CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (2) (2026), Pages 1716–1733
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Spatial Heterogeneity of Poverty Determinants in Indonesia: A Hierarchical Geographically Weighted Regression Approach

Debora Dwi Kurniawati, Henny Pramoedyo*, and Suci Astutik

Department of Statictics, Faculty of Science, Technology, and Mathematics (FSTeM), Brawijaya University, Indonesia

Abstract

This study employs the Hierarchical Geographically Weighted Regression (HGWR) model to analyze poverty determinants in Indonesia, addressing spatial heterogeneity and hierarchical data structures simultaneously. Using data from 34 provinces and 507 regencies/cities, HGWR with a Gaussian kernel (bandwidth = 15) substantially outperforms global OLS regression, raising R^2 from 0.502 to 0.754. Moran's I on the global model's residuals (0.2216, $p < 0.001$) confirms spatial nonstationarity, while HGWR residuals show this autocorrelation is eliminated (-0.000641, $p = 0.416$).At the regency/city level, the poverty line and mean years of schooling are significantly associated with lower poverty rates, while adjusted per capita expenditure shows no significant fixed-effect relationship. At the provincial level, the Human Development Index (HDI) consistently reduces poverty (mean coefficient -0.8082) but varies spatially, with the strongest impact in eastern Indonesia (below -1.043) and weakest in Java (above -1.043). Expected years of schooling shows a positive mean coefficient (3.4755), concentrated mainly in Java.These findings indicate that place-based, rather than uniform, poverty reduction strategies may be more effective, as HDI gains in eastern Indonesia correspond to larger marginal poverty reductions.

Keywords: Poverty; Hierarchical Geographically Weighted Regression (HGWR); spatial heterogeneity; multilevel modeling; Indonesia.

1. Introduction

Poverty remains a fundamental challenge in Indonesia's development, directly affecting welfare inequality and social policy effectiveness. The Indonesian Central Bureau of Statistics (BPS) recorded a poverty rate of 8.57 percent in September 2024, with a poverty line of IDR 595,242 per capita per month [1]. The spatial availability of poverty data at the regency/city level makes regional analysis feasible [2], yet the uneven distribution of poverty across regions indicates that development has not proceeded uniformly, remaining tied to persistent territorial inequality [3].

Global regression methods most notably OLS are commonly applied to analyze poverty, but assume spatially constant relationships, often oversimplifying Indonesia's diverse spatial reality. Studies have shown that inter-regency poverty dynamics are closely linked to spatial dependence and convergence processes that vary across regions [3]. Poverty in Indonesia encompasses both chronic and transient dimensions driven by differing socioeconomic conditions [4], with substantial

*Corresponding author. E-mail: hennyp@ub.ac.id

Received: May 26, 2026 Revised: June 28, 2026 Accepted: July 13, 2026
DOI: https://doi.org/10.18860/cauchy.v11i2.42956

disparities identified even within Java alone [5]. These findings suggest that the effects of factors such as per capita expenditure, educational attainment, and the poverty line on the poverty headcount rate vary across regions.

A further methodological challenge arises from the hierarchical structure of the data: regencies and cities are nested within provinces. As [6] emphasizes, lower-level units are influenced by upper-level contexts and cannot be treated as independent. Standard multilevel models accommodate nested structures but assume spatially constant parameters [7], while Geographically Weighted Regression (GWR) [8] captures local spatial heterogeneity—as applied, for example, to human development indicators [9]but does not explicitly account for hierarchical nesting [10]. Consequently, both approaches have distinct limitations when applied separately to Indonesian poverty data, which exhibits spatial and hierarchical characteristics simultaneously [11].

Hierarchical Geographically Weighted Regression (HGWR) [12] addresses this gap by integrating hierarchical structure and spatial heterogeneity into a unified analytical framework. HGWR simultaneously estimates global fixed effects, local fixed effects, and random effects, enabling certain variable effects to vary spatially at the group level—making it more appropriate for analyzing regencies nested within provinces that differ in their development characteristics. Prior spatial studies on Indonesian poverty have employed non-hierarchical approaches [5] or conventional spatial panel models [3, 13], including evidence of spatial spillover effects from government expenditure on poverty alleviation [13]. However, research that jointly accounts for both hierarchical nesting and spatial heterogeneity in the analysis of Indonesian poverty remains limited.

This study models the poverty headcount rate (P_0) as the dependent variable, with adjusted per capita expenditure, mean years of schooling, and the poverty line as regency/city-level variables, and Human Development Index (HDI) and expected years of schooling as province-level variables, using 2024 BPS data for 34 provinces and 507 regencies/cities. The research objectives are to: 1) identify the effects of egency/city-level economic and educational variables on the poverty headcount rate; 2) examine the role of provincial-level HDI and expected years of schooling; and 3)investigate the spatial variation of provincial-level variable effects across regions of Indonesia.

The novelty of this study is threefold. First, unlike standard GWR studies [9] that ignore hierarchical nesting, or multilevel models [6] that assume spatial stationarity, this study integrates both structures simultaneously through HGWR. Second, unlike previous spatial econometric studies on Indonesian poverty [3, 5, 13] that operate at a single administrative level, this study explicitly models cross-level interactions between regency/city and provincial variables. Third, this is the first study to apply HGWR specifically to the poverty headcount rate (P_0) in Indonesia using the most recent 2024 BPS cross-sectional data covering all 34 provinces, providing more current empirical evidence than prior panel-based approaches.

2. Methods

This study uses 2024 secondary data obtained from Badan Pusat Statistik (BPS). The data consist of two levels: regency/city and province. The regency/city level is the main unit of analysis, while the province level is used to represent the hierarchical structure in the HGWR model. The dataset includes 507 regencies/cities nested within 34 provinces in Indonesia. The response variable is the poverty rate at the regency/city level. The explanatory variables include per capita expenditure, mean years of schooling, and poverty line at the regency/city level, as well as Human Development Index (HDI) and expected years of schooling at the province level.

Table 1: Description of Variables Used in the Study

Variable	Description	Level	Source
Y	Poverty Rate	Regency/City	BPS
$X1$	Per capita expenditure	Regency/City	BPS
$X2$	Mean years of schooling	Regency/City	BPS
$X3$	Poverty line	Regency/City	BPS
$Z0$	Human Development Index	Province	BPS
$Z1$	Expected years of schooling	Province	BPS

The analytical procedure of this study using the Hierarchical Geographically Weighted Regression (HGWR) approach consists of the following steps:

1. **Data preparation:** Secondary data for 2024 were obtained from Badan Pusat Statistik (BPS), consisting of poverty rates and explanatory variables for 508 regencies/cities nested within 34 provinces in Indonesia. However, only 507 regencies/cities could be matched consistently with the administrative boundary shapefile used for spatial analysis. Therefore, the final analytic sample consisted of 507 regencies/cities.Table 1 presents the description of all variables used in this study.
2. **Descriptive analysis:** Descriptive statistics ere calculated to summarize the general characteristics of poverty rates and all explanatory variables across regencies/cities and provinces.
3. **Multicollinearity test:** The Variance Inflation Factor (VIF) was employed to detect multicollinearity among explanatory variables, as specified in Eq. (1).
4. **Spatial autocorrelation test:** Moran's I test was applied to examine the spatial dependence of poverty rates across regencies/cities, as specified in Eq. (2) and Eq. (3).
5. **Spatial heterogeneity test:** The Breusch–Pagan test was used to identify spatial heterogeneity in the regression relationship, as specified in Eq. (4).
6. **Hierarchical structure identification:** The Intraclass Correlation Coefficient (ICC) was calculated to quantify the proportion of variance in poverty rates attributable to between-province differences, as specified in Eq. (5). A multilevel model was also estimated to confirm the relevance of the hierarchical structure prior to HGWR estimation.
7. **HGWR model estimation:** The HGWR model was estimated by incorporating regency/city-level variables (X_1, X_2, X_3), province-level variables (Z_0, Z_1), and the hierarchical structure of regencies/cities nested within provinces, as shown in Eq. (6). Parameter estimation was performed using a backfitting maximum likelihood algorithm implemented via the `hgwr` package in R, with a Gaussian kernel and adaptive bandwidth of $h^* = 15$ nearest-neighbor provinces selected by minimizing AIC.
8. **Spatial mapping of local coefficients:** The estimated local coefficients for provincial-level variables (Z_0 and Z_1) were visualized using choropleth maps created in ArcGIS Pro 3.0. The WGS 1984 UTM Zone 50S projection was used for cartographic display purposes only and does not affect the HGWR distance calculations,which are based on unprojected geographic coordinates (WGS 1984 decimal degrees).
9. **Model evaluation and interpretation:** Model performance was evaluated by comparing four models OLS, Multilevel Model, standard GWR, and HGWR using AIC, R^2, and residual Moran's I. The local spatial and hierarchical effects on poverty rates were then interpreted based on both statistical results and spatial maps.All policy implications are stated as associations suggested by the model rather than causal effects, given the cross-sectional observational design.

2.1. Spatial Analysis

2.1.1. Multicollinearity Diagnostic

Multicollinearity among explanatory variables was examined using the Variance Inflation Factor (VIF), which detects strong linear relationships between independent variables. For each variable X_k, VIF is calculated as:

$$VIF_k = \frac{1}{1 - R_k^2}, \tag{1}$$

where R_k^2 is the coefficient of determination from regressing X_k on the other variables. Although VIF thresholds vary across studies, variables with high values are considered to have potential multicollinearity and require further evaluation [14].

2.1.2. Spatial Autocorrelation

Spatial autocorrelation was assessed using Moran's I to examine whether model residuals exhibit spatial dependence [15]. Moran's I is expressed as:

$$I = \frac{n}{W} \frac{\sum_i \sum_j w_{ij}(e_i - \bar{e})(e_j - \bar{e})}{\sum_i (e_i - \bar{e})^2}, \tag{2}$$

where n is the number of spatial units, w_{ij} is the spatial weight, e_i is the residual, $\bar{e}$ is the mean residual, and $W = \sum_i \sum_j w_{ij}$. An insignificant result indicates no residual spatial autocorrelation, confirming that the spatial weighting structure adequately accounts for spatial dependence.

The significance of Moran's I is evaluated using:

$$Z_I = \frac{I - \mathbb{E}[I]}{\sqrt{\mathrm{Var}(I)}}. \tag{3}$$

An insignificant result indicates no residual spatial autocorrelation, suggesting that the spatial weighting structure adequately accounts for spatial dependence.

2.1.3. Heteroscedasticity Test

The Breusch Pagan (BP) test was used to examine whether residual variance is constant [16], based on the relationship between squared residuals and explanatory variables:

$$\mathrm{BP} = \sum_{i=1}^{n} \hat{u}_i^2 \mathbf{z}_i^\top \left(\mathbf{Z}^\top \mathbf{Z}\right)^{-1} \mathbf{z}_i. \tag{4}$$

Under the null hypothesis of homoskedasticity, the BP statistic follows a chi-square distribution. A p-value greater than 0.05 indicates no evidence of heteroscedasticity.

2.2. Hierarchical Structute Identification

The hierarchical structure in this study was identified using the Intraclass Correlation Coefficient (ICC). According to [17], ICC is used to measure the proportion of variance in the response variable that is attributable to clustering or group-level differences. In this study, regencies/cities are treated as level-1 units, while provinces are treated as level-2 units. The ICC is calculated as follows:

$$ICC = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\epsilon^2} \tag{5}$$

where σ_μ^2 represents the between-province variance and σ_ϵ^2 represents the within-province variance. A higher ICC value indicates that variations in poverty rates are substantially influenced by differences between provinces, thereby supporting the use of a hierarchical model.

2.3. Hierarchical Geographically Weighted Regression (HGWR)

Hierarchical Geographically Weighted Regression (HGWR) is applied in this study to model the poverty headcount rate by simultaneously accounting for spatial heterogeneity across provinces and hierarchical data structure wherein regencies/cities are nested within provinces. The HGWR model integrates three components: local fixed effects whose coefficients vary spatially across provincial locations, global fixed effects with coefficients assumed constant across all regions, and province-level random effects [11, 18]. The general form of the HGWR model is:

$$\mathbf{y} = \mathbf{G}\boldsymbol{\gamma} + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\mu} + \boldsymbol{\epsilon} \tag{6}$$

where $\mathbf{y}$ is the response variable vector, $\mathbf{G}\boldsymbol{\gamma}$ denotes the local fixed effects with spatially varying coefficients, $\mathbf{X}\boldsymbol{\beta}$ represents the global fixed effects with location-invariant coefficients, $\mathbf{Z}\boldsymbol{\mu}$ captures the province-level random effects, and $\boldsymbol{\epsilon}$ is the idiosyncratic error term. Applied to this study, the HGWR model specification is:

$$Y_{ij} = \gamma_0(u_i, v_i) + \gamma_1(u_i, v_i)Z_{0j} + \gamma_2(u_i, v_i)Z_{1j} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \mu_j + \varepsilon_{ij} \tag{7}$$

where Y_{ij} is the poverty headcount rate of regency/city i in province j; X_{1ij}, X_{2ij}, X_{3ij} are regency/city-level variables (per capita expenditure, mean years of schooling, and poverty line, respectively); Z_{0j} and Z_{1j} are province-level variables (HDI and expected years of schooling); $\gamma_k(u_i, v_i)$ are local fixed coefficients that vary according to the geographic coordinates (u_i, v_i) of the provincial centroid; β_l are global fixed coefficients; μ_j is the province-level random intercept; and ε_{ij} is the residual error.

2.3.1. Euclidean Distance and Coordinate System

In the HGWR model, spatial proximity between provinces is quantified using Euclidean distance computed from WGS 1984 geographic coordinates (unprojected decimal degrees). The Euclidean distance between two provincial centroids i and j is:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \tag{8}$$

The use of Euclidean distance is justified on three grounds. First, spatial weighting in HGWR operates at the provincial level ($n = 34$) using an adaptive nearest-neighbor bandwidth, so only the rank ordering of distances matters rather than absolute values making results insensitive to the specific distance metric. Second, [19] found no significant difference between Euclidean and Haversine distances in regional-scale analyses within Indonesia. Third, Euclidean distance is the default metric implemented in the `hgwr` R package [18]. As a sensitivity check, the rank ordering of nearest-neighbor provinces under Euclidean distance was verified to be consistent with Haversine distance for all 34 centroids, and no change in nearest-neighbor assignments was found under bandwidth $h^* = 15$. We nonetheless acknowledge that Indonesia's approximately 47° longitude span means Euclidean distance does not account for the Earth's curvature; future work may explore great-circle or network-based alternatives as the `hgwr` package develops. The WGS 1984 UTM Zone 50S projection is used solely for cartographic display and does not affect HGWR distance calculations.

