

Multiscale Geographically Weighted Spline Nonparametric Regression for Spatial Modeling of Poverty in Eastern Indonesia

Indi Rizqy Fahrani¹, Henny Pramodyo¹, Atiek Iriany^{1*}

¹Department of Statistics, Brawijaya University, Indonesia

atiekiriany@ub.ac.id

ABSTRACT

Article History:

Received : 01-07-2026

Revised : 15-07-2026

Accepted : 03-09-2026

Online : 23-10-2026

Keywords:

Poverty;

Spatial Heterogeneity;

Multiscale;

Geographically

Weighted Regression;

Truncated Spline.



Poverty in Eastern Indonesia remains spatially heterogeneous, and its determinants are unlikely to operate at a single geographic scale or to follow a predetermined functional form. This study applies a quantitative, spatial nonparametric regression approach, namely the Multiscale Geographically Weighted Spline Nonparametric Regression (MS-GWSNR) model, to poverty rates in 176 regencies/cities across Eastern Indonesia in 2025. The model uses secondary cross-sectional data from Statistics Indonesia (BPS) and the Geospatial Information Agency (BIG), with the Human Development Index (HDI), open unemployment rate (OUR), gross regional domestic product (GRDP) per capita, and population density as covariates. The model was estimated via a backfitting algorithm assigning a component-specific bandwidth to each covariate, benchmarked against single-bandwidth GWSNR and the linear MGWR. The spatial scale of each covariate was characterized by pairing its optimum bandwidth with the effective number of parameters (ENP). The results showed that the covariates operated at markedly different spatial scales: the intercept, HDI, OUR, and population density were local (bandwidth 10–14 nearest regencies), whereas GRDP per capita was nearly global (bandwidth 175, lowest ENP). HDI was the dominant determinant, significant in 89.8% of locations and generally associated with lower poverty. MS-GWSNR achieved the best fit among the compared models, raising the coefficient of determination from 85.69% (MGWR) and 87.54% (GWSNR) to 88.70% and reducing the RMSE from 3.3467 and 3.1232 to 2.9741, while also producing more stable local estimates. Notably, both MGWR and MS-GWSNR independently identified GRDP per capita as a near-global process, supporting the robustness of the multiscale finding.



<https://doi.org/10.31764/jtam.v10i4.40625>



This is an open access article under the [CC-BY-SA](#) license

A. INTRODUCTION

Poverty alleviation remains one of the central challenges of national development in Indonesia. Although the national poverty rate reached 8.25% in September 2025, its distribution across provinces is far from uniform. Nine of the ten provinces with the highest poverty rates are located in Eastern Indonesia (Kawasan Timur Indonesia, KTI), ranging from Central Papua (29.45%) and Highland Papua (27.21%) to Gorontalo (12.62%), while several other KTI provinces such as West Nusa Tenggara, Central Sulawesi, West Sulawesi, and Southeast Sulawesi also recorded rates above the national average (BPS, 2025). Following the regional classification of the Ministry of National Development Planning, KTI comprises Sulawesi, Nusa Tenggara, Maluku, and Papua (Kementerian PPN/Bappenas, 2014). The wide

disparity among regions indicates that poverty in KTI is not spatially homogeneous. This pattern is consistent with recent evidence of pronounced regional economic inequality across Eastern Indonesia (Hamrullah et al., 2026), as well as findings that a mixed geographically weighted model captured pronounced local variation in poverty determinants across districts in Nusa Tenggara Timur (Ola et al., 2024). Such heterogeneity calls for an analytical approach able to accommodate local variation across space.

Earlier studies of poverty in KTI have repeatedly detected spatial heterogeneity using Geographically Weighted Regression (GWR) and its extensions. In Papua, Lumbantoruan et al. (2023) found that GWR outperformed ordinary regression, with the influence of socioeconomic factors on poverty varying significantly across districts. In Nusa Tenggara Timur, Ola et al. (2024) similarly showed that a mixed geographically weighted model captured both local and global determinants of poverty more accurately than ordinary GWR. A related study on Sulawesi Island similarly documented pronounced spatial heterogeneity in poverty determinants using a geographically weighted approach (Laome et al., 2022). These findings suggest that each determinant may exert its influence at a different spatial scale and through relationships that are not necessarily linear.

GWR is a local regression technique that explores spatial non-stationarity by allowing parameters to vary across geographic locations, with weights formed through a kernel function governed by a single optimum bandwidth (Oshan et al., 2019). The use of a single bandwidth implicitly assumes that all relationships operate at the same spatial scale, which may not hold in practice: some processes act over broad regions while others are confined to small neighborhoods. As a result, GWR yields a best-on-average scale that can mask genuine differences among covariates (Comber et al., 2023). To overcome this limitation, GWR with flexible bandwidths was later formalized as Multiscale Geographically Weighted Regression (MGWR), in which each local coefficient is estimated with its own bandwidth interpreted as the spatial scale at which the corresponding relationship operates (Fotheringham et al., 2017). These bandwidths are obtained through model calibration, where a smaller bandwidth indicates a more local spatial process, whereas a larger bandwidth implies a more global reach (Kang & Oshan, 2025). MGWR has improved the understanding of spatial variation in diverse settings, including county-level economic development in China (Huang et al., 2023) and the analysis of environmental factors affecting cardiovascular diseases in Nanning, China (Deng et al., 2025). It has likewise proved useful for examining health outcomes in developed-country settings, including the spatial variations of COVID-19 morbidity and mortality rates across England (Morasae et al., 2024) and the spatial predictors of anemia among women in Mozambique (Gebrehana et al., 2025).

