E-Jurnal Matematika Vol. 15(3), Agustus 2026, pp. 137-146 ISSN: 2303-1751
DOI: https://doi.org/10.24843/MTK.2026.v15.i03.p513

EXPLORING SPATIAL HETEROGENEITY OF LIFE EXPECTANCY IN WEST JAVA PROVINCE USING GEOGRAPHICALLY WEIGHTED REGRESSION

Vera Maya Santi[1§], Muhamad Rifan[2], Widyanti Rahayu[3]

[1]Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta [Email: vmsanti@unj.ac.id]
[2]Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta [Email: muhamadrifan1305@gmail.com]
[3]Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta [Email: wrahayu@unj.ac.id]
[§]*Corresponding Author*

ABSTRACT

Life Expectancy Rate (LER) is a key indicator of population health and regional welfare. Significant disparities in LER across districts and municipalities in West Java suggest that the factors influencing life expectancy may vary geographically. Conventional regression models often assume spatial homogeneity and may therefore be inadequate for capturing local variations. This study aims to identify the determinants of LER in West Java by applying Geographically Weighted Regression (GWR) with an adaptive kernel bisquare weighting scheme. Secondary data were obtained from Statistics Indonesia (BPS) of West Java Province and the Open Data Jabar portal. The analysis involved 27 districts and municipalities and examined 11 potential explanatory variables. The results indicate that the influence of predictor variables differs across locations, confirming the presence of spatial heterogeneity in LER determinants. The GWR approach produced 27 local regression models and classified the study area into 10 groups based on similarities in significant factors. Furthermore, the proposed model achieved a coefficient of determination (R^2) of 98.86%, indicating a strong ability to explain regional variations in life expectancy. These findings highlight the importance of location-specific policy interventions to improve public health outcomes across West Java.

Keywords: *Adaptive Kernel Bisquare, Spatial Effect, Geographically Weighted Regression, Life Expectancy Rate*

1. INTRODUCTION

Advancing health initiatives is crucial for attaining the Sustainable Development Goals (SDGs), a pursuit that inherently relies on public engagement in maintaining well-being. Consequently, a population's quality of life serves as a direct indicator of both societal welfare and the efficacy of localized development policies (Bangun, 2019). As a fundamental human right, health significantly shapes the caliber of human resources. A workforce that is both mentally and physically robust is essential for driving economic and social progress. To evaluate this health status and the corresponding effectiveness of state interventions, Life Expectancy, defined as the projected average lifespan of an individual from birth, is frequently utilized as a primary metric. Elevated Life Expectancy generally corresponds to superior public health conditions. Furthermore, literature suggests a strong link between this demographic indicator and regional socioeconomic advancement; areas with robust economic growth typically observe prolonged lifespans (Anggraini & Lisyaningsih, 2010). Thus, upward trends in this metric reflect the triumphant implementation of broader health and economic strategies.

Given its status as Indonesia's most populous province with over 50 million residents in 2024, West Java necessitates extensive developmental planning. Although the province's overall Life Expectancy rose to 74.07 years in 2024 from

73.80 years previously, notable intra-provincial disparities persist. For instance, Bekasi City reported a figure of 76.12 years, contrasting sharply with Tasikmalaya Regency's 70.39 years (Badan Pusat Statistik Provinsi Jawa Barat, 2025). Such variances imply that the determinants influencing longevity differ significantly across localized jurisdictions. Geographically Weighted Regression (GWR) offers a robust analytical framework for examining Life Expectancy by incorporating spatial dynamics. As an advancement of traditional linear regression, GWR integrates spatial coordinates to generate localized parameter estimates (Fotheringham et al., 2002). This capability is critical for addressing Spatial Heterogeneity, enabling the model to account for diverse geographical and environmental influences on longevity.

Several prior investigations have identified various regional determinants of Life Expectancy (Tanadjaja et al., 2017; Nisa & Hakim, 2022; Madeira et al., 2024). Collectively, these studies demonstrate the effectiveness of the GWR method in modeling these regional variations. Building on this foundation, the current study aims to analyze Life Expectancy across 27 regencies and cities in West Java. By employing GWR with an adaptive bisquare kernel weighting function, this study is expected to uncover the significant, localized factors influencing longevity in each area.