2.3.2. Gaussian Kernel and Spatial Weights

Spatial weights w_{ij} determine how much each province j contributes to the estimation of local coefficients at target province i. Provinces closer to i receive higher weights, while more distant provinces receive lower weights. The **Gaussian kernel** function is:

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{h^2}\right) \tag{9}$$

where w_{ij} is the spatial weight assigned to province j when estimating local coefficients at target location i, d_{ij} is the Euclidean distance from Eq. (8), and h is the bandwidth parameter. The Gaussian kernel assigns *non-zero* weights to all provinces, with weights decaying smoothly as distance increases. This continuous decay makes it appropriate for capturing gradual spatial transitions in coefficient values [8].

The spatial weight matrix for target province i is:

$$\mathbf{W}_i = \operatorname{diag}(w_{i1}, w_{i2}, \ldots, w_{i,34}) \tag{10}$$

This matrix is constructed separately for each target province and enters the local fixed-effect estimation in Section 2.3.4.

2.3.3. Bandwidth Selection Using AIC

Bandwidth selection determines the degree of spatial smoothing when estimating local parameters. In this study, an adaptive nearest-neighbor bandwidth was used rather than a fixed-distance bandwidth, following the geographically weighted regression framework [8, 20]. Therefore, the selected value of 15 does not represent a fixed distance scale. Instead, it refers to the number of nearest neighboring provinces used to define the local bandwidth for each target province.

Let q denote the nearest-neighbor bandwidth and let $d_{i,(q)}$ be the Euclidean distance from target province i to its q-th nearest provincial centroid. For a given value of q, the adaptive bandwidth distance for province i is defined as:

$$h_i(q) = d_{i,(q)} \tag{11}$$

Thus, when the selected bandwidth is $q^* = 15$, the bandwidth distance for each target province is the Euclidean distance from that province to its 15th nearest provincial centroid. This target-specific bandwidth distance is then used in the Gaussian kernel to calculate the spatial weight between target province i and province j [8]:

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{h_i(q)^2}\right) \tag{12}$$

where d_{ij} is the Euclidean distance between the centroids of provinces i and j, and $h_i(q)$ is the adaptive bandwidth distance for target province i. Under the Gaussian kernel, all provinces receive non-zero weights; however, provinces located farther from the target province receive progressively smaller weights.

The optimal nearest-neighbor bandwidth was selected by minimizing the Akaike Information Criterion (AIC) over a candidate range $q \in \{10, 11, \ldots, 25\}$ [20]:

$$\text{AIC} = -2\ln(\hat{L}) + 2k_{\text{eff}} \tag{13}$$

where $\hat{L}$ is the maximized log-likelihood and k_{eff} is the effective number of parameters estimated by the model, commonly represented through effective degrees of freedom in geographically weighted models [8, 20]. The optimal bandwidth is expressed as:

$$q^* = \arg\min_q \ \text{AIC}(q) \tag{14}$$

Both Gaussian and Bisquared kernels were evaluated across the candidate bandwidth range. The Gaussian kernel with an adaptive nearest-neighbor bandwidth of $q^* = 15$ produced the lowest AIC and was therefore used in the final HGWR model [12, 18].

2.3.4. Parameter Estimation of HGWR

1. Local Fixed Effect Parameter Estimation
   Parameter estimation uses the weighted least squares (WLS) approach. For the local fixed effect component, parameters at location (u_i, v_i) are estimated by minimizing:
$$Q(\boldsymbol{\gamma}) = (\mathbf{y} - \mathbf{G}\boldsymbol{\gamma})^T \mathbf{W}_i (\mathbf{y} - \mathbf{G}\boldsymbol{\gamma}) \tag{15}$$
   where $\mathbf{W}_i$ is the spatial weight matrix for location i. Taking the first partial derivative of $Q(\boldsymbol{\gamma})$ with respect to $\boldsymbol{\gamma}$ and setting it to zero yields the local parameter estimator: Thus, the local parameter estimator for location (u_i, v_i) is given by:
$$\hat{\boldsymbol{\gamma}}(u_i, v_i) = (\mathbf{G}^T \mathbf{W}_i \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}_i \mathbf{y} \tag{16}$$
2. Global FIxed Effect Parameter Estimation
   For the global fixed effect parameter $\boldsymbol{\beta}$, the local and random effect components are first separated from the response variable, defining a transformed response vector $\mathbf{y}^* = \mathbf{y} - \mathbf{G}\boldsymbol{\gamma} - \mathbf{Z}\boldsymbol{\mu}$, which reduces the model to standard linear form $\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. The OLS estimator for the global fixed effect parameter is then: Thus, the Ordinary Least Squares (OLS) estimator for the global fixed effect parameter is given by:
$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}^*$$
   By substituting back the definition of $\mathbf{y}^* = \mathbf{y} - \mathbf{G}\boldsymbol{\gamma} - \mathbf{Z}\boldsymbol{\mu}$, the estimator can be explicitly written as:
$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{y} - \mathbf{G}\boldsymbol{\gamma} - \mathbf{Z}\boldsymbol{\mu}) \tag{17}$$
   This estimator produces global coefficients that do not vary by location, representing the net effect of explanatory variables assumed to have a homogeneous influence on poverty rates across all regencies/cities under study.

2.3.5. Backfitting Algorithm for Joint Estimation

The three components local fixed effects $\boldsymbol{\gamma}$, global fixed effects $\boldsymbol{\beta}$, and random effects $\boldsymbol{\mu}$ are estimated jointly using a backfitting maximum likelihood algorithm [11, 18]. The algorithm proceeds as follows:

Step 1 (Initialization): Set initial values $\hat{\boldsymbol{\beta}}^{(0)} = \mathbf{0}$ and $\hat{\boldsymbol{\mu}}^{(0)} = \mathbf{0}$.

Step 2 (Local fixed effects): Given current estimates of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\mu}}$, compute adjusted response $\mathbf{y}^* = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{Z}\hat{\boldsymbol{\mu}}$, then estimate $\hat{\boldsymbol{\gamma}}(u_i, v_i)$ at each provincial location using Eq. (16).

Step 3 (Global fixed effects): Given current estimates of $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\mu}}$, compute $\mathbf{y}^* = \mathbf{y} - \mathbf{G}\hat{\boldsymbol{\gamma}} - \mathbf{Z}\hat{\boldsymbol{\mu}}$, then re-estimate $\hat{\boldsymbol{\beta}}$ using Eq. (17).

Step 4 (Random effects): Given current estimates of $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\beta}}$, estimate province-level random intercepts $\hat{\boldsymbol{\mu}}$ via REML.

Step 5 (Convergence): Repeat Steps 2–4 until the maximum change in all parameter estimates falls below a convergence tolerance of 10^{-6}.

This iterative procedure ensures all three components are mutually adjusted until final estimates reflect their joint contribution to explaining the poverty headcount rate, implemented using the `hgwr` package in R [11].

Local significance testing. The significance of local coefficients $\hat{\gamma}_k(u_i, v_i)$ at each provincial location was assessed using t-tests based on standard errors derived from the backfitting estimator, with significance thresholds at $p < 0.05$, $p < 0.01$, and $p < 0.001$.