A second concern is the functional form of the relationships. Parametric regression presumes a known form and may produce biased estimates when that assumption is violated, whereas nonparametric regression lets the data determine the shape of the regression curve (Mahmoud, 2021). The truncated spline is a widely used nonparametric basis whose strength lies in following local patterns flexibly through knots that mark changes in data behavior (Putra et al., 2023), and has been applied to the achievement status of unmet need targets in East Java and the percentage of the poor population in Indonesia (Suriaslan et al., 2025). Several studies of poverty in KTI provinces support the relevance of such flexibility: a truncated spline model

of rural poverty, unemployment, and agricultural growth across Indonesia's major regions, including Sulawesi and Papua-Maluku-Nusa Tenggara, found pronounced nonlinear, region-specific effects that classical linear models could not capture (Fattah et al., 2025), and a study on Sulawesi Island found that a geographically weighted mixed model incorporating truncated splines was more appropriate than global linear regression in addressing spatial heterogeneity of poverty (Laome et al., 2022).

The integration of spline nonparametric regression with geographic weighting is known as Geographically Weighted Spline Nonparametric Regression (GWSNR) (Sifriyani et al., 2017), and has been applied to analyze the crude birth rate across 38 areas in Eastern Java (Sifriyani et al., 2018). GWSNR is flexible in functional form but still relies on a single spatial scale for all covariates, so it cannot fully represent scale differences among the explanatory variables. MGWR, conversely, accommodates multiple spatial scales but remains parametric and therefore limited in handling complex nonlinear relationships. This leaves a methodological gap. Existing multiscale approaches are constrained to linear functional forms, while existing nonparametric spatial approaches are constrained to a single spatial scale, so neither can jointly capture heterogeneity across space and nonlinearity in functional form. This complementary gap motivates a model that combines both strengths, namely the Multiscale Geographically Weighted Spline Nonparametric Regression (MS-GWSNR), whose theoretical development and estimator properties are presented in a companion paper (Fahrani et al., 2026). Rather than re-deriving the estimator here, the present work builds on that companion paper and instead emphasizes the empirical contribution: a spatially explicit account of poverty determinants in Eastern Indonesia.

Beyond poverty analysis, this combination of properties is broadly relevant because the two elements are independently well-documented across a range of fields. Nonlinear, spatially heterogeneous relationships have recently been reported in housing markets (Chen et al., 2025), environmental health risk analysis (Czarnota et al., 2015), and transportation research (Guo et al., 2025). Scale-dependent processes, in which different factors operate at different spatial extents, have similarly been documented using multiscale geographically weighted regression in public health (Deng et al., 2025; Gebrehana et al., 2025; Morasae et al., 2024) and traffic emissions modelling (Omrani et al., 2025). A model capable of jointly capturing both properties therefore addresses a combination of characteristics that, while individually common across fields, has rarely been modelled simultaneously. To the best of our knowledge, an approach that simultaneously accommodates multiscale spatial heterogeneity and nonlinear relationships has not yet been applied to poverty across KTI as a whole; previous multiscale work covered only a single province, while nonparametric studies were conducted province by province.

This study therefore applies the MS-GWSNR model to the 2025 poverty rate across all 176 regencies/cities of KTI. The objectives are threefold: to characterize the spatial scale at which each determinant of poverty operates by pairing the optimum bandwidth with the effective number of parameters (ENP); to assess the local significance and direction of each covariate; and to compare the empirical performance and estimation stability of MS-GWSNR against the linear multiscale MGWR and the single-bandwidth GWSNR.

B. METHODS

1. Data and Study Area

This study adopts a quantitative research design employing spatial nonparametric regression analysis, using secondary cross-sectional data on the 2025 poverty rate in Eastern Indonesia (KTI) obtained from Statistics Indonesia (BPS). Following the classification of the Ministry of National Development Planning, the study area covers Sulawesi, Nusa Tenggara, Maluku, and Papua, comprising 176 regencies/cities as the observation units. The geographic coordinates of each regency were obtained from the Ina-Geoportal of the Geospatial Information Agency (BIG). The variables are summarized in Table 1.

Table 1. Research Variables

Symbol	Variable	Unit	Scale
Y	Percentage of poor population (P_0)	Percent (%)	Ratio
X_1	Human Development Index (HDI)	Index (0–100)	Ratio
X_2	Open unemployment rate (OUR)	Percent (%)	Ratio
X_3	GRDP per capita (constant prices)	Thousand Rp/capita	Ratio
X_4	Population density	Persons/km ²	Ratio

2. The MS-GWSNR Model

MS-GWSNR extends GWSNR by allowing every model component to have its own bandwidth, combining the local nonlinear flexibility of the truncated spline with the multiscale spatial structure of MGWR. For an observation located at coordinates (u_i, v_i) , the model is (Fahrani et al., 2026):

$$y_i = \sum_{k=1}^p \left[\sum_{j=0}^m \beta_{kj}^{(bw_k)}(u_i, v_i) x_{ik}^j + \sum_{h=1}^r \delta_{kh}^{(bw_k)}(u_i, v_i) (x_{ik} - K_{kh})_+^m \right] + \varepsilon_i \quad (1)$$

where y_i is the response, x_{ik} the k -th predictor, $\beta_{kj}^{(bw_k)}$ and $\delta_{kh}^{(bw_k)}$ the spatially varying polynomial and truncated spline coefficients, and bw_k the component-specific bandwidth.

Because the components use different bandwidths and spatial weight matrices, the model is estimated iteratively with a backfitting algorithm initialized from the single-bandwidth GWSNR fit and iterated until convergence based on the score-of-change (SOC) criterion (Figure 1). A tolerance of 10^{-5} was adopted following convention in backfitting-based estimation, where the SOC is expected to approach zero as component estimates stabilize. This threshold balances computational cost against estimation precision.

