



Geographically Weighted Regression Modeling Using Fixed and Adaptive Kernel Weights for the Human Development Index Case in West Java Province

Karin Dita Widiyanti¹, Alwan Fadlurohman², Dannu Purwanto³
^{1,2,3} Universitas Muhammadiyah Semarang, Indonesia

DOI: <https://doi.org/10.26714/jodi.v3i2.887>

Article Information

Article information:

Submitted 22 November 2025

Revised 15 December 2025

Accepted 29 December 2025

Keywords:

Human Development Index;
Geographically Weighted
Regression; West Java

Abstract

This study investigates spatial heterogeneity in the determinants of the Human Development Index (HDI) across 27 districts and municipalities in West Java Province using the Geographically Weighted Regression (GWR) framework. The analysis incorporates four socio-economic variables—Open Unemployment Rate (TPT), School Participation Rate for ages 16–18 (APS 16–18), Population Density, and Gross Regional Domestic Product at constant prices (PPK). Preliminary diagnostics from the global OLS model confirm the absence of heteroscedasticity and multicollinearity, supporting the validity of further spatial modeling. Several GWR specifications were evaluated, and the adaptive Gaussian kernel emerged as the most robust, yielding the lowest AIC and superior explanatory power compared to the global model. Local parameter estimates reveal clear spatial non-stationarity: APS 16–18 and Population Density exhibit consistently significant effects in many regions, PPK shows localized significance, and TPT remains insignificant throughout the study area. These findings underscore the advantage of GWR in capturing spatially varying relationships and emphasize the need for region-specific development strategies.

✉ Corresponding Author:
E-mail: karinndw23@gmail.com

e-ISSN: 2988 - 2109

I. INTRODUCTION

The success of a country can be measured using the Human Development Index (HDI). Human development is a concept first introduced by the United Nations Development Programme (UNDP) in 1990 through the publication of the Human Development Report. UNDP states that the main goal of development is to create an enabling environment that allows people to live long, healthy, and productive lives. There are three dimensions that form the basis of human development: long and healthy life, knowledge, and a decent standard of living (UNDP, 1990).

One of the models that can be applied to identify factors affecting HDI is spatial regression analysis. Spatial regression is an extension of simple linear regression. Several assumptions must be met in a linear regression model, including normality, homoscedasticity, non-multicollinearity, and non-autocorrelation (Alvin C. Rencher and G. Bruce Schaalje, 2008). Spatial development influences the analyzed data. Spatial data refers to information related to the position, objects, and their relationships within the Earth's space. Spatial data is a type of information that includes elements of the Earth, such as its surface, subsurface, waters, oceans, and the lower atmosphere (A. Rajabifard and I. Williamson, 2001). Therefore, the author is interested in modelling HDI using a point-based approach through Geographically Weighted Regression (GWR).

Research related to the Geographically Weighted Regression (GWR) model has been conducted by Andriani, who examined the modelling of the number of poor people in Central Java using GWR. The study found that the factors influencing the number of poor people include the Human Development Index (HDI), the Open Unemployment Rate (TPT), and the District/City Minimum Wage (UMK). In this study, the GWR model was more effective in explaining the number of poor people in Central Java, as indicated by the R^2 and MAPE values, and the resulting model varied across regions.

Several researchers Agustina (2015), K. Amelia (2020), Lutfiani (2019), and Ispriyanti (2005) have investigated GWR models using the Gaussian kernel weighting function. In GWR, determining the weighting scheme is crucial for calculating the value at each location. The fixed kernel model in GWR, consisting of the fixed Gaussian kernel and fixed bi-square kernel, has a significant influence on estimating nearby locations. Kernel weighting functions are one of several approaches to constructing the GWR weighting matrix. For data smoothing, kernel sensitivity functions are often used by incorporating appropriate weights whose values depend on the characteristics of the data. In some studies, the Gaussian kernel weighting function is used because the continuous distance values between observation locations allow for more satisfactory analysis.