3. Results and Discussion

3.1. Descriptive Statistics

Descriptive analysis was conducted to provide an overview of the characteristics of the poverty rate and the explanatory variables across 507 regencies/cities in Indonesia using 2024 data from the Badan Pusat Statistik (BPS).

Table 2: Descriptive Statistics

Variable	Min	Median	Mean	Max
Y	2.230	9.170	10.970	41.420
X_1	8.433	9.336	9.315	10.149
X_2	2.900	8.620	8.757	13.100
X_3	12.610	13.200	13.220	13.980
Z_0	53.420	73.880	72.990	83.080
Z_1	9.970	13.340	13.400	15.700

Table 2 presents the summary statistics. The poverty rate (Y) has a mean of 10.97 and a range of 2.23–41.42, indicating substantial regional variation. Adjusted per capita expenditure (X_1) and mean years of schooling (X_2) also vary considerably (ranges: 8.43–10.15 and 2.90–13.10, respectively), suggesting pronounced educational and economic disparities. The poverty line (X_3) is relatively homogeneous (range 12.61–13.98), while the Human Development Index (Z_0) ranges widely (53.42–83.08), reflecting substantial inter-provincial disparities in human development. Expected years of schooling (Z_1) ranges from 9.97 to 15.70. Overall, these results confirm considerable regional heterogeneity, particularly in poverty and education indicators, providing empirical motivation for a spatially heterogeneous analytical approach.

To illustrate the spatial distribution of poverty in Indonesia in 2024, a thematic map is presented in Figure 1.

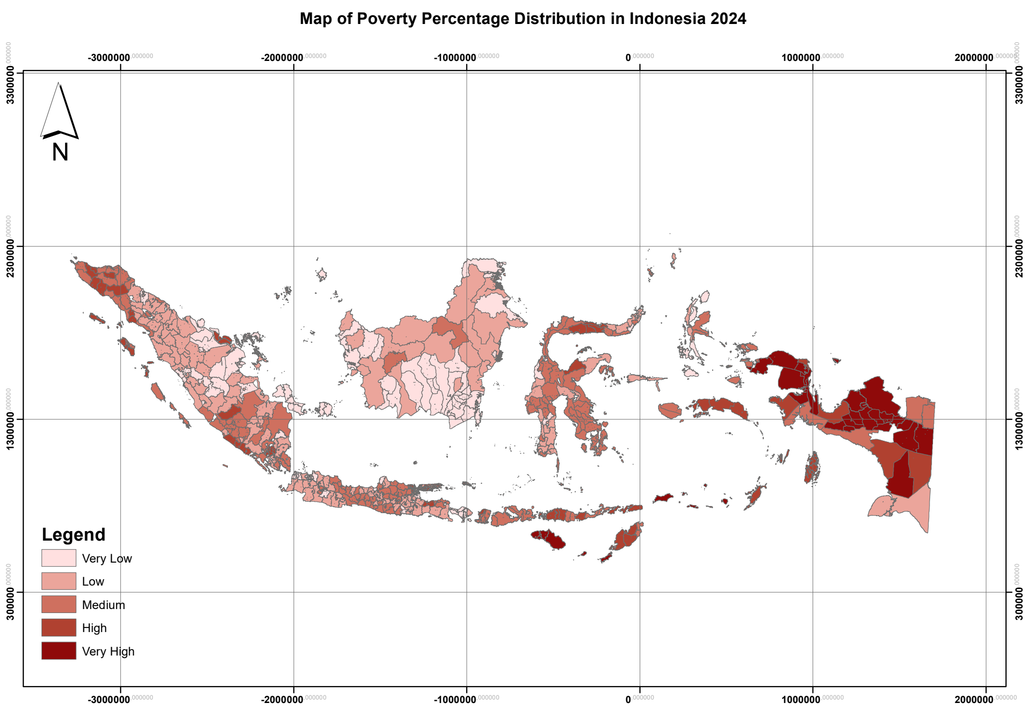

Fig. 1: Map of Poverty Percentage Distribution Across Indonesian Provinces in 2024

Based on Figure 1, poverty levels exhibit clear spatial variation across regions. Low- to moderate-poverty categories are generally concentrated in western Indonesia, particularly in Sumatra and Java. In contrast, high to very high poverty categories are concentrated in eastern Indonesia, especially in Papua as well as parts of Nusa Tenggara and Maluku.

In addition, spatial variation is also observed within the same regions, such as Sumatra and Sulawesi, where regencies/municipalities display a combination of low to high poverty categories. Overall, this pattern indicates the existence of spatial disparities in poverty levels across Indonesia and provides a basis for further analysis using spatial approaches.

3.2. Multicollinearity Test

The multicollinearity test was conducted on the regency/city-level independent variables using the Variance Inflation Factor (VIF). Meanwhile, the province-level variables were examined through inter-variable correlation to ensure that no excessively strong relationships existed.

Table 3: Multicollinearity Test

Variable	VIF
X_1	1.278900
X_2	1.634277
X_3	1.103801
Z_0	1.649448
Z_1	1.573386

All VIF values are below 10 (range: 1.10–1.65), indicating no severe multicollinearity among the explanatory variables. All variables are therefore retained for HGWR estimation.

3.3. Spatial Autocorrelation

Before applying the HGWR model, a spatial autocorrelation test was conducted on the residuals of the initial global model (Ordinary Least Squares). The Moran's I test was employed to detect the presence of spatial dependence. Table 4 presents the test results.

Table 4: Moran's I Test for Global Model Residuals

Statistic	Value	Expectation	p-value
Moran's I	0.2216	-0.0013	< 0.001

Moran's $I = 0.2216$ ($p < 0.001$) confirms significant positive spatial clustering in the OLS residuals, indicating that adjacent regions share similar unmodelledcharacteristics. This provides strong justification for employing HGWR, which is designed to handle spatial non-stationarity while accounting for hierarchical structures.

3.4. Hierarchical Linear Model

Multilevel models are utilized to accommodate the hierarchical structure of the data, wherein regencies/cities are nested within provinces. The estimation results of the multilevel model are presented in Table 5.

Table 5: Multilevel Model Estimation Results

Parameter	Estimate	Std. Error	t-value
Fixed Effects			
Constant	36.1506	10.4830	3.449
X_1	0.2448	0.6582	0.372
X_2	-1.7819	0.1149	-15.504

Bersambung ke halaman berikutnya...

Table 5 – Lanjutan dari halaman sebelumnya

Parameter	**Estimate**	**Std. Error**	***t*-value**
X_3	-0.9254	0.7253	-1.276
Random Effects (Variance | Std. Dev.)			
Intercept (Between-province)	24.46	4.945	–
Residual (Within-province)	10.09	3.177	–

Based on Table 5, the mean years of schooling (X_2) has a negative and significant effect on the poverty rate. This finding aligns with previous studies indicating that higher levels of education play a crucial role in reducing poverty. Conversely, per capita expenditure (X_1) and the poverty line (X_3) do not show substantial influences within the multilevel framework. Regarding the random components, the between-province intercept variance of 24.46 reveals considerable disparities in poverty rates across provinces. Meanwhile, the residual variance of 10.09 indicates that while within-province variation remains, it is notably smaller than the variation observed between provinces.