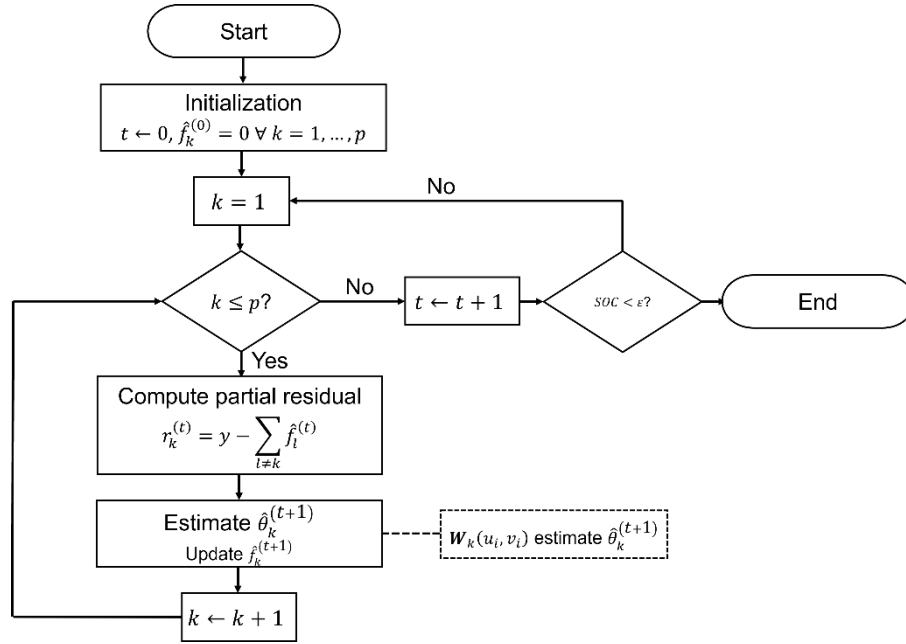


Figure 1. Flowchart of the backfitting algorithm for MS-GWSNR estimation

Following the linear smoother framework (Buja et al., 1989), the spatial scale of each component is summarized by its effective number of parameters, defined as the trace of its smoother matrix $\mathbf{S}_k$:

$$ENP_k = \text{tr}(\mathbf{S}_k) \quad (2)$$

A small bandwidth yields a large ENP, indicating a locally flexible component, while a large bandwidth yields a small ENP, indicating a near-global or stationary effect. Pairing the optimum bandwidth with the ENP thus indicates scale differences among covariates (Yu et al., 2019).

3. Analysis Procedure

The following is a flowchart of the MS-GWSNR analysis procedure, as shown in Figure 2.

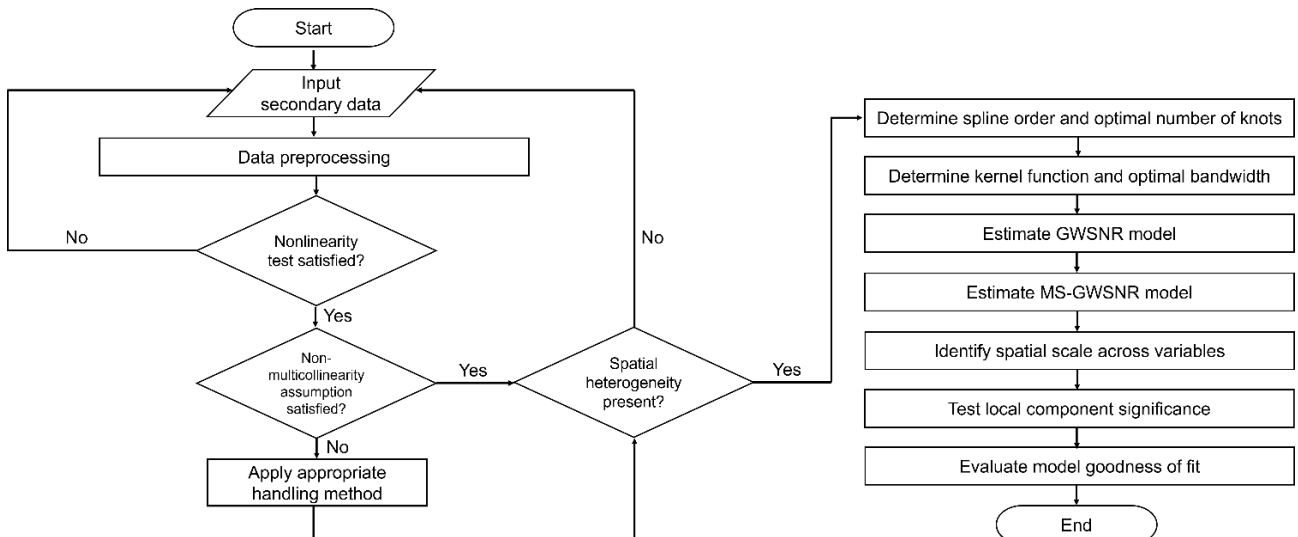


Figure 2. Flowchart of the MS-GWSNR analysis procedure

The analysis followed the stages summarized in Figure 2, from data preprocessing through model estimation to goodness-of-fit evaluation. Data preprocessing comprised a Ramsey RESET linearity test for each covariate, a variance inflation factor check for multicollinearity, and a spatial heterogeneity test. The spline order and number of knots were selected by minimizing the Generalized Cross-Validation (GCV) criterion, and the kernel and bandwidth by minimizing the Cross-Validation (CV) criterion under an adaptive weighting scheme expressed as a number of nearest neighbors. GCV and CV are standard data-driven criteria that balance fit against overfitting without requiring a pre-specified functional form. An adaptive rather than a fixed-distance bandwidth was used because regencies across KTI vary substantially in size and spatial density, a fixed distance would oversmooth dense clusters and undersmooth sparse areas, whereas a nearest-neighbor bandwidth adjusts to local point density. Inter-location distances were computed geodesically, since the study area spans several UTM zones; covariates were standardized using the Z-score transformation given their widely differing ranges; and missing OUR values in seven regencies were imputed using the weighted average of the five nearest neighbors, consistent with the principle that nearby locations share similar characteristics.

Following data preprocessing, GWSNR was estimated first as the starting value for backfitting, followed by MS-GWSNR, whose optimum bandwidth and ENP per component were examined to identify spatial scales. Local significance at each location was tested using a t-statistic from the component standard error, allowing a covariate to be significant in some regions but not others. As a benchmark, the linear multiscale MGWR was estimated using the same adaptive kernel and cross-validated bandwidths; since MGWR shares MS-GWSNR's multiscale structure but assumes linear relationships, the contrast isolates the contribution of the spline component. Goodness of fit across MGWR, GWSNR, and MS-GWSNR was compared using the residual sum of squares (RSS), the coefficient of determination (R^2), and the root mean square error (RMSE), complemented by maps of each covariate's spatial contribution.