Based on this background, this study aims to model the Human Development Index in West Java Province using the GWR method by comparing the performance of fixed and adaptive kernels. By identifying the best model and the dominant spatial factors influencing HDI in the province, the findings are expected to serve as a foundation for regional development decision-making based on spatial data.

II. METHODS

2.1. Human Development Index

The Human Development Index (HDI) was introduced by UNDP in 1990 as a comprehensive measurement tool to assess the quality of human development, focusing on three key dimensions: long and healthy life, knowledge, and a decent standard of living (UNDP, 1990). This measurement was developed as an alternative to purely economic indicators such as per capita income, which are considered insufficient to fully reflect societal well-being (Haq, 1996).

HDI consists of three components: life expectancy for the health dimension; average years of schooling and expected years of schooling for the education dimension; and real expenditure per capita for the standard of living dimension (BPS, 2023). These indicators are used because they represent the fundamental capabilities required for individuals to live productive and dignified lives (UNDP, 1990).

2.2. Multiple Linear Regression

Multiple linear regression is an analysis used to estimate or predict the value of a response variable Y based on measurements of several predictor variables X_1, X_2, \dots, X_r (Walpole, 1993). The multiple linear regression equation can be written as follows (Drapper, N.R., dan Smith H. , 1992):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

Where $\beta_1, \beta_2, \dots, \beta_k$ are parameters (regression coefficients) that must be estimated, and ε is the error term [$\varepsilon \sim N(0, \sigma^2)$]. The equation above can be expressed in matrix form as follows:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_3 \end{pmatrix}$$

In its simplified form, the equation above can be denoted as follows:

$$Y = X\beta + \varepsilon$$

2.3. Spatial Data

Spatial data refers to data associated with the position, objects, and the relationships among them within the Earth's space. Spatial data is a type of information that includes details about the Earth, such as its surface, subsurface, waters, oceans, and the lower atmosphere (A. Rajabifard and I. Williamson, 2001). Spatial data and its derived information are used to determine the position and identification of elements on the Earth's surface (Radjabidfard, and Abbas., 2001).

2.4. Geographically Weighted Regression Model

Geographically Weighted Regression (GWR) is a method used to estimate regression parameters in data that exhibit spatial heterogeneity or variability. According to Colombia Public Health (2019), GWR is also used to visualize the relationship between independent and dependent variables across space. The GWR model analyzes each observation point by incorporating its spatial coordinates (longitude and latitude). The Geographically Weighted Regression (GWR) model is expressed as follows:

$$Y_i = \beta_0(u_i, v_j) + \sum_{k=1}^p \beta_k(u_i, v_j) X_{ik} + \varepsilon_i ; i = 1, 2, \dots, n$$

Where (u_i, v_j) represents the coordinate points (longitude and latitude) of each observation i , and $\beta_k(u_i, v_j)$ is the regression coefficient of the k -th independent variable at location (u_i, v_j) .

2.5. Parameter Estimation of the Geographically Weighted Regression Model

Parameter estimation in the GWR model can be carried out using the Weighted Least Squares (WLS) method. This method assigns weights to the parameter estimation through a spatial weighting matrix. The first step in WLS is constructing a spatial weighting matrix with dimensions $m \times m$ at location i , formulated as follows. According to Calton (2009), observational units that are geographically close to one another receive larger weights than units that are farther apart. The parameter estimation for the k variable at each observation location using GWR is expressed as follows:

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y$$

2.6. Weighting in Geographically Weighted Regression

Spatial weighting functions can be determined using kernel functions. In general, there are two types of kernel functions: Fixed Kernel and Adaptive Kernel.