To quantify the magnitude of the hierarchical structure's effect, the Intraclass Correlation Coefficient (ICC) is calculated as follows:

$$\text{ICC} = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\epsilon^2} = \frac{24.46}{24.46 + 10.09} = 0.708 \tag{18}$$

The ICC value of 0.708 indicates that approximately 70.8% of the variance in poverty rates can be explained by between-province differences. This result suggests that the hierarchical structure plays a crucial role in explaining variations in poverty across regencies/cities in Indonesia. Consequently, the multilevel model provides a better fit than a standard global linear regression model. However, since this model does not yet account for spatial non-stationarity across regions, the analysis is extended using the Hierarchical Geographically Weighted Regression (HGWR) model.

3.5. Estimation of Hierarchical Geographically Weighted Regression (HGWR)

The first step in calibrating the HGWR model is determining the Euclidean distance for the local model. The bandwidth selection was performed using two kernel functions: Gaussian and Bisquared. Table 6 presents the bandwidth selection results based on AIC and R-squared values.

Table 6: Bandwidth Selection Results

Kernel	**Bandwidth**	**AIC**	R^2
Bisquared	18	1374.6	0.753
Gaussian	15	1372.6	0.754

Based on Table 6, the Gaussian kernel produced the lowest AIC with an adaptive nearest-neighbor bandwidth of $h^* = 15$. This value does not represent a fixed distance scale; rather, it means that for each target province, the bandwidth distance was defined by the Euclidean distance to its 15th nearest provincial centroid. Therefore, the same number of nearest provinces was used as the main information source for each local estimation, while farther provinces received smaller weights through the Gaussian kernel. The estimated HGWR model is expressed in Equation 19:

$$\hat{Y}_{ij} = 11.925327 + \gamma_1(u_i, v_i)Z_{0j} + \gamma_2(u_i, v_i)Z_{1j} + 0.060145X_{1ij} - 2.196809X_{2ij} - 0.307432X_{3ij} + \mu_j \tag{19}$$

For example, the model specification for Jambi province is presented in Equation 20:

$$\hat{Y}_{Jambi} = 11.925327 - 0.47101Z_{0j} + 1.55398Z_{1j} + 0.060145X_{1ij} - 2.196809X_{2ij} - 0.307432X_{3ij} + \mu_j \tag{20}$$

Having presented the general HGWR model and a specific example for Jambi province, the following subsections elaborate on the estimation results for each model component: **global fixed effects**, **local fixed effects**, and **random effects**.

1. **Fixed Effects**
   Table 7 presents the estimation results for the global fixed effects. Global fixed effects are variables whose coefficients are assumed constant across all regions (provinces) in Indonesia. In this model, three regency/city-level variables (X_1, X_2, and X_3) are specified as global fixed effects.

   Table 7: Fixed Effects Estimation Results

Variable	Estimate	Std. Error	t-value	Pr(>|t|)	Significance
Intercept	11.9253	0.1488	80.1499	<0.001	Significant
X_1	0.0601	0.0512	1.1754	0.240	Not Significant
X_2	−2.1968	0.0552	−39.8183	<0.001	Significant
X_3	−0.3074	0.0489	−6.2853	<0.001	Significant

   The global fixed-effect estimates at the regency/city level indicate the following:
   (a) **Mean years of schooling (X_2):** $\hat{\beta}_2 = -2.197$, $t = -39.818$, $p < 0.001$ — Statistically significant. A one-year increase in mean years of schooling is associated with a decrease of approximately 2.197 percentage points in the poverty headcount rate, holding other variables constant. This finding is consistent with human capital theory, wherein higher educational attainment enhances individual productivity and expands economic opportunities, thereby reducing poverty at the regional level.
   (b) **Poverty line (X_3):** $\hat{\beta}_3 = -0.307$, $t = -6.285$, $p < 0.001$ — Statistically significant. A one-unit increase in the poverty line is associated with a decrease of approximately 0.307 percentage points in the poverty headcount rate. This negative association may reflect the role of the poverty line as a proxy for regional cost-of-living adjustments: regions with higher poverty lines tend to be more economically developed, where households are more capable of meeting basic needs. This result should be interpreted cautiously, however, as the poverty line is both a threshold measure and a policy instrument whose relationship with measured poverty is partly mechanical.
   (c) **Adjusted per capita expenditure (X_1):** $\hat{\beta}_1 = 0.060$, $t = 1.175$, $p = 0.240$ Not statistically significant at the conventional level. This result indicates that, after controlling for mean years of schooling, the poverty line, and provincial-level effects, adjusted per capita expenditure does not show a statistically significant association with the poverty headcount rate in the HGWR fixed-effect component. A possible explanation is that the effect of per capita expenditure on poverty may operate primarily through spatial variation at the provincial level captured by the local effects of Z_0 and Z_1 rather than as a uniform global effect. This finding differs from the multilevel model results (Table 5), where X_1 also did not reach significance ($p = 0.372$), suggesting that this result is consistent across model specifications rather than an artefact of the HGWR estimation.

2. **Local Effects**
   Local effects in the HGWR model are represented by Z_0 (Human Development Index) and Z_1 (Expected Years of Schooling). These variables are modeled as local effects because their impacts on poverty rates are assumed to vary spatially across regions.

Table 8: Summary of Local Coefficient Estimates

Variable	Mean Est.	Mean SD	Significance (%)			
			p<0.001	p<0.01	p<0.05	p<0.1
Z_0	-0.8082	0.2007	58.8%	14.7%	20.6%	5.9%
Z_1	3.4755	0.9771	58.8%	35.3%	5.9%	0%

Based on Table 8, Z_0 (HDI) has a negative mean local coefficient of -0.8082, indicating that higher HDI is consistently associated with lower poverty rates across all provinces. The coefficient of Z_0 is statistically significant in most provinces, with 58.8% significant at $p < 0.001$, 14.7% at $p < 0.01$, and 20.6% at $p < 0.05$.
Meanwhile, Z_1 (Expected Years of Schooling) has a positive mean local coefficient of 3.4755, indicating that its effect on poverty rates is predominantly positive and varies across provinces. The coefficient of Z_1 is significant across all provinces, with 58.8% at $p < 0.001$, 35.3% at $p < 0.01$, and 5.9% at $p < 0.05$. The counterintuitive positive direction of Z_1 is discussed further in Subsection 3.7.
Overall, the spatial variation in both local coefficients confirms that a spatially varying specification is more appropriate than a globally fixed model for provincial-level variables.

3. **Random Effect**
The random effects in the HGWPR model are used to capture province-level variation in poverty rates. In this study, regencies/cities are nested within provinces, so the random intercept for province is included in the model. The random effect results are presented in Table 9.

Table 9: Random Effects Estimation Results

Group	Effect	Mean	Std. Dev.
Province	Intercept	0.0000	2.8512
Residual	–	0.1755	3.4965

Based on Table 9, the province-level random intercept standard deviation is 2.8512, confirming that poverty rates vary across provinces even after controlling for explanatory variables. The residual standard deviation of 3.4965 represents the remaining variation at the regency/city level. These findings indicate that poverty rates are shaped by both regency/city-level spatial factors and higher-level provincial differences, justifying the hierarchical structure of the HGWPR model.

3.6. Model Selection

Table 10 compares four models: OLS, Hierarchical Linear Model (HLM), standard GWR, and HGWR, evaluated using R^2 and AIC.