C. RESULT AND DISCUSSION

1. Data Description

The 176 KTI regencies displayed substantial inter-regional variation in 2025. The percentage of poor population ranged from 3.34% to 42.56% with a mean of 15.34%, reflecting a wide poverty gap. The most striking variation appeared in GRDP per capita, whose standard deviation exceeded its mean, indicating sharp economic inequality, while population density ranged from 1 to more than 8,000 persons per km². Descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of the Research Variables

Variable	Minimum	Maximum	Mean	Std. dev.
Y	3.34	42.56	15.344	8.873
X_1	40.35	86.36	70.683	7.262
X_2	1.24	11.37	3.868	1.908
X_3	4,590.00	621,817.00	38,208.648	66,772.062
X_4	1.00	8,334.00	326.301	941.353

The spatial distribution of poverty (Figure 3) revealed a clear west-to-east gradient. Sulawesi and parts of Maluku showed relatively low poverty (3.34 – 13.22%), whereas Papua, particularly Highland, Central, and South Papua, was dominated by very high rates (28.1 – 42.56%). This large spatial variation provided an initial indication that the relationships between the explanatory factors and poverty are unlikely to be uniform across regions, making a location-aware modeling approach relevant.

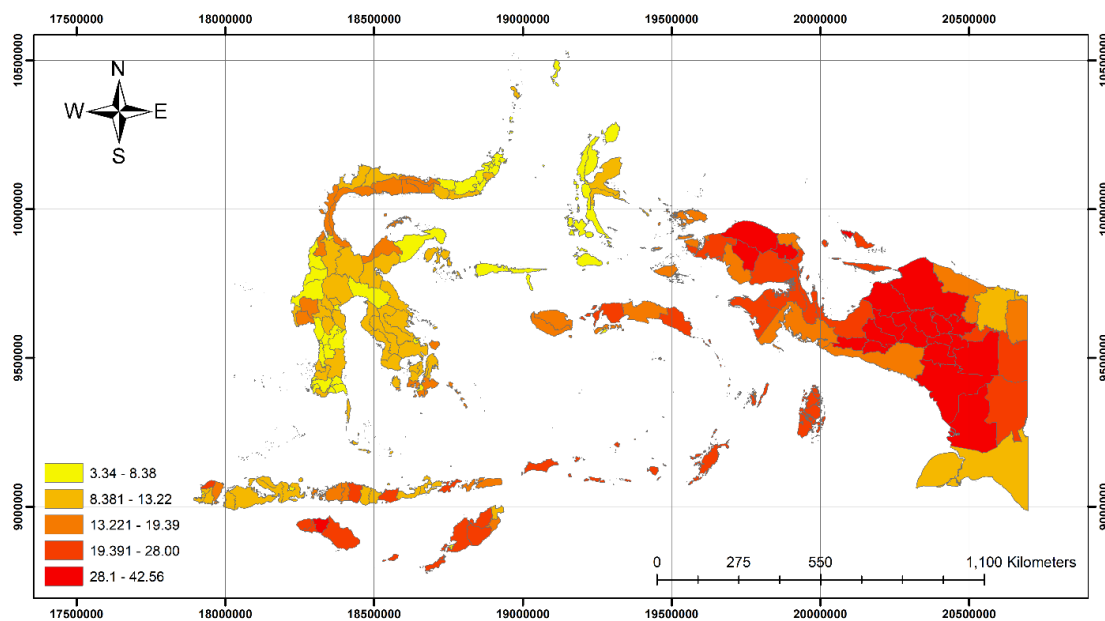


Figure 3. Choropleth map of poverty distribution across 176 KTI regencies, 2025

2. Preliminary Tests

The Ramsey RESET test indicated nonlinear relationships for HDI, GRDP per capita, and population density, while OUR appeared linear (Table 3). Because at least some covariates departed from linearity, a flexible spline-based specification was justified. Classical assumption checks further supported a geographically weighted treatment: the spatial heterogeneity test confirmed significant non-stationarity, indicating that a single global model would be inadequate and that local estimation was warranted.

Table 3. Ramsey RESET Linearity Test

Variable	F-statistic	p-value	Conclusion
X_1	7.6630	0.000649	Nonlinear
X_2	0.4493	0.638800	Linear
X_3	20.1024	0.000000	Nonlinear
X_4	12.7450	0.000007	Nonlinear

3. Spline Specification and Bandwidth Selection

Evaluation of the GCV criterion yielded an optimum order $m^* = 2$ with $r^* = 1$ knot for every covariate, producing a minimum GCV of 29.4067. Each covariate was therefore modeled with a second-order truncated spline with one knot. Under the adaptive scheme, four kernels were compared; the Exponential kernel achieved the smallest CV (21.2659) at a bandwidth of 17

nearest regencies (Table 4) and was thus adopted both for GWSNR and as the kernel within the MS-GWSNR backfitting algorithm.

Table 4. Kernel and Optimum Bandwidth Selection

Kernel	Bandwidth (regencies)	CV
Exponential	17	21.2659
Gaussian	24	21.9420
Bisquare	104	22.5809
Tricube	106	22.5981

4. GWSNR and MS-GWSNR Estimation

The single-bandwidth GWSNR model (Exponential kernel, $bw = 17$) explained 87.54% of the variation in poverty ($RSS = 1716.7862$; $RMSE = 3.1232$) and provided the initial values for the backfitting algorithm. The MS-GWSNR backfitting converged at the 160th iteration with a final SOC of 9.96×10^{-6} , below the 10^{-5} tolerance, with the SOC decreasing consistently across iterations. The resulting component-specific bandwidths ranged from 10 to 175 nearest regencies (Table 5). The intercept, HDI, OUR, and population density had small bandwidths (10–14), indicating locally varying relationships, whereas GRDP per capita had a markedly larger bandwidth of 175, indicating a broad, nearly homogeneous effect. This contrasts sharply with GWSNR, which forced all components onto a single scale of 17.