A. Fixed Kernel Function

A fixed kernel function uses the same bandwidth for every observation location. There are three types of fixed kernel functions, namely:

1. Fixed Kernel Gaussian

$$W_{ij}(u_i, v_i) = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{b} \right)^2 \right]$$

2. Fixed Kernel Bisquare

$$W_{ij}(u_i, v_i) = \left\{ 1 - \left(\frac{d_{ij}}{b} \right)^2 \right\}^2, d_{ij} < b \text{ 0, lainnya}$$

3. Fixed Kernel Tricube

$$W_{ij}(u_i, v_i) = \left\{ 1 - \left(\frac{d_{ij}}{b} \right)^3 \right\}^3, d_{ij} < b \text{ 0, lainnya}$$

B. Adaptive Kernel Function

An adaptive kernel function has a varying bandwidth for each observation point. This is due to the adaptive kernel's ability to adjust according to the location of each observation point. There are three types of adaptive kernel functions, namely:

1. Adaptive Kernel Gaussian

$$W_{ij}(u_i, v_i) = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{b} \right)^2 \right]$$

2. Adaptive Kernel Bisquare

$$W_{ij}(u_i, v_i) = \left\{ 1 - \left(\frac{d_{ij}}{b_{i(q)}} \right)^2 \right\}^2, d_{ij} < b_{i(q)} \text{ 0, lainnya}$$

3. Adaptive Kernel Tricube

$$W_{ij}(u_i, v_i) = \left\{ 1 - \left(\frac{d_{ij}}{b_{i(q)}} \right)^3 \right\}^3, d_{ij} < b_{i(q)} \text{ 0, lainnya}$$

2.7. Bandwidth Geographically Weight Regression

In this modeling process, the optimal bandwidth is determined using cross-validation (CV). Cross-validation is one of the methods used to calculate the optimal window width. The optimal bandwidth is the one that produces the minimum cross-validation value, calculated using the following formula:

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2$$

where $\hat{y}_{\neq i}(b)$ is the fitted value obtained by excluding the observation at location i from the prediction process. The optimal bandwidth is obtained through an iterative procedure until the minimum CV value is reached.

2.8. Parameter Testing in the GWR Model

Parameter testing in the GWR model consists of evaluating the model's goodness of fit and conducting partial tests for the GWR parameters.

1. Simultaneous Test

The simultaneous test is conducted to determine whether all independent variables collectively have a significant effect on the dependent variable. The hypotheses used in this test are as follows:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0 \text{ (no simultaneous effect)}$$

$$H_1 : \text{at least one } \beta_i \neq 0, \text{ for } i = 1, 2, \dots, k \text{ (there is a simultaneous effect).}$$

2. Partial Test of the GWR Model

This test aims to determine which parameters significantly affect the response variable. The partial test uses the following hypotheses:

$$H_0 : \beta_{kh}(u_i, v_i) = 0$$

$$H_1 : \beta_{kh}(u_i, v_i) \neq 0 \text{ for all } k \text{ and } h, \text{ where } k = 1, 2, \dots, q$$

Decision criteria:

If $|t_{hit}| > t_{\frac{\alpha}{2}(n-p-1)}$ or $p\text{-value} < \alpha$ then H_0 is rejected. This means that the parameter is not significant partially (Harini, S., Purhadi, Mashuri, M., & Sunaryo, S., 2012) (Saputri, S., Ispriyanti, D., & Wuryandari, T., 2015).

2.9. Selection of the Best Model

In selecting the best GWR model, two parameters are used: the Akaike Information Criterion (AIC) and the coefficient of determination (R^2).

A. Akaike Information Criterion (AIC)

Selecting the best model is an evaluation process used to determine how well each constructed model fits the data. In addition to determining the optimal bandwidth, AIC can also be used for model selection to identify the most suitable model. The best model is the one with the smallest AIC value. The AIC value is calculated using the following formula:

$$AIC = 2n \log(\hat{\sigma}) + n \log(2\pi) + n \text{tr}(L)$$

Where:

$\hat{\sigma}$ = estimator of the residual standard deviation

L = projection matrix, where $\hat{y} = Ly$

B. Coefficient of Determination (R^2)

The coefficient of determination is used to measure the proportion of variation in the observed data that can be explained by the model in a classical linear regression framework. A small or near-zero (R^2) value indicates that the model has very limited ability to explain the dependent variable. The (R^2) value in the GWR model can be determined using the following equation:

$$R^2 = \frac{SST_{GWR} - SSE_{GWR}}{SST_{GWR}}$$

2.10. Data Source and Research Variables

This study uses data from 27 regencies and cities in West Java Province for the year 2024. All data were obtained from the Badan Pusat Statistik of West Java Province, which can be accessed online through the website <https://jabar.bps.go.id/id>