Table 10: Model Comparison Results

Model	R^2	AIC
OLS	0.502	3051.3
HLM (R_c^2)	0.763	2735.8
GWR	0.851	2516.8
HGWR	0.754	1372.6

[a] R^2 for HLM reported as conditional R^2 (R_c^2).
[b] AIC values are not directly comparable across models; see text.

All alternative models show improvement over OLS, indicating that the assumption of spatially homogeneous relationships is not adequate for these data. The HLM accounts for the hierarchical structure of regencies/cities nested within provinces, but it still assumes that regression coefficients are spatially constant. In contrast, GWR allows coefficients to vary spatially, which explains its higher R^2 value. However, GWR does not explicitly accommodate the nested structure of regencies/cities within provinces and may overfit the data by allowing all coefficients to vary locally.

HGWR was therefore selected as the final model because it is more consistent with the analytical structure of this study. The model simultaneously incorporates global fixed effects, local fixed effects, and province-level random effects, allowing it to account for both hierarchical dependence and spatial heterogeneity. In addition, the residual Moran's I test indicates that spatial autocorrelation is no longer significant after applying HGWR, suggesting that the model adequately captures the spatial structure of the data.

The AIC values reported in Table 10 are used as supporting information rather than as the sole basis for model selection. This is because AIC values from OLS, HLM, GWR, and HGWR are not strictly comparable across different model classes due to differences in likelihood formulation and degrees-of-freedom calculation. Thus, the selection of HGWR is based primarily on its theoretical suitability for hierarchical spatial data, its ability to address spatial non-stationarity, and its residual diagnostic performance.

3.7. Spatial Variation of Local Coefficients

Human Development Index (Z_0). Figure 2 presents the spatial distribution of local coefficients for HDI.

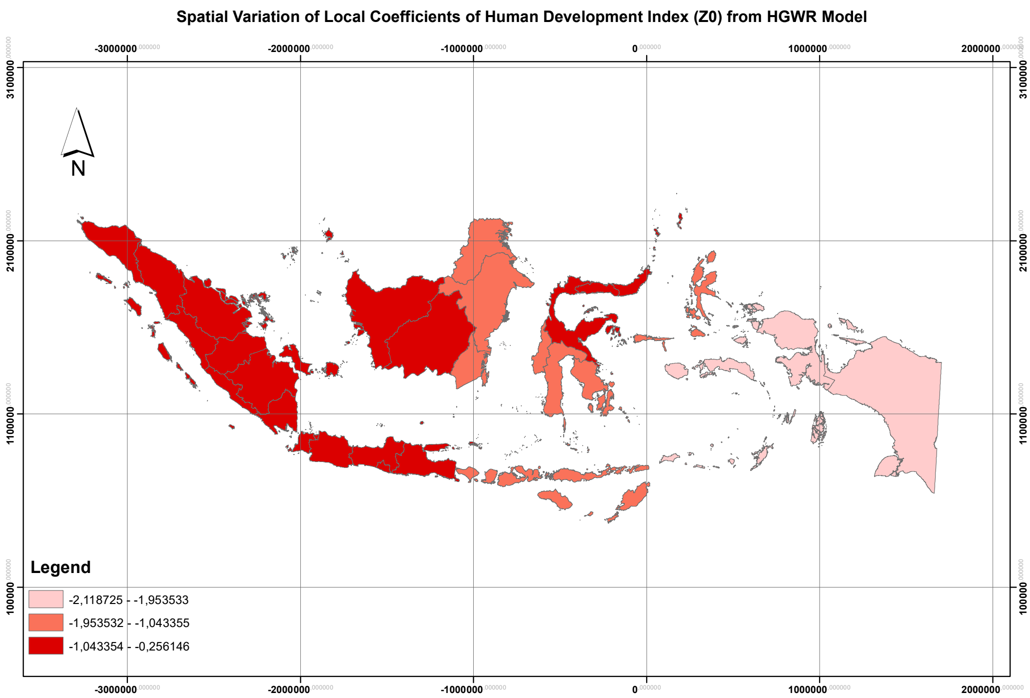

Fig. 2: Spatial Distribution of Local Coefficients for HDI (Z_0)

1. **Negative coefficients across all provinces:** HDI consistently reduces poverty nationwide.
2. **Strongest effects in Eastern Indonesia:** Provinces such as Papua, West Papua, Maluku, and North Maluku exhibit the most negative coefficients (below -1.043, corresponding to the darkest class in the legend), indicating that HDI improvements in these regions yield the largest poverty reductions.
3. **Weakest effects in Java:** Provinces in Java, particularly DKI Jakarta, West Java, and

East Java,show the least negative coefficients (above -1.043, in the lightest class of the legend), suggesting diminishing returns to HDI improvements in more developed regions.

4. **Moderate effects in Sumatra and Kalimantan:** These regions show coefficients in the intermediate range (-1.953 to -1.043), representing a transitional category..

Expected Years of Schooling (Z_1). Figure 3 presents the spatial distribution of local coefficients for expected years of schooling.

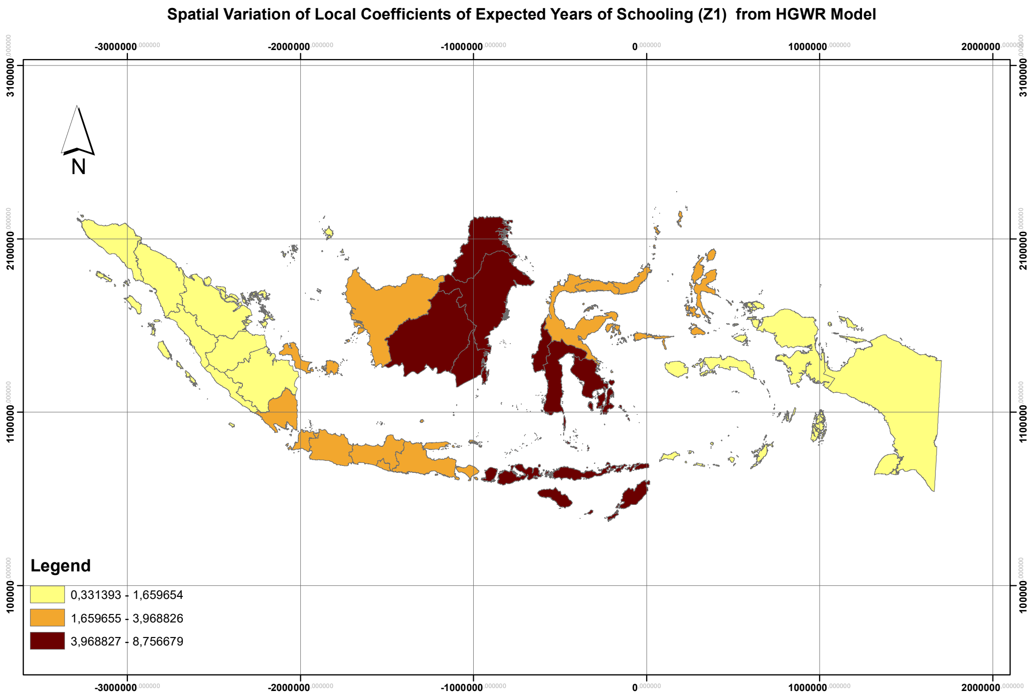

Fig. 3: Spatial Distribution of Local Coefficients for Expected Years of Schooling (Z_1)

Based on Figure 3, the following patterns are observed for Z_1:

1. **Positive coefficients dominate:** Unlike HDI, the coefficients for expected years of schooling are predominantly positive across Indonesian provinces, with a mean value of 3.4755.
2. **Strongest positive effects in Java:** Provinces in Java, particularly DKI Jakarta, East Java, and Central Java,exhibit the highest positive coefficients (in the range 3.97 to 8.76, the darkest class in the legend). This suggests that in regions with already high education levels, additional increases in schooling years are associated with higher poverty rates—a finding that warrants further investigation.
3. **Negative or weak effects in Eastern Indonesia:** Some provinces in eastern Indonesia, such as Papua and West Papua, show the lowest positive coefficients for Z_1 (in the range 0.33 to 1.66, corresponding to the lightest class in the legend), indicating that the positive association between expected years of schooling and poverty is weakest in these regions.
4. **Possible explanations:** The positive association between expected years of schooling and poverty in developed regions (Java) may reflect labor market saturation, where higher education does not guarantee employment, or measurement issues where education captures different socioeconomic phenomena in more developed contexts.