2. RESEARCH METHODS

2.1 Identification of Spatial Effects

Spatial effect identification verifies whether geographical properties intrinsically influence the observed variables. Based on Anselin (1988), these spatial dynamics manifest as either spatial dependence or spatial heterogeneity. To systematically detect these phenomena, Moran's I test is utilized to measure spatial autocorrelation, while the Breusch-Pagan test is applied to identify spatial heterogeneity:

2.1.1 Spatial Autocorrelation Test

Spatial autocorrelation in the dataset was examined using Moran's I statistic (Anselin, 1988; Nadya et al., 2017; Wahyudi et al., 2023). The hypotheses used in this test are as follows: $H_0: I = 0$ (there is no spatial autocorrelation) versus $H_1: I \neq 0$ (there is spatial autocorrelation). Decisions can be made based on Moran's I test statistics using the following calculations (Anselin, 1988):

$$Z_{count} = \frac{I - E(I)}{\sqrt{Var(I)}} \qquad (1)$$

The decision made is to reject the null hypothesis at a significance level of α if the value of $|Z_{count}| > Z_{\frac{\alpha}{2}}$ or if the value of p_{value} is less than α (Mukrom et al., 2021).

2.1.2 Spatial Heterogeneity Test

To test for spatial heterogeneity, the Breusch-Pagan (BP) test was conducted with the following hypothesis (Santi et al., 2025b; Anselin, 1988), H_0: No heterogeneity observed versus H_1: Heterogeneity observed. The statistic for this test is as follows:

$$BP = \frac{1}{2}\boldsymbol{f}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{f} \qquad (2)$$

The null hypothesis is retained at significance level α provided the BP statistic is $\leq x_k^2$, or the p_{value} exceeds α. Should the test reject the null hypothesis, it confirms the existence of regional heteroscedasticity, thereby justifying the integration of spatial effects into the analytical framework (Wahyudi et al., 2023; Ramadhani et al., 2020).

2.2 Geographically Weighted Regression

GWR serves as a robust spatial modeling technique specifically designed to address spatial heterogeneity (Fotheringham et al., 2002). The GWR model as described by Fotheringham et al. (2002) is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i\,; \quad i = 1,2,\ldots,n \qquad (3)$$

where y_i is observed value of response variable i, x_{ik} is observed value of predictor variable k at location i, $\beta_0(u_i, v_i)$ is the intercept value at observation i, (u_i, v_i) is geographical coordinates (longitude, latitude) of location i, $\beta_k(u_i, v_i)$ Observed value of predictor variable k at location i, ε_i is error in observation i.

2.2.1 Parameter Estimation (Weighted Least Square)

Parameter estimation in GWR is performed using the Weighted Least Square (WLS) method (Fotheringham et al., 2002). Based on the WLS method, GWR parameter estimation is formulated as follows:

$$\hat{\boldsymbol{\beta}}(u_i, v_i) = [\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} \qquad (4)$$

Since there are n sample locations, the β matrix consists of n sets of local parameters and has the following matrix structure:

$$\boldsymbol{\beta}(u_i, v_i) = \begin{bmatrix} \beta_0(u_1, v_1) & \beta_1(u_1, v_1) & \cdots & \beta_p(u_1, v_1) \\ \beta_0(u_2, v_2) & \beta_1(u_2, v_2) & \cdots & \beta_p(u_2, v_2) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_0(u_n, v_n) & \beta_1(u_n, v_n) & \cdots & \beta_p(u_n, v_m) \end{bmatrix} \quad (5)$$

Based on equation (4), if $\boldsymbol{G}$ is the notation for $[X^T W(u_i, v_i) X]^{-1} X^T W(u_i, v_i)$, then $\hat{\beta}(u_i, v_i)$ can be denoted as:

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = \boldsymbol{GY} \quad (6)$$

Thus, the variance for parameter estimation is denoted by:

$$Var(\widehat{\boldsymbol{\beta}}(u_i, v_i)) = \boldsymbol{GG^T\sigma^2} \quad (7)$$

2.2.2 Bandwidth Selection

The bandwidth delineates a circular radius (h) around a focal observation point (i), postulating that proximate data points exert stronger leverage on parameter estimations (Wahyudi et al., 2023).

One approach commonly used in determining the optimal bandwidth value is the Cross Validation (CV) method, as explained by Caraka & Yasin (2017), formally expressed as follows:

$$CV = \sum_{i=1}^{n} (y_i - \hat{y}_{\neq i}(h))^2 \quad (8)$$

where $\hat{y}_{\neq i}(h)$ is the estimated value of y_i where observations in region i are omitted from the estimation process. The optimum bandwidth (h) will be obtained through an iterative process so that the minimum CV value is obtained.

2.2.3 Model Weighting

Methodologically, configuring these spatial relationships frequently involves applying a kernel function, with the two predominant variations being the Fixed Kernel and the Adaptive Kernel (Fotheringham et al., 2002; Wahyudi et al., 2023).

