, indicating a significant difference between the MGTWR model and the classical regression model.

2. Testing of Model Parameters (T-Test)

   The MGTWR parameter test is divided into two steps, involving tests for both global and local variables.

a. Global Parameter

Table 10. Global Parameter Test Result

Weight	**Parameter**	**Estimated Value**	**t**	t_{table}	**Decision**
Fixed Kernel Gaussian	$\hat{\beta}_1$	-5.2968	0.9606	1.97852	Accept H_0
	$\hat{\beta}_2$	-0.1313	0.4755		Accept H_0
	$\hat{\beta}_3$	-0.0496	4.0624		Reject H_0
	$\hat{\beta}_4$	0.0036	1.9794		Reject H_0
	$\hat{\beta}_6$	-0.0182	2.1808		Reject H_0
Adaptive Kernel Gaussian	$\hat{\beta}_2$	0.0509	0.3173	1.97838	Accept H_0
	$\hat{\beta}_3$	-0.0511	4.8877		Reject H_0
	$\hat{\beta}_4$	0.0007	0.5289		Accept H_0
	$\hat{\beta}_6$	0.0194	1.9698		Accept H_0

Based on Table 10 in the fixed kernel Gaussian weighting function, the variables of the percentage of poverty (x_3), population density(x_4), and unemployment rate (x_6) have $t_{statistic}>$ t_{table} (1,97867), so it can be concluded that H_0 is rejected, meaning that the variables significantly influence on the school dropout rate.

Meanwhile, for the adaptive kernel Gaussian weighting function, the variables of the percentage of poor population (x_3) have a significant impact on the school dropout rate.

b. Local Parameter

The purpose of the local parameter test is to identify local variables that significantly impact school dropout rates in each location. In the MGTWR parameter test, a t-test statistic was used for each model in each district/city in West Java. For example, we conducted the test for Bogor Regency in 2019.

Table 11. Local Parameter Test Result for Bogor Regency in 2019

Weight	**Parameter**	**Estimated Value**	**t**	t_{table}	**Decision**
Fixed Kernel Gaussian	$\hat{\beta}_0(u_i v_i t_i)$	8.8856	4.3633	1.97796	Reject H_0
	$\hat{\beta}_5(u_i v_i t_i)$	-5.7221	3.3823		Reject H_0
Adaptive Kernel Gaussian	$\hat{\beta}_0(u_i v_i t_i)$	4.3432	2.8066	1.97824	Reject H_0
	$\hat{\beta}_1(u_i v_i t_i)$	-0.4348	0.0678		Accept H_0
	$\hat{\beta}_5(u_i v_i t_i)$	-2.8922	3.8335		Reject H_0

Based on Table 11 in the fixed kernel Gaussian weighting function, the variable of Average Years of Schooling (x_5), have $t_{statistic}>$ t_{table} (1,97796), so it can be concluded that H_0 is rejected, meaning that the variable significantly influences on the school dropout rate.

The same thing applies to the adaptive kernel Gaussian weighting that the Average Years of Schooling (x_5), influences the dropout rate.

3.6 Selection of Best Model

Table 12. R² and RMSE Value Model MGTWR

Weight	R^2	**RMSE**
Fixed Kernel Gaussian	0.9209	0.0755
Adaptive Kernel Gaussian	0.2768	0.2286

Based on Table 12, it is concluded that the MGTWR model for Bogor Regency in 2019 with the fixed Gaussian kernel weighting function shows the smallest RMSE value, which is 0.0755, and an R^2 value of 0.9209. Therefore, it can be suggested that the MGTWR model with a fixed Gaussian kernel weighting is more effective in modeling dropout rates data in West Java.

The results of the global and local parameter tests produce the MGTWR model for Bogor Regency in 2019 with a fixed Gaussian kernel weighting as follows:

$$\hat{y}_i = 8.8856 - 5.7221x_5 - 0.0496x_3 + 0.0036x_4 - 0.0182x_6$$

In the model, the variables with negative coefficient values are given in the

1. Average length of schooling (x_5),
2. Percentage of population in poverty (x_3), and
3. Unemployment rate (x_6),

which influence the dropout rate with coefficient values of -5.7221, -0.0496, and -0.0182 respectively. This indicates that as the average length of schooling, percentage of population in poverty, and unemployment rate decrease, it will increase the dropout rate in West Java.

The province of West Java still has a high number of child laborers. A study conducted by [20] in the Ciomas District, Bogor Regency, West Java found that child laborers are aged 11-16 years. This indicates that child laborers are children who are still within the school age. The lack of interest in learning becomes the main reason for dropping out of school. In addition to being classified as dropouts, child laborers also have low levels of education. The low level of education is due to the lack of awareness among parents, the belief that higher education does not guarantee someone to get a decent job and earn a lot of money. Another reason is the strong motivation from the child's own desire to work independently and earn money.

Meanwhile, the population density variable (x_4) affects the dropout rate with a coefficient value of 0.0036, which is positive. This indicates that as population density increases, it will increase the dropout rate.