The variables analyze represent the socioeconomic welfare level in West Java Province. The study includes variables such as Per Capita Income (constant prices), Population Density, Open Unemployment Rate (TPT), and School Participation Rate (APS). In addition, the coordinates u (latitude) and v (longitude) are used for spatial regression analysis under the GWR method.

2.11. Analysis Method

The analysis methods used in this study are as follows:

1. Collecting dependent and independent variable data from 27 regencies/cities in West Java.
2. Conducting descriptive statistical analysis for each variable.
3. Performing assumption tests on the variables before constructing a multivariate regression model, including tests for independence, multivariate normality, and multicollinearity.
4. After the regression model is established, performing residual assumption tests and hypothesis testing to assess the overall significance of the model.
5. Determining the spatial weighting function by testing various kernel combinations (Gaussian, Bisquare, Tricube) using both fixed and adaptive bandwidth approaches, and calculating the optimal bandwidth using the Cross-Validation (CV) method.
6. Selecting the best kernel type and computing the optimal bandwidth based on the smallest AIC value, then constructing the weighting matrix based on Euclidean distances for each location.

7. Conducting Geographically Weighted Regression parameter hypothesis tests simultaneously and partially at each location to determine the local significance of each independent variable.
8. Grouping regencies/cities based on variables that are significant for each dependent variable and presenting the results in thematic maps to visualize the resulting spatial patterns.
9. Visualizing the model results in the form of thematic maps.
10. Interpreting the spatial patterns of the influence of variables on the Human Development Index (HDI).

III. RESULTS AND DISCUSSION

3.1. Descriptive Statistics

An overview of the data regarding the Human Development Index and the factors influencing it in West Java Province is presented. The descriptive statistics are shown as follows:

Table 1. Descriptive Statistics

Variable	Min	Median	Mean	Max
HDI	68.39	75.32	74.68	83.75
Unemployment Rate (%)	2.48	3.74	3.91	11.47
APS 16–18 (%)	5.59	67.13	62.52	82.02
Population Density	0.69	7.416	7.035	85.10
Constant Price GRDP	36.54	146.53	146.85	370.16

Based on the descriptive statistics, it can be observed that there is variation among regions in terms of human development, education, health, and economic conditions. These variations indicate the need for a model capable of accommodating spatial heterogeneity within regional data. One such model is Geographically Weighted Regression (GWR), which can detect spatial variation in the relationships between variables. The following section presents the classical regression analysis used as the initial step prior to spatial regression modeling.

3.2. Classical Linear Regression Modeling

Classical regression modeling was conducted to examine the relationship between the Human Development Index (HDI) as the dependent variable and several independent variables, namely the unemployment rate (TPT), the gross enrollment rate for ages 16–18 (APS 16–18), population density, and constant-price GRDP. The estimation results of the classical regression model are as follows:

$$\hat{y} = 74,676 + 0,071x_1 + 0,682x_2 + 3,196x_3 + 0,834x_4$$

Overall, the model yields a R-squared value of 0.864 and an adjusted R-squared value of 0.839, indicating that it explains the variation in HDI reasonably well. However, the regression results show that only population density is statistically significant at the 0.05 level, while the other variables are not, suggesting that the model may have limitations and further spatial analysis (e.g., GWR) is warranted.

3.3. Classical Assumption Testing

A. Normality Test

To assess whether the residuals follow a normal distribution, the Shapiro–Wilk test was conducted. Based on the test results, the obtained p-value is 0.6285. Since the p-value is greater than 0.05, the residuals are considered to follow a normal distribution. Therefore, the normality assumption of the classical linear regression model is fulfilled.