Table 5. Optimum Bandwidth and ENP per Component (MS-GWSNR)

Component	MS-GWSNR bw	GWSNR bw	ENP
Intercept	10	17	13.168
X_1	14	17	17.771
X_2	10	17	27.579
X_3	175	17	4.236
X_4	14	17	18.024

The ENP values confirmed the inverse relationship with bandwidth: OUR, with the smallest bandwidth (10), had the highest ENP (27.579), while GRDP per capita, with the largest bandwidth (175), had the lowest ENP (4.236). The total model ENP was 80.7785. Because each spline basis is estimated locally through geographic weighting, the effective number of parameters can greatly exceed the number of basis functions when the bandwidth is small. The large dispersion of ENP across components provided empirical support for a multiscale approach, since the single-bandwidth assumption of GWSNR cannot accommodate such scale differences. The score-of-change decreased monotonically across iterations and the backfitting algorithm converged at the 160th iteration (final SOC = 9.96×10^{-6} , below the 10^{-5} tolerance).

5. Local Significance of Components

Because estimation is local, each component produced 176 location-specific values and t-statistics. Using the model degrees of freedom ($df = 176 - 80.7785 = 95.2215$) and an error standard deviation of 4.04339, the critical value at the 5% level was 1.9853. Table 6 summarizes the distribution of component values and the share of locations at which each component was significant. The intercept was significant everywhere, HDI in 89.8% of

locations, population density in 44.9%, while OUR and GRDP per capita were significant in only 5.1% and 4.5% of locations, respectively.

Table 6. Component Summary and Share of Significant Locations (MS-GWSNR)

Component	Min	Median	Max	Mean SE	Sig. (%)
Intercept	12.2121	14.2643	19.9701	0.6912	100
X_1	-9.1466	-0.5992	11.4644	0.7144	89.8
X_2	-7.9386	-0.0835	5.4724	0.9793	5.1
X_3	-1.6265	-0.1666	1.2997	0.4430	4.5
X_4	-3.2181	-0.6800	5.4105	0.7863	44.9

6. Model Comparison

Table 7 compares the goodness of fit of the three multiscale and spline variants. MS-GWSNR achieved the best fit on all measures, with the lowest RSS (1556.780), the highest R^2 (88.70%), and the lowest RMSE (2.9741). Reading the table along its two design axes is informative. Moving from MGWR to MS-GWSNR isolates the effect of adding spline flexibility while holding the multiscale structure fixed: the RSS fell from 1971.269 to 1556.780 and R^2 rose from 85.69% to 88.70%, indicating that the nonlinear component contributes materially beyond what a linear multiscale model captures. Moving from GWSNR to MS-GWSNR isolates the effect of adding multiscale bandwidths while holding the spline form fixed, with a further reduction in RSS from 1716.786 to 1556.780. Both contrasts favor MS-GWSNR, supporting the view that the two ingredients are complementary rather than redundant.

Table 7. Goodness-of-Fit Comparison of MGWR, GWSNR, and MS-GWSNR

Model	RSS	R^2 (%)	RMSE
MGWR (multiscale, linear)	1971.269	85.69	3.3467
GWSNR (single bandwidth, spline)	1716.786	87.54	3.1232
MS-GWSNR (multiscale, spline)	1556.780	88.70	2.9741

Two qualifications keep this comparison honest. First, the gains in fit come with greater model complexity, since MS-GWSNR estimates a larger effective number of parameters than the linear MGWR; the choice between them therefore depends on whether the analytical goal is parsimony or a richer description of spatial and functional structure. Second, MGWR was estimated through a separate implementation (the GWmodel package) rather than the backfitting framework used for MS-GWSNR, so the comparison should be read as a benchmark against a standard tool rather than a controlled ablation. The strongest independent justification for the spline component does not rest on these fit metrics at all but on the Ramsey RESET results, which identified nonlinear relationships in three of the four covariates before any model was fitted.

A noteworthy point of agreement strengthens the multiscale finding: MGWR and MS-GWSNR, despite differing in functional form and implementation, independently identified GRDP per capita as a near-global process (optimum bandwidth of 174 and 175 nearest regencies, respectively) while assigning local bandwidths to the remaining covariates. The recovery of the same spatial-scale structure by two distinct methods indicates that this structure is a property of the data rather than an artifact of a single model. This pattern is

consistent with multiscale poverty studies elsewhere, such as in Hubei Province, China, which likewise report that broad geographical and economic-endowment factors tend to have global effects, while social-informatization and institutional determinants act more locally (Liu et al., 2023). The following are paired spatial-contribution maps for GRDP per capita and population density, as shown in Figure 4.