Table 1. Kernel Function Formula

Kernel Gaussian	Kernel Bi-Square
Fixed	
$w_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right]$	$w_{ij} = \begin{cases}\left[1-\left(\left(\frac{d_{ij}}{b}\right)^2\right)\right]^2, & if\ d_{ij} < h \\ 0, & others\end{cases}$
Adaptive	
$w_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b_i}\right)^2\right]$	$w_{ij} = \begin{cases}\left[1-\left(\left(\frac{d_{ij}}{b_i}\right)^2\right)\right]^2, & if\ d_{ij} < h_i \\ 0, & others\end{cases}$

where:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (9)$$

h_i = bandwidth at observation i

2.2.4 Significance Test of Model Parameters

The corresponding hypotheses are specified as follows: $H_0: \beta_k(u_i, v_i) = 0$, whereas H_1 $H_1: \beta_k(u_i, v_i) \neq 0$, for $i = 1,2, \dots, n; k = 1,2, \dots, p$. The null hypothesis is rejected when $|T_{count}| > t_{\left(\frac{\alpha}{2}; n-p-1\right)}$. T_{count} can be obtained (Caraka & Yasin, 2017):

$$T_{count} = \frac{\widehat{\boldsymbol{\beta}}_{\boldsymbol{k}}(u_i, v_i)}{\hat{\sigma}\sqrt{g_{kk}}} \quad (10)$$

where $\widehat{\boldsymbol{\beta}}_k(u_i, v_i)$ is estimated parameter values at location (u_i, v_i) and g_{kk} is the diagonal element k of the matrix GG^T, where $\boldsymbol{G} = (\boldsymbol{X^T W}(u_i, v_i)\boldsymbol{X})^{-1}\boldsymbol{X^T}\ \boldsymbol{W}(u_i, v_i)$.

2.2.5 Goodness of Fit Test

The goodness-of-fit test assesses how well the model parameters collectively represent the observed data. The hypothesis is as follows: $H_0: \beta_k(u_i, v_i) = \beta_k$, $k = 1,2,3, \dots, p$ and $i = 1,2,3, \dots, n$ versus H_1: at least one $\beta_k(u_i, v_i) \neq \beta_k$. The decision criterion for this test is to reject H_0 if $F^* \geq F_{tabel}$ (Fotheringham et al., 2002; Santi et al., 2024; Santi et al., 2025a; Santi et al., 2026):

$$F^* = \frac{SSE(H_0)/df_1}{SSE(H_1)/df_2} \quad (11)$$

$$\boldsymbol{S} = \begin{bmatrix} x_1^T[\boldsymbol{X^T W}(u_1, v_1)\boldsymbol{X}]^{-1}\boldsymbol{X^T W}(u_1, v_1) \\ x_2^T[\boldsymbol{X^T W}(u_2, v_2)\boldsymbol{X}]^{-1}\boldsymbol{X^T W}(u_2, v_2) \\ \vdots \\ x_n^T[\boldsymbol{X^T W}(u_n, v_n)\boldsymbol{X}]^{-1}\boldsymbol{X^T W}(u_n, v_n) \end{bmatrix}$$

$\boldsymbol{S}$ is the projection matrix of the GWR model, which is a matrix that projects the value of y into $\hat{y}$ at the location (u_i, v_i) (Caraka & Yasin, 2017).

2.3 Research Data

Secondary data used in this study were collected from the BPS of West Java Province and Open Data Jabar. The gathered information focuses on Life Expectancy alongside the various variables presumed to affect it. In total, the dataset comprises 27 regional observations, specifically covering 18 regencies and 9 cities within West Java Province. Additionally, the required spatial data takes the form of shapefiles acquired from https://www.indonesia-geospasial.com/.

2.4 Research Variables

These independent variables were carefully selected based on a comprehensive review of relevant prior literature, as detailed in Table 2.

Table 2. Research Variables

Symbol	Scale	Description
Y	Ratio	Life Expectancy
X_1	Ratio	Percentage of Poor Population
X_2	Ratio	Midwife Ratio per 10,000 Population
X_3	Ratio	Expected Years of Schooling
X_4	Ratio	Percentage of Illiteracy
X_5	Ratio	Percentage of Unemployment Rate
X_6	Ratio	Percentage of Population with Improved Sanitation Services
X_7	Ratio	Percentage of Households with Clean and Healthy Living Behaviors
X_8	Ratio	Population Density
X_9	Ratio	Number of Low Birth Weight Babies
X_{10}	Ratio	Number of BCG Immunizations in Babies
X_{11}	Ratio	Number of Stunted Toddlers

In addition to the response variable and predictor variable, there are two geographic variables, namely latitude and longitude coordinates (u_i, v_i) regarding the location of regencies/cities in West Java.