B. Multicollinearity Test

Multicollinearity was tested using the Variance Inflation Factor (VIF). All independent variables have a VIF value below 10, as shown in the following table:

Table 2. VIF Values for Each Variable

Variable	VIF
Unemployment Rate	1.45
APS 16–18	1.71
Population Density	2.29

Constant Price GRDP 1.51

Multicollinearity was assessed using the Variance Inflation Factor (VIF). All independent variables have VIF values below 10, indicating that none of the predictors exhibit problematic collinearity. This means each variable provides distinct information to the model, and the regression estimates can be interpreted reliably without concerns of multicollinearity.

C. Heteroscedasticity Test

The heteroscedasticity test was conducted to examine whether a regression model contains variance inequality between one observation and another. The test was carried out using the Glesjer method, where each independent variable was regressed against the absolute residual values. If the significance value is greater than $\alpha = 0.05$, the Glesjer test concludes that there is no heteroscedasticity. The results are shown in Table 3.

Table 3. Heteroscedasticity Test Results

Variable	Significance
Unemployment Rate	0.896
APS 16–18	0.126
Population Density	0.191
Constant Price GRDP	0.579

The heteroscedasticity test was performed using the Glesjer method by regressing the absolute residuals on each independent variable. If the significance value exceeds 0.05, the model is considered free from heteroscedasticity. As shown in Table 3, all variables have significance values above 0.05, indicating that the regression model does not exhibit heteroscedasticity.

3.4. Spatial Heterogeneity Test

The spatial heterogeneity test aims to examine whether the relationship between the dependent variable and the independent variables varies across spatial units. Spatial heterogeneity indicates that the variance of the regression residuals is not constant across regions, suggesting that the classical linear regression model may not adequately capture spatial variations in the data. To assess spatial heterogeneity, the Breusch–Pagan test was employed. This test evaluates whether the residual variance is homogeneous (constant) or varies across observations. The hypotheses for the Breusch–Pagan test are as follows:

- H_0 : There is no spatial heterogeneity (homoskedastic residual variance).
- H_1 : There is spatial heterogeneity (heteroskedastic residual variance)..

Table 4. Spatial Autocorrelation Test Results

Breusch-Pagan Test	P-value
5.5428	0.236

The Breusch–Pagan test shows a p-value of 0.236, which is above the 0.05 threshold. This means the null hypothesis cannot be rejected, indicating that the model does not exhibit significant spatial heterogeneity. Residual variance is relatively uniform across regions.

3.5. Geographically Weighted Regression (GWR) Modeling

The Geographically Weighted Regression (GWR) model was used to identify spatial effects of the independent variables on the Human Development Index (HDI) across 27 districts/cities in West Java Province. The GWR model was tested using various combinations of kernel functions (Gaussian, Bisquare, and Tricube) and bandwidth types (fixed and adaptive). The best GWR model was selected based on the AIC, t-statistic, and R^2 values.

A. Comparison of GWR Models

The comparison of all GWR kernel and bandwidth combinations (Gaussian, Bisquare, Tricube; fixed and adaptive) is shown in Table 5:

Table 5. Comparison Results of the GWR Models

Bandwidth	Kernel	AIC
-----------	--------	-----

Adaptive	Gaussian	140.6470
	Bisquare	130.4790
	Tricube	127.3190
Fixed	Gaussian	2985.5696
	Bisquare	624.2662
	Tricube	547.0813

The AIC values show clear differences in model performance across each setting. Among the adaptive bandwidth options, the Tricube kernel yields the lowest AIC (127.3190), followed by Bisquare and Gaussian. Meanwhile, all fixed-bandwidth kernels produce substantially higher AIC values, indicating poorer model fit.

Overall, the adaptive Tricube kernel emerges as the best-performing GWR specification because it achieves the lowest AIC value (127.3190) across all combinations. This result suggests that the model provides the most efficient fit and stronger predictive accuracy compared to other kernel bandwidth configurations.