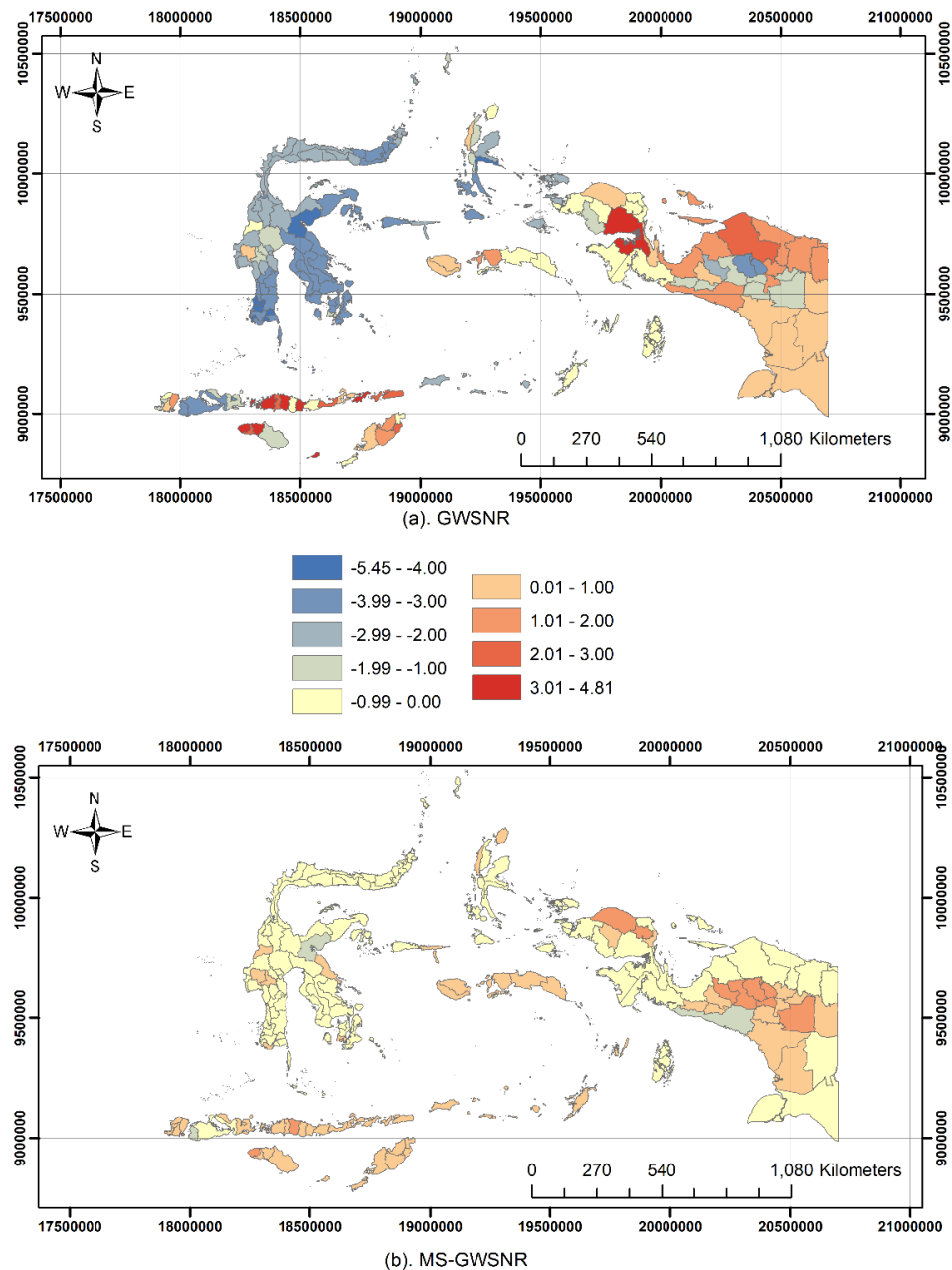


Figure 4. Paired spatial-contribution maps for GRDP per capita and population density under (a) GWSNR and (b) MS-GWSNR

Maps of the spatial contribution of each covariate made the difference clearer. For GRDP per capita in particular, the GWSNR model's local coefficients exhibit a wide and unstable variation, ranging from -5.45 to 4.81 (Figure 4a), where the most pronounced negative contributions are clustered in Sulawesi, while the largest positive impacts are observed in NTT and Papua. This

pattern reflects a scale mismatch, where the fixed bandwidth of GWSNR forces GRDP per capita to be modeled locally, resulting in erratic estimates. In contrast, the MS-GWSNR model (Figure 4b) drastically compresses this variation, with almost all regions displaying a narrow contribution range close to zero. This massive reduction in variance provides the strongest evidence that GRDP per capita operates at a global scale, as confirmed by its optimum bandwidth of 175 and the smallest ENP of 4.236. When GRDP per capita is modeled at its appropriate scale, its effect on poverty proves to be relatively minor and homogeneous across Eastern Indonesia (KTI). Ultimately, these results confirm that the principal advantage of MS-GWSNR is not only a better overall model fit but also more stable, realistic, and interpretable local estimates.

7. Interpretation of Determinants

The four determinants operate at distinct spatial scales. HDI was the dominant, locally varying factor (bandwidth 14, ENP 17.771), significant in 89.8% of regions with a negative median component, so higher human development was generally associated with lower poverty, though its strength varied spatially (range -9.15 to 11.46). Population density was also local (bandwidth 14, ENP 18.024) and significant in 44.9% of regions, with both direction and magnitude differing across space. OUR had the smallest bandwidth (10) and the highest ENP (27.579), a highly local pattern, yet was significant in only 5.1% of regions with a near-zero median. GRDP per capita behaved oppositely, operating near-globally (bandwidth 175, ENP 4.236) and significant in only 4.5% of regions, so its influence on poverty was weak and relatively homogeneous. These scale differences justify a multiscale approach in which each covariate is estimated at its most relevant spatial scale, consistent with the broader spatial-poverty literature in Indonesia that reports locally varying determinants (Ola et al., 2024; Saifudin et al., 2025).

The dominance of HDI is consistent with Ola et al. (2024), who similarly found school participation and access to basic services among the most frequently significant local determinants of poverty in Nusa Tenggara Timur. The weak and highly localized significance of OUR is consistent with Lumbantoruan et al. (2023), who similarly found that the unemployment rate had no significant effect on poverty at any location in Papua, suggesting that conventional unemployment measures may be a persistently weak predictor of poverty across KTI regardless of the specific spatial modelling approach used.

The local scale of HDI and population density is consistent with the fact that education, health, and infrastructure services are administered and delivered at the district level, so their influence on poverty is inherently bound to local institutional capacity; this implies that human development interventions are best designed and evaluated district by district rather than through uniform national targets. The near-global scale of GRDP per capita, in contrast, reflects the diffuse nature of macroeconomic growth, which spreads through inter-regional trade and fiscal transfers rather than remaining confined to the district where it is generated; this suggests that GRDP-focused growth policies alone are unlikely to be an effective lever for poverty reduction in KTI, as their effect is both weak and spatially homogeneous. The highly local yet largely insignificant effect of OUR points to a similar caution: conventional open-unemployment measures may inadequately capture labor market conditions in regions where

informal employment dominates, a limitation worth accounting for in future poverty-alleviation targeting.