2.5 Analysis Procedures

The following are the analysis steps for modeling Life Expectancy in West Java using GWR:

1. Create a spatial distribution map of life expectancy across 27 regencies/cities in West Java and perform a descriptive analysis of the response variable and all predictor variables.
2. Multiple Linear Regression Analysis (Global Model).
3. Testing Classical Assumptions.
   a. Testing residual normality using the Anderson Darling test.
   b. Detection of multicollinearity with VIF values.
   c. Assess error variance homogeneity by examining the residuals versus fitted response value plot.
4. Identify the presence of spatial effects in the observed data.
5. GWR Modeling using equation (3).
6. GWR Procedure: compute Euclidean distances among locations, select optimal bandwidth via CV, construct adaptive bisquare kernel weights, estimate parameters using WLS, assess model fit, perform parameter significance tests, and interpret predictor effects.
7. Classifying regencies/cities according to significant variable influence patterns.

3. RESULTS AND DISCUSSION

3.1 Data Description

West Java Province was administratively divided into 27 regencies/cities. Descriptive statistics, specifically the minimum, maximum, mean, median, and standard deviation, were employed to summarize the study's dataset. This dataset comprises a single response variable, Life Expectancy (Y), alongside eleven predictor variables: the Percentage of Poor Population (X_1), Midwife Ratio per 10,000 Population (X_2), Expected Years of Schooling (X_3), Percentage of Illiteracy (X_4), Percentage of Unemployment Rate (X_5), Percentage of Population with Improved Sanitation Services (X_6), Proportion of Households Practicing Clean and Healthy Living (X_7), Population Concentration (X_8), Number of Low Birth Weight Babies (X_9), Number of BCG Immunizations in Babies (X_{10}), and Number of Stunted Toddlers (X_{11}). Summary statistics of all study variables are presented on Table 3.

E-Jurnal Matematika Vol. 15(3), Agustus 2026, pp. 137-146 ISSN: 2303-1751
DOI: https://doi.org/10.24843/MTK.2026.v15.i03.p513

Table 3. Descriptive Statistics of Research Variables

Variable	Minimum	Maximum	Average	Median	Std. Deviation
Y	70.39	76.12	73.26	73.16	1.44
X_1	3.99	15.28	8.48	7.74	2.91
X_2	2.34	11.93	8.01	8.41	2.65
X_3	12.00	14.30	12.92	12.74	0.76
X_4	0.26	6.74	1.46	0.83	1.59
X_5	1.58	8.97	6.53	6.73	1.71
X_6	85.37	99.85	95.10	96.76	4.22
X_7	45.48	84.67	67.87	68.00	10.03
X_8	396.00	15,557.00	3,675.00	1,365.00	4,617.51
X_9	124.00	1,967.00	739.70	518.00	542.96
X_{10}	464.00	21,806.00	5,631.00	3,641.00	5,906.03
X_{11}	2,337.00	89,595.00	26,969.00	23,786.00	19,837.09

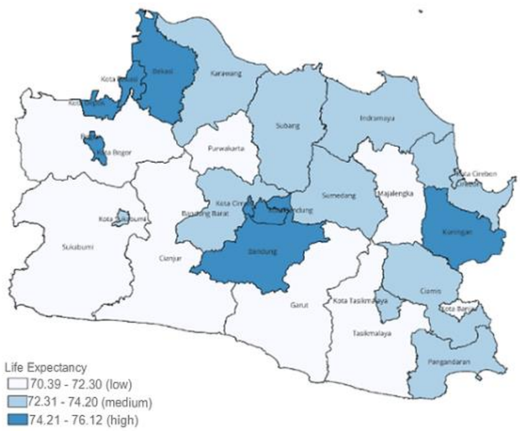

Figure 1. Distribution of Life Expectancy Values in West Java

Table 3 demonstrates that Life Expectancy ranges from 70.39 to 76.12 years. Tasikmalaya Regency recorded the lowest Life Expectancy across the region, whereas Bekasi City achieved the highest. A standard deviation of 1.44 was observed, which is substantially lower than the mean Life Expectancy, suggesting a homogeneous data distribution.

Figure 1 presents a map of Life Expectancy values in West Java in 2024. The grouping of regencies/cities in West Java based on Life Expectancy values was done by classifying Life Expectancy values into three categories. Regions with higher Life Expectancy values are represented by dark blue shading, whereas lower values are indicated in white.