B. Local Parameter Significance (GWR)

The local parameter significance analysis from the best-performing GWR model shows that each variable affects the HDI differently across regions. In other words, the influence of the predictors is not uniform; some variables become significant only in certain districts or cities. Findings show:

Table 6. Significant Variables in the GWR Model

Variable	Number of Significant Districts/Cities
Unemployment Rate	0
APS 16–18	4
Population Density	12
Constant Price GRDP	7

The results show variation in how each variable significantly influences HDI across districts/cities. Population density has the widest spatial impact, being significant in 12 regions, indicating strong spatial variation in its relationship with HDI. Constant Price GRDP follows with significance in 7 regions, suggesting its economic effect on HDI is present but not universal. APS 16–18 is significant in only 4 regions, meaning its influence on human development is more localized. Meanwhile, the unemployment rate shows no significance in any district/city, indicating that its effect on HDI is not spatially detectable within the study area. The spatial distribution of the significant coefficients is presented in Figure 1, illustrating how unemployment affects HDI across regions in West Java.

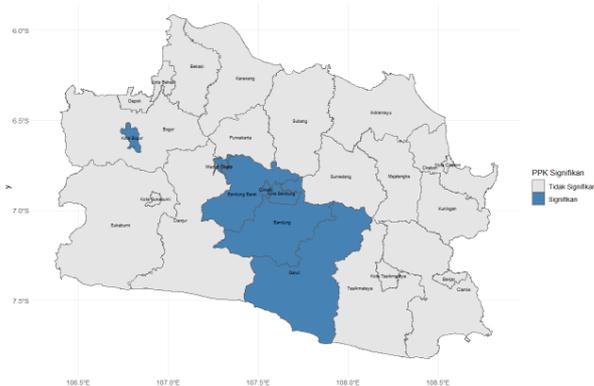


Figure 1. Spatial Distribution of Local GRDP Coefficients on HDI

Figure 1 shows that the influence of constant-price GRDP on the Human Development Index (HDI) is significant only in several central regions of West Java, as indicated by the blue-colored areas. This means that in these specific districts, higher GRDP is strongly associated with improvements in human development. In contrast, the grey areas

represent regions where the GRDP coefficient is not statistically significant, suggesting that GRDP is not a key determinant of HDI in those locations. Overall, the map highlights that the effect of GRDP on human development varies spatially across the province.

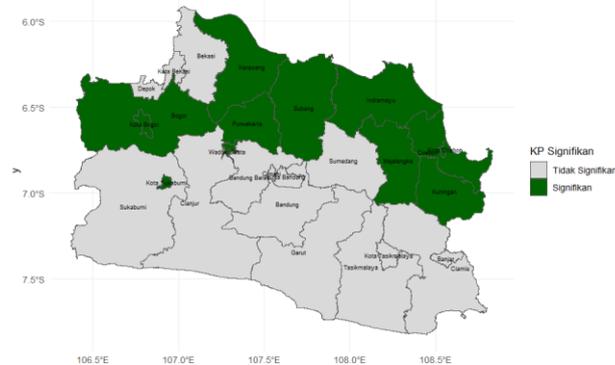


Figure 2. Spatial Distribution of Local Population Density Coefficients on HDI

The map shows that population density significantly affects HDI in several regions highlighted in green, mainly in the northern and western parts of West Java. These areas exhibit a strong spatial relationship between population density and human development. Meanwhile, the grey areas indicate regions where population density does not have a significant effect on HDI, demonstrating that its influence varies across locations.



Figure 3. Spatial Distribution of APS 16–18 Coefficients on HDI

The map shows that the APS 16–18 variable significantly affects HDI only in specific regions, highlighted in purple. This indicates that school participation among individuals aged 16–18 contributes meaningfully to human development in those areas. In contrast, the grey regions represent areas where APS 16–18 does not have a significant influence on HDI, suggesting that its effect varies across different parts of West Java.

3.6. Local Multicollinearity Diagnostics

Local multicollinearity diagnostics can be visualized using the VIF values, which reflect whether there are linear relationships among independent variables at the local level. The results of the local VIF test for all districts/cities in West Java Province are shown in Table 7.