Several limitations should be acknowledged. First, the backfitting algorithm required 160 iterations to converge, and its computational cost may become prohibitive for larger datasets or finer spatial resolutions than the 176 regencies analyzed here. Second, the estimated spatial scales are conditional on the kernel function and bandwidth-selection criterion (CV) used in this study; alternative criteria or kernel forms could plausibly yield different bandwidth estimates, and the sensitivity of the multiscale structure to these choices was not formally assessed. Third, the cross-sectional nature of the data means the model captures a single snapshot of spatial heterogeneity in 2025 and cannot account for how these spatial scales might shift over time as regional economies and infrastructure develop.

D. CONCLUSION AND SUGGESTIONS

Applying MS-GWSNR to the 2025 poverty rate across 176 regencies of Eastern Indonesia showed that the determinants of poverty operate at distinct spatial scales. The optimum bandwidths ranged from 10 to 175 nearest regencies, with the intercept, HDI, OUR, and population density behaving locally (bandwidth 10–14) and GRDP per capita behaving globally (bandwidth 175). HDI was the dominant determinant, significant in 89.8% of regions and generally associated with lower poverty, followed by population density (44.9%), whereas OUR and GRDP per capita were significant in only 5.1% and 4.5% of regions, respectively. MS-GWSNR achieved the best fit among the compared models, improving on both the linear multiscale MGWR (R^2 85.69%, RMSE 3.3467) and the single-bandwidth GWSNR (R^2 87.54%, RMSE 3.1232) to reach an R^2 of 88.70% and an RMSE of 2.9741. Importantly, MGWR and MS-GWSNR independently recovered the same spatial-scale structure, with GRDP per capita near-global and the other covariates local, indicating that this structure reflects the data rather than a single model.

Beyond these empirical results, the principal contribution of this study is methodological: MS-GWSNR is, to our knowledge, the first model to jointly accommodate multiscale spatial heterogeneity and nonlinear functional form within a single nonparametric framework, and its application here is the first to characterize poverty determinants across the whole of KTI rather than a single province. The gains in fit come with greater complexity, so the contribution of MS-GWSNR is best understood as offering a richer and more faithful description of spatial heterogeneity and nonlinear structure rather than a uniformly superior alternative. These findings suggest that policy responses to poverty in KTI should be differentiated according to the spatial scale at which each determinant operates, with particular emphasis on human development.

Future work may extend MS-GWSNR to a spatio-temporal setting to capture how these spatial scales evolve over time, develop bootstrap-based or other formal inference procedures for testing whether a covariate's bandwidth is statistically distinguishable from purely local or purely global, evaluate the sensitivity of the results to alternative missing-data imputation strategies such as multiple imputation or model-based approaches, benchmark MS-GWSNR against alternative nonparametric spatial approaches such as spatially varying coefficient models estimated via generalized additive models or spatial machine-learning methods, and

apply the model to other domains, such as housing markets and environmental health risk, where nonlinear and multiscale properties have been documented independently but not yet jointly modelled.

ACKNOWLEDGEMENT

The authors highly appreciate BPS-Statistics Indonesia and the Geospatial Information Agency (BIG) for providing the baseline socio-economic data and geospatial datasets for this research. Without their comprehensive data support, this spatial analysis would not have been possible.

REFERENCES

- BPS. (2025). *Profil Kemiskinan di Indonesia September 2025*. <https://www.bps.go.id/id/pressrelease/2026/02/05/2536/persentase-penduduk-miskin-september-2025-turun--menjadi-8-25-persen-.html>
- Buja, A., Hastie, T., & Tibshirani, R. (1989). Linear Smoothers and Additive Models. *The Annals of Statistics*, 17(2), 453–510. https://hastie.su.domains/Papers/lsam_annals.pdf
- Chen, Y., Ye, Y., Liu, X., Yin, C., & Jones, C. A. (2025). Examining the nonlinear and spatial heterogeneity of housing prices in urban Beijing: an application of GeoShapley. *Habitat International*, 162, 103439. <https://doi.org/10.1016/j.habitatint.2025.103439>
- Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., Wang, Y., & Harris, P. (2023). A Route Map for Successful Applications of Geographically Weighted Regression. *Geographical Analysis*, 55, 155–178. <https://doi.org/https://doi.org/10.1111/gean.12316>
- Czarnota, J., Wheeler, D. C., & Gennings, C. (2015). Evaluating Geographically Weighted Regression Models for Environmental Chemical Risk Analysis. *Cancer Informatics*, 14s2, CIN.S17296. <https://doi.org/10.4137/CIN.S17296>
- Deng, S., Zhu, S., Chen, X., Liang, J., & Zheng, R. (2025). A Multi-Scale Geographically Weighted Regression Approach to Understanding Community-Built Environment Determinants of Cardiovascular Disease: Evidence from Nanning, China. *ISPRS International Journal of Geo-Information*, 14(9), 362. <https://doi.org/10.3390/ijgi14090362>
- Fahrani, I. R., Pramoedyo, H., & Iriany, A. (2026). A Multiscale Extension of Geographically Weighted Spline Nonparametric Regression: Model Formulation and Theoretical Properties. *CAUCHY: Jurnal Matematika Murni Dan Aplikasi*, 11(2, in press), 1580–1599. <https://doi.org/10.18860/cauchy.v11i2.42954>
- Fattah, S., Rahman Razak, A., Yusuf, M. A., & Fernandes, A. A. R. (2025). Econometric Modelling of the Rural Poverty, Unemployment and the Agricultural Sector Using a Truncated Spline Approach with Longitudinal Data. *Economies*, 13(9), 273. <https://doi.org/10.3390/economies13090273>
- Fotheringham, A., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, 107(6), 1247–1265. <https://doi.org/https://doi.org/10.1080/24694452.2017.1352480>
- Gebrehana, D. A., Tamir, T. T., Molla, G. E., Kebede, Y., Tegegne, D., Nigatu, S. G., & Nigatu, A. M. (2025). Spatial variation and predictors of anemia among women of reproductive age in Mozambique, 2022/23: a multiscale geographically weighted regression. *Frontiers in Public Health*, 13. <https://doi.org/10.3389/fpubh.2025.1502177>
- Guo, K., Tang, R., Pan, H., Zhang, D., Liu, Y., & Shi, Z. (2025). Activity Spaces in Multimodal Transportation Networks: A Nonlinear and Spatial Analysis Perspective. *ISPRS International Journal of Geo-Information*, 14(8), 281. <https://doi.org/10.3390/ijgi14080281>
- Hamrullah, H., Saudi, N. D. S., Djam'an, F., & Hamzah, S. (2026). Spatial Spillovers in Regional Economic Inequality: Evidence from Macro Indicators—Remote Sensing in Eastern Indonesia. *Economies*, 14(4), 109. <https://doi.org/10.3390/economies14040109>
- Huang, Z., Li, S., Peng, Y., & Gao, F. (2023). Spatial Non-Stationarity of Influencing Factors of China's County Economic Development Base on a Multiscale Geographically Weighted Regression