4. CONCLUSION

The best model for modeling dropout rates in West Java is obtained using the fixed Gaussian kernel weighting function which has the smallest RMSE value of 0.0755 and an R^2 of 92.09%. Factors that have a global influence on dropout rates in West Java include variables such as the percentage of the population living in poverty, population density, and the unemployment rate. The variable that has a local influence is the mean years of schooling. The variables that influence the school dropout rate in West Java locally vary in each district/city.

Author Contributions

Prizka Rismawati Arum: conceptualization, Formal Analysis, Funding Acquisition, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing-Original Draft, Writing-Review and Editing. Diandra Fatimahthus Zahra: Data Curation, Formal Analysis, Investigation, Resources, Software, Supervision, Visualization, Writing-Original Draft. Endang Tri Wahyuni Maharani: Conceptualization, Project Administration, Software, Supervision, Validation. Tiani Wahyu Utami: Data Curation, Methodology, Resources, Writing-Review and Editing. All authors discussed the results and contributed to the final manuscript.

Funding Statement

The author(s) disclosed receipt of the following financial support for the research, authorship, and publication of this article: This work was supported by a research grant from Universitas Muhammadiyah Semarang [grant number 0132/UNIMUS.L/PG/PJ.INT/2024].

Acknowledgment

The authors thank the Department of Statistics and the Department of Data Science for being an integral part of the Faculty of Science and Agriculture Technology at Universitas Muhammadiyah Semarang. Their substantial support and dedication in advancing this research to completion are immensely valued. The authors would also like to thank Universitas Muhammadiyah Semarang for the provision of resources, facilities, and academic guidance, which have been pivotal in supporting our research.

Declarations

The authors declares that she/he has no conflicts of interest to report study.

REFERENCES

[1] D. Saepuloh and A. Suherman, "ANALISIS PENYEBAB ANGKA PUTUS SEKOLAH DAN KONDISI SOSIAL EKONOMI KELUARGA DI KOTA TANGERANG," *Jurnal Penelitian Dan Karya Ilmiah*, vol. 18, no. 2, pp. 98–111, 2019, doi: https://doi.org/10.33592/pelita.v18i2.45.

[2] BPS, "JUMLAH PENDUDUK MENURUT KELOMPOK UMUR DAN JENIS KELAMIN, 2022."

[3] Kemdikbud, "Data Peserta Didik", *Kemdikbud*, 25 Desember 2022, [Online]. Available: https://dapo.dikdasmen.go.id/pd [Accessed: 26 September 2024].

[4] Kemdikbud, "JUMLAH SISWA PUTUS SEKOLAH", *Kemdikbud*, 25 Desember 2022, [Online]. Available: https://data.kemdikbud.go.id/dataset/p/peserta-didik/jumlah-siswa-putus-sekolah-menurut-tingkat-tiap-provinsi-indonesia-sd-2023 [Accessed: 26 September 2024].

[5] P. R. Arum, Kintoko, A. O. Winarta, I. M. Nur, and E. A. Purnomo, "MODELLING GENDER-BASED HUMAN DEVELOPMENT INDEX IN INDONESIA USING SPATIAL SEEMINGLY UNRELATED REGRESSION (S-SUR)," 2023, pp. 508–517. doi: https://doi.org/10.2991/978-2-38476-176-0_73.

[6] A. Fadlurohman, N. Ayu Nur Roosyidah, and N. Amalia Annisa, "SOCIAL VULNERABILITY ANALYSIS IN CENTRAL JAVA WITH K-MEDOIDS ALGORITHM," *Parameter: Journal of Statistics*, vol. 4, no. 2, pp. 83–92, Dec. 2024, doi: https://doi.org/10.22487/27765660.2024.v4.i2.17131.

[7] A. S. Fotheringham, B. Lu, M. Charlton, and P. Harris, "GEOGRAPHICALLY WEIGHTED REGRESSION : THE ANALYSIS OF SPATIALLY VARYING INTERNATIONAL JOURNAL OF GEOGRAPHICAL GEOGRAPHICALLY WEIGHTED REGRESSION WITH A NON-EUCLIDEAN DISTANCE METRIC : A CASE STUDY USING HEDONIC HOUSE PRICE DATA," *International Journal of Geographical Information Science*, vol. 00, no. 00, pp. 1–22, 2002, doi: https://doi.org/10.1080/13658816.2013.865739

[8] H. Cao, Z. Wu, and W. Zheng, "IMPACT OF TOURISTIFICATION AND LANDSCAPE PATTERN ON HABITAT QUALITY IN THE LONGJI RICE TERRACE ECOSYSTEM, SOUTHERN CHINA, BASED ON GEOGRAPHICALLY WEIGHTED REGRESSION MODELS," *Ecol Indic*, vol. 166, Sep. 2024, doi: https://doi.org/10.1016/j.ecolind.2024.112259.