Table 7. Summary of Local VIF Values for All Districts/Cities

Variable	Number of Districts/Cities with VIF ≥ 10
Unemployment Rate	0
APS 16–18	0
Population Density	0
Constant Price GRDP	0

The local VIF test results indicate that all independent variables Unemployment Rate, APS 16–18, Population Density, and Constant Price GRDP have VIF values below 10 in all districts/cities. This means that local multicollinearity does not occur within the GWR model. Therefore, each variable can be reliably used for local parameter estimation.

3.7. GWR Model Fit Test

The model adequacy test indicates that the GWR model provides a significant improvement over the OLS model ($F = 6.46$; $p = 0.0485$). The GWR residuals (10.26) are substantially smaller than the OLS residuals (66.29), demonstrating that GWR can capture local spatial variations that cannot be explained by the OLS model. Therefore, the relationships between the predictor variables and the HDI are spatially non-stationary, making the application of GWR more appropriate.

3.8. Model Evaluation

The evaluation of the regression models shows that the Geographically Weighted Regression (GWR) model outperforms the Ordinary Least Squares (OLS) model in capturing spatial variations in the dependent variable. As shown in Table 9, the GWR model has a higher R^2 value (1.000) and a higher AIC (140.647) compared to the OLS model ($R^2 = 0.864$; $AIC = 112.874$). Although the AIC is higher for GWR due to the increased model complexity, the perfect R^2 indicates that GWR captures the local variation in the data very well.

Table 8. Model Evaluation

Model	R^2	AIC
OLS	0.864	112.874
GWR	1.000	140.647

The GWR model incorporates local variations across all regions in West Java. The coefficients of the independent variables vary between regions, indicating that the relationship between the dependent and independent variables is not uniform. This highlights the importance of considering spatial heterogeneity when analyzing the factors affecting the dependent variable.

3.9. Model Result

A. Global Model (OLS)

The OLS model assumes that the relationship between HDI and its predictors, unemployment rate (TPT), gross enrollment rate for ages 16 to 18 (APS 16 to 18), population density, and constant price GRDP, is uniform across all regions. The estimated OLS model is:

$$\widehat{HDI} = 74,676 + 0,071 \cdot TPT + 0,682 \cdot APS_{16-18} + 3,196 \cdot PopulationDensity + 0,834 \cdot ConstantPriceGRDP$$

The model explains 86.4% of the variance in HDI ($R^2 = 0.864$) with an AIC of 112.874. Among the predictors, only population density is statistically significant at the 0.05 level, indicating that higher population density is associated with higher HDI.

B. Local Model (GWR)

The GWR model allows the relationship between HDI and its predictors, unemployment rate (TPT), gross enrollment rate for ages 16 to 18 (APS 16 to 18), population density, and constant price GRDP, to vary across regions. The general GWR model is:

$$\widehat{HDI} = \beta_0(u_i, v_i) + \beta_1(u_i, v_i) \cdot TPT + \beta_2(u_i, v_i) \cdot APS_{16-18} + \beta_3(u_i, v_i) \cdot PopulationDensity + \beta_4(u_i, v_i) \cdot ConstantPriceGRDP$$

For Bekasi, the local coefficients are:

$$\widehat{HDI}_{bekasi} = 74,933 + 0,117 \cdot TPT + 0,399 \cdot APS_{16-18} + 3,534 \cdot PopulationDensity + 0,762 \cdot ConstantPriceGRDP$$

The local R^2 for Bekasi is 0.919, indicating that 91.9% of the variation in HDI is explained by the local model. Population density has the strongest positive effect, followed by APS 16 to 18 and GRDP, while TPT has a smaller effect locally. Bekasi was chosen as a representative region due to its high local R^2 and its urban characteristics, making it suitable for illustrating spatial variation in HDI determinants.