3.2 Multiple Linear Regression Analysis

As an initial evaluation, multiple linear regression utilizing OLS estimation was performed to identify explanatory variables significantly impacting Life Expectancy in West Java, prior to incorporating spatial elements. The model obtained from the multiple linear regression analysis is as follows:

$$\begin{aligned}\hat{y} = 56.1955 - {} & 0.219x_1 + 0.048x_2 \\ & + 0.305x_3 - 0.279x_4 \\ & + 0.127x_5 + 0.227x_6 \\ & + 0.013x_7 + 0.0001x_8 \\ & - 0.002x_9 + 0.0001x_{10} \\ & - 0.000001x_{11}\end{aligned}$$

The overall model was statistically significant, as indicated by an F-statistic of 8.83 and a p-value of 0.0001041 (< 0.05). This result confirms that the explanatory variables jointly contribute to variations in Life Expectancy across West Java. The model produced an R^2 value of 0.8662 and an adjusted R^2 of 0.7641, indicating that 86.62% of the variability in Life Expectancy is explained by the selected predictors, whereas the remaining 13.38% is associated with factors outside the scope of the model.

Next, a partial significance test was conducted to assess the effect of each explanatory variable on the response variable. The results are summarized in Table 4.

Table 4. Partial Significance Test Results

Variable	t_{count}	p_{value}
Intercept	1.847	0.085
X_1	-3.096	0.007
X_2	0.452	0.657
X_3	-0.693	0.499
X_4	-2.12	0.051
X_5	1.056	0.308
X_6	3.71	0.002
X_7	0.684	0.504
X_8	1.018	0.325
X_9	-2.747	0.015
X_{10}	2.378	0.031
X_{11}	0.101	0.921

Based on the results in Table 4 with a α of 5%, it shows that H_0 is rejected for variables X_1, X_6, X_9, X_{10} because the $p_{value} < \alpha$, which means that the variables Percentage of Poor Population, Percentage of Population with Improved Sanitation Services, Number of Low Birth Weight Babies, and Number of BCG Immunizations in Babies significantly affect Life Expectancy in West Java. Meanwhile, for variables $X_2, X_3, X_4, X_5, X_7, X_8, X_{11}$ do not significantly affect Life Expectancy in West Java.

3.3 Assumption Testing

Assumption testing in regression is used to examine the quality of the data so that it can be analyzed further.

3.3.1 Assumption of Normality

The Anderson–Darling test yielded a p-value of 0.4959 (> 0.05), indicating that the residuals followed a normal distribution and the normality assumption was met.

3.3.2 Assumption of Multicollinearity

Multicollinearity detection was performed to see whether the explanatory variables were correlated with other explanatory variables. Multicollinearity is considered present when the VIF exceeds 10 (Santi et al., 2025c). The VIF test results are as follows.

Table 5. Multicollinearity Test

Variable	VIF
X_1	2.312
X_2	4.343
X_3	6.018
X_4	2.364
X_5	2.287
X_6	3.618
X_7	1.879
X_8	3.020
X_9	5.111
X_{10}	2.969
X_{11}	3.382

According to Table 5, indicating no multicollinearity among the explanatory variables.

3.3.3 Assumption of Homogeneity of Variance

Homogeneity testing assesses the consistency of residual variance across observations. The assumption of homogeneity of error variance can be examined by looking at the Scale Location plot in Figure 2.

Figure 2 illustrates a downward trend in the red line corresponding with rising fitted values, alongside an uneven distribution of residuals that cluster toward the center and right. This pattern signifies a violation of the homoscedasticity assumption.

3.4 Identification of Spatial Effects

Spatial effects were identified to see whether the data used met spatial effects, indicating that the ordinary multiple linear regression model was less able to explain existing local variations. The identification of spatial effects was carried out using two tests. The results of the two tests are as follows.

3.4.1 Spatial Autocorrelation Test

Spatial autocorrelation was assessed using Moran's I test. The results of the Moran's I test are presented in Table 6.

Table 6. Spatial Autocorrelation Test

***Moran's I* statistic standard deviate**	p_{value}
-1,6285	0,9483

E-Jurnal Matematika Vol. 15(3), Agustus 2026, pp. 137-146 ISSN: 2303-1751
DOI: https://doi.org/10.24843/MTK.2026.v15.i03.p513

Moran's I test resulted in a p-value of 0.9483 (> 0.05), indicating that Life Expectancy across regencies and cities in West Java did not exhibit significant spatial autocorrelation.

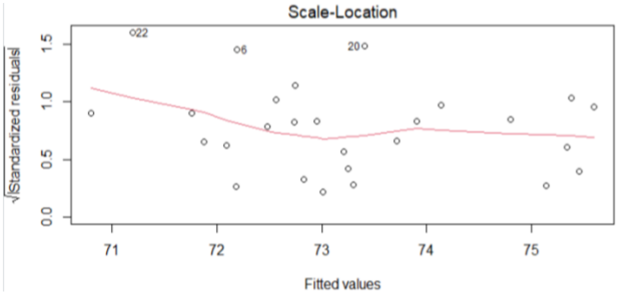

Figure 2. Plot of Homogeneity of Variance Test

3.4.2 Spatial Heterogeneity Test

The Breusch–Pagan (BP) test was performed, and the results are presented in Table 7.