[9] K. Kopczewska and P. Ćwiakowski, "SPATIO-TEMPORAL STABILITY OF HOUSING SUBMARKETS. TRACKING SPATIAL LOCATION OF CLUSTERS OF GEOGRAPHICALLY WEIGHTED REGRESSION ESTIMATES OF PRICE DETERMINANTS," *Land use policy*, vol. 103, Apr. 2021, doi: https://doi.org/10.1016/j.landusepol.2021.105292.

[10] P. R. Arum, I. Fathoni Amri, and S. Amri, "GLS ESTIMATION IN PYTHON TO FORECAST GROSS REGIONAL DOMESTIC PRODUCT USING GENERALIZED SPACE–TIME AUTOREGRESSIVE SEEMINGLY UNRELATED REGRESSION MODEL," *Front Appl Math Stat*, vol. 10, Aug. 2024, doi: https://doi.org/10.3389/fams.2024.1365723.

[11] A. Fadlurohman, A. Choiruddin, and J. Mateu, "INHOMOGENEOUS LOG-GAUSSIAN COX PROCESSES WITH PIECEWISE CONSTANT COVARIATES: A CASE STUDY IN MODELING OF COVID-19 TRANSMISSION RISK IN EAST JAVA," *Stochastic Environmental Research and Risk Assessment*, vol. 38, no. 7, pp. 2891–2901, Jul. 2024, doi: https://doi.org/10.1007/s00477-024-02720-4.

[12] A. S. Fotheringham, R. Crespo, and J. Yao, "GEOGRAPHICAL AND TEMPORAL WEIGHTED REGRESSION (GTWR)," *Geogr Anal*, vol. 47, no. 4, pp. 431–452, 2015, doi: https://doi.org/10.1111/gean.12071.

[13] B. Huang, B. Wu, and M. Barry, "GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION FOR MODELING SPATIO-TEMPORAL VARIATION IN HOUSE PRICES," *International Journal of Geographical Information Science*, vol. 24, no. 3, pp. 383–401, 2010, doi: https://doi.org/10.1080/13658810802672469.

[14] N. Nuramaliyah, A. Saefuddin, and M. N. Aidi, "THE BEST GLOBAL AND LOCAL VARIABLES OF THE MIXED GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION MODEL," *Indonesian Journal of Statistics and Its Applications*, vol. 3, no. 3, pp. 320–330, 2019, doi: https://doi.org/10.29244/ijsa.v3i3.564.

[15] P. R. Arum, R. P. Gautama, and M. Al Haris, "GEOGRAPHICALLY WEIGHTED GENERALIZED POISSON REGRESSION AND GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION MODELING ON PROPERTY CRIME CASES IN CENTRAL JAVA," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 19, no. 3, pp. 1469–1484, Jul. 2025, doi: https://doi.org/10.30598/barekengvol19iss3pp1469-1484.

[16] J. Liu *et al.*, "A MIXED GEOGRAPHICALLY AND TEMPORALLYWEIGHTED REGRESSION: EXPLORING SPATIAL-TEMPORAL VARIATIONS FROM GLOBAL AND LOCAL PERSPECTIVES," *Entropy*, vol. 19, no. 2, 2017, doi: https://doi.org/10.3390/e19020053.

[17] D. L. Sogen, P. R. Arum, and R. Wasono, "PEMODELAN GEOGRAPHICALLY WEIGHTED POISSON REGRESSION (GWPR) DENGAN PEMBOBOT FIXED BISQUARE KERNEL DAN ADAPTIVE BISQUARE KERNEL PADA KASUS DBD DI RIAU," in *Prosiding Seminar Nasional Unimus*, 2023.

[18] P. R. Arum and S. Alfian, "PEMODELAN PERTUMBUHAN EKONOMI DI JAWA BARAT MENGGUNAKAN METODE GEOGRAPHICALLY WEIGHTED PANEL REGRESSION," *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, vol. 15, no. 2, pp. 219–227, Dec. 2022, doi: https://doi.org/10.36456/jstat.vol15.no2.a5506.

[19] P. R. Arum, M. Ridwan, I. Alfidayanti, and R. Wasono, "ROBUST GEOGRAPHICALLY WEIGHTED REGRESSION WITH LEAST ABSOLUTE DEVIATION (LAD) ESTIMATION AND M-ESTIMATION ON GRDP OF WEST JAVA PROVINCE," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 3, pp. 1573–1584, Aug. 2024, doi: https://doi.org/10.30598/barekengvol18iss3pp1573-1584.

[20] A. Avianti and M. Sihaloho, "PERANAN PEKERJA ANAK DI INDUSTRI KECIL SANDAL TERHADAP PENDAPATAN RUMAHTANGGA DAN KESEJAHTERAAN DIRINYA DI DESA PARAKAN, KECAMATAN CIOMAS, KABUPATEN BOGOR, JAWA BARAT," *Sodality: Jurnal Sosiologi Pedesaan*, vol. 1, no. 1, pp. 10–25, 2015.doi: https://doi.org/10.22500/sodality.v1i1.9386