CONCLUSION

This study concludes that the Geographically Weighted Regression (GWR) model provides a better representation of the spatial variations influencing the Human Development Index (HDI) in West Java Province compared to the global Ordinary Least Squares (OLS) model. The adaptive Gaussian kernel was identified as the best-performing model, indicated by the lowest AIC and CV values. The GWR model also demonstrated superior explanatory power with a higher R^2 value, highlighting its ability to capture spatial heterogeneity that the OLS model cannot detect. The local parameter analysis revealed that the Unemployment Rate (TPT), Population Density, Gross Regional Domestic Product (GRDP), and School Participation Rate for ages 16–18 (APS 16–18) show varying levels of significance across districts and cities, confirming non-uniform relationships between predictors and HDI. Population Density and GRDP were significant in the majority of regions, while APS 16–18 showed significance only in selected areas, and TPT was significant in fewer locations. These findings emphasize that the determinants of HDI differ spatially, and thus, region-specific development strategies are essential. Overall, the GWR model proves to be an effective analytical approach for understanding geographically diverse factors affecting human development.

REFERENCES

- UNDP, 1990. *Global Human Development Report*. s.l.:s.n.
- Alvin C. Rencher and G. Bruce Schaalje, 2008. *Linear Models in Statistics*. Hoboken, New Jersey: John Wiley & Sons Inc.
- A. Rajabifard and I. Williamson, 2001. Spatial data infrastructures: Concept, SDI hierarchy and future directions. In: Tehran, Iran: Proceedings of Geomatics 80.
- M. F. Agustina, R. W. a. M. Y. D., 2015. Pemodelan Geographically Weighted Regression (GWR) pada Tingkat Kemiskinan di Provinsi Jawa Tengah. *Statistika*, pp. vol.3, no.2, pp.67-74.
- K. Amelia, L. O. A. a. L. F., 2020. "Pemodelan Incident Rate Demam Berdarah Dengue di Indonesia yang Berkaitan dengan Faktor Lingkungan Menggunakan Metode Geographically Weighted Regression (Gwr). *Ekologia*, pp. vol. 20, no. 2, pp. 64–73, doi: 10.33751/ekologia.v20i2.2167..
- N. Lutfiani, S. a. S. M., 2019. Pemodelan Geographically Weighted Regression (GWR) dengan Fungsi Pembobot Kernel Gaussian dan Bi-Square. *Unnes Journal of Mathematics*, pp. vol. 8, no. 1, pp. 82–91, doi: 10.15294/ujm.v8i1.17103..
- Ispriyanti, I. M. D. & D. M. Y., 2005. Pemodelan Data Kemiskinan di Provinsi Sumatera Barat dengan Metode Geographically Weighted Regression (GWR). *Media Statistika*, pp. 8(1), 1-12.
- Haq, M., 1996. *Reflections on Human Development, 1st Edition*.. New York: Oxford University Press.
- BPS, B. P. S., 2023. *Statistik Kemiskinan Indonesia 2023*, Jakarta: Badan Pusat Statistik.
- Walpole, R., 1993. *Pengantar Statistika. Edisi ketiga. Alih bahasa Ir. Bambang Sumantri*. Jakarta: PT Gramedia.
- Drapper, N.R., dan Smith H. , 1992. *Analisis Regresi Terapan. Edisi Kedua*. New York : s.n.
- Radjabidfar, and Abbas., 2001. *SDI Hierarchy, from Local to Global SDI Initiatives*.. Melbourne: Victoria: Spatial Data Research Group, Departement of Geomatics..

- Harini, S., Purhadi, Mashuri, M., & Sunaryo, S. , 2012. Statistical test for multivariate geographically weighted regression model using the method of maximum likelihood ratio test. *International Journal of Applied Mathematics & Statistics*, pp. 29(5), 1–15.
- Saputri, S., Ispriyanti, D., & Wuryandari, T., 2015. Penentuan Model Kemiskinan di Jawa Tengah dengan Multivariate Geographically Weighted Regression (MGWR).. *JURNAL GAUSSIAN*, pp. (2), 161-170.