Table 7. Spatial heterogeneity Test

Statistik Uji BP	p_{value}
21,925	0,025

The obtained p-value of 0.025 was below the 5% significance level, indicating rejection of the null hypothesis. This result suggests the presence of spatial heterogeneity, meaning that residual variances differ across observation locations.

The results of spatial effect identification show that there is no spatial autocorrelation but there is heterogeneity of residual variance between observation locations. For this reason, the results of the spatial effect identification can be used as a basis for spatial modeling, namely Geographically Weighted Regression (GWR).

3.5 GWR Modeling

The first step is to determine the distance matrix between 27 regencies/cities using the Euclidean distance formula. The resulting distance matrix is 27 × 27 as follows:

$$\boldsymbol{D}_{27\times27} = \begin{bmatrix} 0 & 0.284 & \dots & 0.666 \\ 0.284 & 0 & \dots & 0.0001 \\ \vdots & \vdots & \ddots & \vdots \\ 0.666 & 0.046 & \dots & 0 \end{bmatrix}_{27\times27}$$

Subsequently, the optimal bandwidth was derived utilizing Cross-Validation (CV) criteria coupled with an adaptive kernel bisquare function, resulting in variable bandwidths tailored to each specific location.

Table 8. Bandwith Value

Regency/City	Bandwith Value
Bandung City	0.9287
Bandung Regency	0.9985
Banjar City	16.709
Bekasi City	16.439
Bekasi Regency	15.240
Bogor City	16.152
Bogor Regency	16.601
Ciamis Regency	15.123
Cianjur Regency	11.705
Cimahi City	0.8637
Cirebon City	14.753
Cirebon Regency	14.740
Depok City	17.093
Garut Regency	10.353
Indramayu Regency	12.104
Karawang Regency	12.953
Kuningan Regency	14.994
Majalengka Regency	11.935
Pangandaran Regency	17.645
Purwakarta Regency	11.107
Subang Regency	0.9711
Sukabumi City	13.470
Sukabumi Regency	15.786
Sumedang Regency	10.014
Tasikmalaya City	13.776
Tasikmalaya Regency	13.875
West Bandung Regency	0.9369

Using the distance matrix and optimal bandwidth, the spatial weighting matrix for the 27 regencies/cities is constructed. This results in a 27×27 order matrix for the life expectancy data in West Java, as follows:

$$\boldsymbol{W}_{27\times27} = \begin{bmatrix} 1 & 0.5120 & \dots & 0.1105 \\ 0.4851 & 1 & \dots & 0.0001 \\ \vdots & \vdots & \ddots & \vdots \\ 0.4280 & 0.2585 & \dots & 1 \end{bmatrix}_{27\times27}$$

3.5.1 Goodness of Fit Test

A goodness-of-fit evaluation was performed to ascertain the overall adequacy of the estimated parameter coefficients. The results are presented in Table 9:

Table 9. Goodness of Fit Test

Source of Variance	Df	Sum Sq	Mean Sq	F^*	F_{table}
OLS Residuals	12.000	7.193			
GWR Model	11.994	6.583	0.549		
GWR Residuals	3.007	0.610	0.203	2.706	2.687

Table 9 shows that the F^* value is 2.706. The F_{table} value obtained using a significance level of 5% is 2.687. Consequently, the null hypothesis is rejected $(F^* > F_{table})$, establishing that the GWR approach yields a significantly different and superior fit compared to the global regression model.

3.5.2 Model Significance Test

A significance level of 5% was used, with parameter testing performed using the T-test to show differences in characteristics between districts/cities based on factors affecting Life Expectancy. This is done by looking at the t-test statistical value. If the predictor variable has $|T_{count}| > t_{(0.025;25)} = 2.0595$, then reject the null hypothesis, which means that the variable affects the model. The following are the significant variables in each Regency/City:

Table 10. Parameter Significance Test Results

No.	Regency/City	Significance
1.	Bandung City	$X_1, X_4, X_6, X_7, X_8, X_9$
2.	Bandung Regency	$X_1, X_5, X_6, X_9, X_{10}$
3.	Banjar City	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_9, X_{10}, X_{11}$
4.	Bekasi City	X_1, X_6, X_9, X_{10}
5.	Bekasi Regency	X_1, X_6, X_9, X_{10}
6.	Bogor City	X_1, X_6, X_9, X_{10}
7.	Bogor Regency	X_1, X_6, X_9, X_{10}
8.	Ciamis Regency	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_9, X_{10}, X_{11}$
9.	Cianjur Regency	X_1, X_6, X_9, X_{10}
10.	Cimahi City	X_1, X_6, X_7, X_8, X_9
11.	Cirebon City	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_{10}$
12.	Cirebon Regency	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_{10}, X_{11}$
13.	Depok City	X_1, X_6, X_9, X_{10}
14.	Garut Regency	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_9, X_{10}, X_{11}$
15.	Indramayu Regency	$X_1, X_2, X_4, X_5, X_6, X_8$
16.	Karawang Regency	X_1, X_6, X_9, X_{10}
17.	Kuningan Regency	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_9, X_{10}, X_{11}$
18.	Majalengka Regency	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_{10}, X_{11}$
19.	Pangandaran Regency	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_9, X_{10}, X_{11}$
20.	Purwakarta Regency	$X_1, X_5, X_6, X_9, X_{10}$
21.	Subang Regency	$X_1, X_4, X_5, X_6, X_8, X_9, X_{11}$
22.	Sukabumi City	X_1, X_6, X_9, X_{10}
23.	Sukabumi Regency	X_1, X_6, X_9, X_{10}
24.	Sumedang Regency	$X_1, X_2, X_4, X_6, X_8, X_{10}$
25.	Tasikmalaya City	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_9, X_{10}, X_{11}$
26.	Tasikmalaya Regency	$X_1, X_2, X_3, X_4, X_6, X_7, X_8, X_9, X_{10}, X_{11}$
27.	West Bandung Regency	X_1, X_6, X_9, X_{10}

E-Jurnal Matematika Vol. 15(3), Agustus 2026, pp. 137-146 ISSN: 2303-1751
DOI: https://doi.org/10.24843/MTK.2026.v15.i03.p513

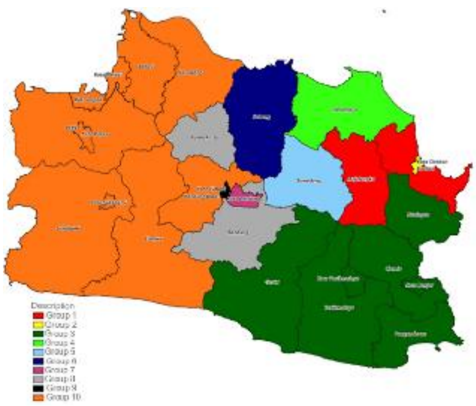

Figure 3. Regional Grouping

As displayed in Table 10 and Figure 3, regions can be classified into distinct clusters based on the specific variables driving their Life Expectancy. Locations sharing identical significant predictors are aggregated into the same group.

3.5.3 Selection of the Best Model

Model selection was predicated on comparing the Akaike Information Criterion (AIC) (Santi et al., 2024; Santi et al., 2026) and R^2 (Santi et al., 2025a) metrics between the OLS and GWR frameworks. The optimal model minimizes the AIC while maximizing the R^2 value (Santi et al., 2025a). The AIC and R^2 values for both models are presented in Table 9.

Table 9. Comparison of the Best Model Criteria Values

Method	AIC	R^2
Ordinary Least Square	66.9077	0.7681
Geographically Weighted Regression	-3.5301	0.9886

Table 9 shows that the GWR method has an AIC value of -3.53, which is smaller than the OLS AIC value of 66.91. The R^2 value in the GWR method is 0.989. These results demonstrate that the Geographically Weighted Regression (GWR) framework successfully captures 98.86% of the variance in West Java's Life Expectancy, leaving a mere 1.14% to unobserved factors. Based on the AIC and R square values obtained, it is evident that the best method to use is GWR, where the parameter estimator used is Weighted Least Squares.

4. CONCLUSIONS

This study demonstrates that the Geographically Weighted Regression (GWR) model effectively captures spatial heterogeneity in Life Expectancy Rate (LER) across West Java. The analysis produced 27 local regression models, indicating that the factors influencing life expectancy vary among regencies and cities. Using the Adaptive Bisquare Kernel, the regions were further classified into 10 groups based on similarities in their significant explanatory variables.

The results reveal that the percentage of poor population and the percentage of population with access to improved sanitation services consistently influence LER across all regions. However, additional significant determinants differ geographically, highlighting the importance of location-specific policy interventions to improve life expectancy outcomes.

Future studies are encouraged to compare alternative spatial weighting schemes, such as Adaptive Gaussian, Fixed Gaussian, and Fixed Bisquare kernels, to assess their relative effectiveness and enhance the robustness of spatial modeling results.