The positive mean coefficient of Z_1 (expected years of schooling) is counterintuitive and therefore requires cautious interpretation. To examine whether this result reflects a substantive positive association or is influenced by overlap between expected years of schooling and HDI, additional robustness checks were conducted.

First, the Pearson correlation between Z_0 and Z_1 was examined. The correlation was $r = 0.546$ ($t = 14.649$, $df = 505$, $p < 0.001$, 95% CI: [0.482, 0.604]), indicating a moderate positive association. Although the VIF values for both variables remain below the conventional threshold, this correlation suggests that Z_0 and Z_1 share overlapping information, particularly because education is one component of the Human Development Index.

Second, a reduced HGWR model excluding Z_0 was estimated while retaining Z_1. In the full HGWR model, the mean local coefficient of Z_1 was 3.4755. After excluding Z_0, the mean local coefficient of Z_1 decreased to 1.9843. Its statistical strength also weakened substantially, with only 14.7% of provinces significant at $p < 0.05$ and 58.8% marginally significant at $p < 0.10$. This decrease in both magnitude and statistical strength indicates that the positive coefficient of expected years of schooling is sensitive to the inclusion of HDI in the model.

These results suggest that the positive coefficient of Z_1 should not be interpreted as evidence that higher expected years of schooling directly increases poverty. Rather, the result may reflect shared information between Z_0 and Z_1, the composite nature of HDI, and the distinction between expected schooling and realized socioeconomic outcomes. Expected years of schooling represents future educational opportunity rather than current labor-market absorption or actual household welfare. Therefore, the positive association is interpreted as an exploratory statistical pattern rather than a causal effect.

3.8. Residual Diagnostics

To validate the HGWR model, the residuals were tested for spatial autocorrelation using Moran's I test. The test was performed under the randomization assumption with a spatial weight matrix based on regency/city-level neighbors.Table 11 presents the results.

Table 11: Moran's I Test for HGWR Residuals

Statistic	Value	Expectation	Variance	p-value
Moran's I	-0.000641	-0.001976	0.0000396	0.416

The Moran's I test result is -0.000641 with a p-value of 0.416, which is not statistically significant. This indicates that the residuals are randomly distributed across space, with no evidence of spatial clustering. Thus, the HGWR model has effectively captured the spatial heterogeneity in the data, and no substantial spatial dependence remains in the residuals.

4. Conclusion

This study applied the HGWR model to analyze determinants of the poverty headcount rate (P_0) across 507 regencies/cities in Indonesia using 2024 BPS cross-sectional data. HGWR demonstrated substantially better fit than the global OLS model, with R^2 increasing from 0.502 to 0.754 and AIC decreasing from 3051.3 to 1372.6. The elimination of significant residual spatial autocorrelation—from Moran's $I = 0.2216$ ($p < 0.001$) in OLS residuals to $I = -0.000641$ ($p = 0.416$) in HGWR residuals confirms that the model effectively captures spatial heterogeneity in Indonesian poverty data.

At the regency/city level, mean years of schooling ($\hat{\beta} = -2.197$, $p < 0.001$) and the poverty line ($\hat{\beta} = -0.307$, $p < 0.001$) are significant negative determinants of poverty, while adjusted per capita expenditure does not reach statistical significance ($\hat{\beta} = 0.060$, $p = 0.240$). At the provincial level, HDI consistently shows a negative association across all provinces (mean local coefficient $= -0.8082$), with the strongest effects in eastern Indonesia (coefficients < -0.95) and weakest in Java (coefficients > -0.65). Expected years of schooling exhibits a positive mean local coefficient (3.4755), highest in Java, interpreted cautiously as it may reflect overlap with HDI, labor market conditions in more developed regions, or limitations of the cross-sectional design rather than a direct causal relationship.

The spatial patterns in this study suggest that place-based approaches to poverty reduction may be worth further consideration. In particular, HDI improvements in eastern Indonesia appear to be associated with larger marginal reductions in poverty compared to Java, where the HDI–poverty relationship appears weaker. Since this study is based on cross-sectional observational data, these associations do not establish causal effects; policy decisions should be informed by additional longitudinal studies or program evaluations.

This study has several limitations. First, the cross-sectional design precludes causal inference, and observed associations may reflect reverse causality or omitted variable bias. Second, Euclidean distance may not fully capture spatial relationships in Indonesia's archipelagic geography. Third, the model comparison includes OLS, HLM, GWR, and HGWR; future work may consider additional specifications such as Multiscale GWR (MGWR) or spatial panel models. Fourth, the positive coefficient of expected years of schooling warrants further investigation using panel data or instrumental variable approaches. Future research could extend this framework by incorporating time-series data, additional socioeconomic variables such as infrastructure access or employment rates, and alternative distance metrics better suited to Indonesia's geographic context.

CRediT Authorship Contribution Statement

Debora Dwi Kurniawati: Conceptualization, methodology, writing—original draft. **Henny Pramoedyo**: Supervision, validation, editing. **Suci Astutik**: Supervision, validation, editing.

Declaration of Generative AI and AI-assisted technologies

During the preparation of this manuscript, the author used ChatGPT (OpenAI, GPT-4) and DeepSeek (DeepSeek AI) for language refinement, proofreading, and formatting assistance. All data analysis, model estimation, and mapping were performed by the author without AI assistance. The author takes full responsibility for the content of this manuscript.

Declaration of Competing Interest

The author declares no competing interests.

Funding and Acknowledgments

This research received no external funding. The author thanks the supervisors and the institution for their support and guidance. The data used in this study were obtained from the Badan Pusat Statistik (BPS), Indonesia.

Data and Code Availability

The data used in this study are publicly available from the Badan Pusat Statistik (BPS) Indonesia through the following specific sources:

- **Poverty headcount rate (P_0) and poverty line (X_3):** [1]
- **Adjusted per capita expenditure (X_1):** *Pengeluaran per Kapita Disesuaikan (Rupiah), 2024*, BPS Strategic Data.[2]

[1] https://www.bps.go.id/id/statistics-table/2/NjIxIzI=/persentase-penduduk-miskin-menurut-kabupaten-kota.html

[2] https://www.bps.go.id/id/statistics-table/2/NDE2IzI=/-metode-baru-pengeluaran-per-kapita-disesuaikan.html

- **Mean years of schooling (X_2), expected years of schooling (Z_1), and Human Development Index (Z_0):** *Indeks Pembangunan Manusia 2024*, BPS, November 2024.[3]

The spatial weight construction used Euclidean distances between WGS 1984 provincial centroids, adaptive Gaussian kernel, $h^* = 15$ nearest-neighbor provinces, implemented using the `hgwr` package (version [0.6-2]) in R (version [4.4.1]) [18]. The R analysis code and ArcGIS project files are available from the corresponding author upon reasonable request.