- Model. *ISPRS International Journal of Geo-Information*, 12, 109. <https://doi.org/10.3390/ijgi12030109>
- Kang, W., & Oshan, T. M. (2025). Scale and correlation in multiscale geographically weighted regression (MGWR). *Journal of Geographical Systems*, 27, 399–424. <https://doi.org/10.1007/s10109-025-00468-1>
- Kementerian PPN/Bappenas. (2014). *Pembangunan Kawasan Timur Indonesia Dalam Konteks Kekinian* (D. Dachlan & S. Suhab, Eds.). Puslitbang Kebijakan dan Manajemen Lembaga Penelitian dan Pengabdian Masyarakat Universitas Hasanuddin. https://perpustakaan.bappenas.go.id/e-library/file_upload/koleksi/migrasi-data-publikasi/file/Policy_Paper/Buku%20BPN-UH-FINAL.pdf
- Khedmati Morasae, E., Derbyshire, D. W., Amini, P., & Ebrahimi, T. (2024). Social determinants of spatial inequalities in COVID-19 outcomes across England: A multiscale geographically weighted regression analysis. *SSM - Population Health*, 25, 101621. <https://doi.org/10.1016/j.ssmph.2024.101621>
- Laome, L., Budiantara, I. N., & Ratnasari, V. (2022). Estimation Curve of Mixed Spline Truncated and Fourier Series Estimator for Geographically Weighted Nonparametric Regression. *Mathematics*, 11(1), 152. <https://doi.org/10.3390/math11010152>
- Liu, M., Ge, Y., Hu, S., & Hao, H. (2023). The Spatial Effects of Regional Poverty: Spatial Dependence, Spatial Heterogeneity and Scale Effects. *ISPRS International Journal of Geo-Information*, 12(12), 501. <https://doi.org/10.3390/ijgi12120501>
- Lumbantoruan, J. F. A. K., Ohlyver, M., & Moniaga, J. V. (2023). Developing a Poverty Model in Papua using Geographically Weighted Regression. *Procedia Computer Science*, 227, 243–252. <https://doi.org/10.1016/j.procs.2023.10.522>
- Mahmoud, H. F. F. (2021). Parametric Versus Semi and Nonparametric Regression Models. *International Journal of Statistics and Probability*, 10(2), 90–109. <https://doi.org/10.5539/ijsp.v10n2p90>
- Ola, P. K., Iriany, A., & Astutik, S. (2024). Mixed Geographically Weighted Regression (MGWR) With Adaptive Weighting Function in Poverty Modeling in NTT Province. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 18(3), 2035–2046. <https://doi.org/10.30598/barekengvol18iss3pp2035-2046>
- Omran, F., Shad, R., & Ziaee, S. A. (2025). A multiscale geographically weighted regression approach to emphasize the effects of traffic characteristics on vehicular emissions. *Atmospheric Environment: X*, 25, 100315. <https://doi.org/10.1016/j.aeaoa.2025.100315>
- Oshan, T., Li, Z., Kang, W., Wolf, L., & Fotheringham, A. (2019). mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS International Journal of Geo-Information*, 8(6), 269. <https://doi.org/10.3390/ijgi8060269>
- Putra, R., Fadhlurrahman, M. G., & Gunardi. (2023). Determination of the best knot and bandwidth in geographically weighted truncated spline nonparametric regression using generalized cross validation. *MethodsX*, 10, 101994. <https://doi.org/10.1016/j.mex.2022.101994>
- Saifudin, T., Chamidah, N., Ramadhina, F. S. I., Al Hasri, I. M., Trisa, N. L. H., Valida, H., & Setyawan, M. D. B. (2025). Mapping Food Insecurity: Spatial Modelling of Undernourishment Prevalence in Indonesia using Geographically Weighted Regression. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, 9(4), 1165–1181. <https://doi.org/10.31764/jtam.v9i4.32063>
- Sifriyani, Haryatmi, Budiantara, I. N., & Gunardi. (2017). Geographically Weighted Regression with Spline Approach. *Far East Journal of Mathematical Sciences (FJMS)*, 101(6), 1183–1196. <https://doi.org/10.17654/MS101061183>
- Sifriyani, Kartiko, S. H., Budiantara, I. N., & Gunardi. (2018). Development of nonparametric geographically weighted regression using truncated spline approach. *Songklanakarin J. Sci. Technol*, 40(4), 909–920. <https://doi.org/https://doi.org/10.14456/sjst-psu.2018.98>
- Suriaslan, A. S., Budiantara, I. N., & Ratnasari, V. (2025). Nonparametric regression estimation using multivariable truncated splines for binary response data. *MethodsX*, 14, 103084. <https://doi.org/10.1016/j.mex.2024.103084>

- Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., Kang, W., & Wolf, L. J. (2019). Inference in Multiscale Geographically Weighted Regression. *Geographical Analysis*, 52(1), 87–106. <https://doi.org/10.1111/gean.12189>