REFERENCES

Anggraini, E., U. Lisyaningsih. 2010. Disparitas Spasial Angka Harapan Hidup di Indonesia Tahun 2010. *Jurnal Bumi Indonesia* 3(2), p.71–80.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Springer Netherlands, Dordrecht.

Badan Pusat Statistik Provinsi Jawa Barat. 2025. Angka Harapan Hidup (Hasil SP2010) Berdasarkan Kabupaten/Kota di Jawa Barat. *Open Data Jabar*. Tersedia pada: https://opendata.jabarprov.go.id/id/dataset/usia-harapan-hidup-menggunakan-sp2010-berdasarkan-kabupatenkota-di-jawa-barat.

Bangun, R.H. 2019. Analisis Determinan Angka Harapan Hidup Kabupaten Mandailing Natal. *Jurnal Kuntansi & Ekonomi FE UN PGRI Kediri* 4(3), p.22–31.

Caraka, R.E., H. Yasin. 2017. *Geographically Weighted Regression (GWR): Sebuah Pendekatan Regresi Geografis*. Mobius, Yogyakarta.

Fotheringham, A.S., C. Brunsdon, M. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*.

Madeira, I.D.J., P.S. Syukur, R.M. Istiawati, G. Anuraga. 2024. Pemodelan Angka Harapan Hidup di Indonesia Menggunakan Geographically Weighted Regression (GWR). *Seminar Nasional Hasil Riset dan Pengabdian* 6, p.444–450.

Mukrom, M.H., H. Yasin, A.R. Hakim. 2021. Pemodelan Angka Harapan Hidup Provinsi Jawa Tengah Menggunakan Robust Spatial Durbin Model. *Jurnal Gaussian* 10(1), p.44–54.

Nadya, M., W. Rahayu, V.M. Santi. 2017. Analisis Geographically Weighted Regression (GWR) Pada Kasus Pneumonia Balita di Provinsi Jawa Barat. *Jurnal Statistika dan Aplikasinya*.

Nisa, R.F., A.R. Hakim. 2022. Pemodelan Mixed Geographically Weighted Regression dengan Adaptive Bandwidth untuk Angka Harapan Hidup (Studi Kasus: Angka Harapan Hidup di Jawa Tengah). *Jurnal Gaussian* 11(1), p.67–76.

Ramadhani, E., N. Salwa, M.S. Mazaya. 2020. Identifikasi Faktor-Faktor yang Memengaruhi Angka Harapan Hidup di Sumatera Tahun 2018 Menggunakan Analisis Regresi Spasial Pendekatan Area. *Journal of Data Analysis* 3(2), p.62–75.

Santi, V.M., I. Hasari, D. Handayani. 2024. Linear Mixed Models to Analyze Indonesia's PISA Reading Literacy Score. *AIP Conference Proceedings* 2982(1).

Santi, V.M., A.M. Riyantobi, W. Rahayu. 2025a. Analysis of Indonesian Students' Reading Literacy Using The Smoothly Clipped Absolute Deviation (SCAD) Penalty. *Jurnal Statistika dan Aplikasinya* 9(1), p.126–139.

Santi, V.M., D. Indriana, F. Ladayya, Tonah. 2025b. Linear Mixed Model-Lasso with MLE and REML Estimation on Poverty Data in Java Island. *Jurnal Statistika dan Aplikasinya* 9(2), p.14–24.

Santi, V.M., K. Helmina, Y. Mahatma. 2025c. Pemodelan Penyakit Tuberkulosis di Jawa Barat Menggunakan Linear Mixed Model Dengan Penalti Smoothly Clipped Absolute Deviation. *Jurnal Riset dan Aplikasi Matematika (JRAM)* 9(2), p.150–159.

Santi, V.M., K.A. Notodiputro, Indahwati, B. Sartono, B. Susetyo, A. Kurnia. 2026. Two-Level Random Effects in Multivariate Linear Mixed Model for Longitudinal Data. *Engineering Letters* 34(1), p.368–379.

Tanadjaja, A., I. Zain, W. Wibowo. 2017. Pemodelan Angka Harapan Hidup di Papua dengan Pendekatan Geographically Weighted Regression. *Institut Teknologi Sepuluh Nopember*, Surabaya.

Wahyudi, R., Y. Fauzi, J. Rizal. 2023. Analisis Kemiskinan Ekstrem Provinsi Bengkulu Menggunakan Metode Geographically Weighted Regression (GWR) dengan Pembobot Adaptive Gaussian Kernel dan Adaptive Bi-Square. *Journal of Mathematics Universitas Negeri Padang* 8(2), p.134–149.