References

[1] Badan Pusat Statistik. *Profil Kemiskinan di Indonesia September 2024*. Berita Resmi Statistik No. 05/01/Th. XXVIII. Jakarta, Indonesia: Badan Pusat Statistik, Jan. 2025. https://www.bps.go.id/id/pressrelease/2025/01/15/2401/persentase-penduduk-miskin-september-2024-turun-menjadi-8-57-persen-.html.

[2] Chaeruniza Fitriyani, Rusiti Rusiti, Qori'atul Septiavin, and Euis Ratna Sari. "Analisis Determinan Ketimpangan di Indonesia pada Masa Covid-19". In: *Bappenas Working Papers* 7.3 (2024), pp. 293–307. DOI: 10.47266/bwp.v7i3.354.

[3] Ragdad Miranti. "Is regional poverty converging across Indonesian districts? A distribution dynamics and spatial econometric approach". In: *Asia-Pacific Journal of Regional Science* 5 (Apr. 2021). DOI: 10.1007/s41685-021-00199-3.

[4] Lilik Sugiharti, Miguel Angel Esquivias, Mohd Shahidan Shaari, Ari Dwi Jayanti, and Abdul Rahim Ridzuan. "Indonesia's poverty puzzle: Chronic vs. transient poverty dynamics". In: *Cogent Economics & Finance* 11.2 (2023), p. 2267927. DOI: 10.1080/23322039.2023.2267927.

[5] Muhammad Arif, Lutfi Muta'ali, and Rijanta. "Mapping poverty traps in Indonesia: a spatial perspective". In: *Regional Statistics* 15 (Jan. 2025), pp. 341–364. DOI: 10.15196/RS150207.

[6] Francis Huang. *Practical Multilevel Modeling Using R*. Sage Publications Inc, Dec. 2022. DOI: 10.4135/9781071846162.

[7] Taylor Oshan, Jordan Smith, and Alexander Fotheringham. "Targeting the spatial context of obesity determinants via multiscale geographically weighted regression". In: *International Journal of Health Geographics* 19 (Apr. 2020). DOI: 10.1186/s12942-020-00204-6.

[8] A. Stewart Fotheringham, Chris Brunsdon, and Martin Charlton. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, England: Wiley, 2002. https://books.google.com/books/about/Geographically_Weighted_Regression.html?id=GFjFX-3dH68C.

[9] Devi Hasibuan, Heribertin Teku, Maria Putri, Yudi Setyawan, and Rokhana Bekti. "Application of Geographically Weighted Regression Method on the Human Development Index of Central Java Province". In: *Enthusiastic : International Journal of Applied Statistics and Data Science* (Oct. 2023), pp. 189–201. DOI: 10.20885/enthusiastic.vol3.iss2.art6.

[10] Thierry Feuillet, Etienne Cossart, Hélène Charreire, Arnaud Banos, Hugo Pilkington, Virginie Chasles, Serge Hercberg, Mathilde Touvier, and Jean Michel Oppert. "Hybridizing geographically weighted regression and multilevel models: a new approach to capture contextual effects in geographical analyses". In: *Geographical Analysis* 56.3 (2024), pp. 554–572. DOI: 10.1111/gean.12385.

[3] https://www.bps.go.id/id/statistics-table/2/NDE1IzI=/-metode-baru-rata-rata-lama-sekolah.html, https://www.bps.go.id/id/statistics-table/2/NDEzIzI=/-metode-baru-indeks-pembangunan-manusia.html,https://www.bps.go.id/en/statistics-table/2/NDE3IzI=/-metode-baru--harapan-lama-sekolah.html

[11] Yigong Hu, Richard Harris, Richard Timmerman, and Binbin Lu. "A Hierarchical and Geographically Weighted Regression Model and Its Backfitting Maximum Likelihood Estimator". In: *12th International Conference on Geographic Information Science (GIScience 2023)*. Ed. by Roger Beecham, Jed A. Long, Dianna Smith, Qunshan Zhao, and Sarah Wise. Vol. 277. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 39:1–39:6. DOI: 10.4230/LIPIcs.GIScience.2023.39.

[12] Yigong Hu, Binbin Lu, Yong Ge, and Guanpeng Dong. "Uncovering spatial heterogeneity in real estate prices via combined hierarchical linear model and geographically weighted regression". In: *Environment and Planning B: Urban Analytics and City Science* 49.6 (2022), pp. 1715–1740. DOI: 10.1177/23998083211063885.

[13] Mega Novitasari, Doddy Aditya Iskandar, et al. "Spatial spillover impact of sectoral government expenditure on poverty alleviation in South Kalimantan Province". In: *The Journal of Indonesia Sustainable Development Planning* 3.3 (2022), pp. 207–221. DOI: 10.46456/jisdep.v3i3.361.

[14] Román Salmerón, Catalina Garcia, and Jose Pérez. "A Redefined Variance Inflation Factor: Overcoming the Limitations of the Variance Inflation Factor". In: *Computational Economics* 65 (Mar. 2024), pp. 337–363. DOI: 10.1007/s10614-024-10575-8.

[15] Daojun Zhang and Yu Zhang. "Moran's I of VRPAD: A human activity-sensitive spatial pattern index for vegetation restoration evaluation". In: *Journal of Environmental Management* 387 (2025), p. 125948. DOI: 10.1016/j.jenvman.2025.125948.

[16] Bülent Gülöğlu, Süleyman Taşpınar, Osman Doğan, and Anil K. Bera. "Testing Homoskedasticity in Spatial Panel Data Models". In: *Econometrics and Statistics* (2024). DOI: 10.1016/j.ecosta.2024.04.003.

[17] Denny Kerkhoff and Fridtjof Nussbeck. "Obtaining Sound Intraclass Correlation and Variance Estimates in Three-Level Models: The Role of Sampling-Strategies". In: *Methodology* 18 (Mar. 2022), pp. 5–23. DOI: 10.5964/meth.7265.

[18] Yigong Hu, Richard Harris, Richard Timmerman, and Binbin Lu. "A Backfitting Maximum Likelihood Estimator for Hierarchical and Geographically Weighted Regression Modelling, with a Case Study of House Prices in Beijing". In: *International Journal of Geographical Information Science* 38.12 (2023), pp. 2458–2491. DOI: 10.1080/13658816.2024.2391412.

[19] Vinka Haura Nabilla, Dina Fitria, Dony Permana, and Fadhilah Fitri. "Comparison of Haversine and Euclidean Distance Formulas for Calculating Distance Between Regencies in West Sumatra". In: *UNP Journal of Statistics and Data Science* 1.3 (2023), pp. 120–125. DOI: 10.24036/ujsds/vol1-iss3/39.

[20] Tuba Koç. "Bandwidth Selection in Geographically Weighted Regression Models via Information Complexity Criteria". In: *Journal of Mathematics* 2022 (2022), p. 1527407. DOI: 10.1155/2022/1527407.