Jurnal Varian
Vol. 8, No. 3, October 2025, pp 293–306
e-ISSN: 2581-2017

Mixed Geographically Weighted Regression Modeling Using the MM-Estimator Method on Data of Poverty

Amalia Mentari Djalumang, Raupong Raupong, Siswanto Siswanto
Universitas Hasanuddin, Makassar, Indonesia

Article Info

Article history:

Received : 05-31-2025
Revised : 09-21-2025
Accepted : 10-05-2025

Keywords:

Method of Moment-Estimator;
Mixed Geographically Weighted Regression;
Outliers;
Percentage of Poor Population;
Tukey Bisquare.

ABSTRACT

The mixed geographically weighted regression model combines a global linear regression model with a geographically weighted regression model, with some parameters global and others local. When analyzing data with this model, outliers are common, which can significantly affect the regression coefficients and lead to biased parameter estimates. Therefore, a more robust estimation method that is resistant to outliers is needed to improve accuracy. This study aims to estimate the parameters of the mixed geographically weighted regression model using the Method of Moments (MM) Estimator method, which is more robust to outliers, and to identify the factors that significantly influence the percentage of the poor population in South Sulawesi Province in 2023. The results show that the poverty depth index has a significant global effect on the percentage of the population living in poverty. Meanwhile, the percentage of the population, the open unemployment rate, and the expected years of schooling have significant local effects. Based on these findings, it can be concluded that neighboring regions share common factors influencing poverty rates. These findings can assist policymakers in designing poverty-alleviation programs that account for regional differences and support further research on robust spatial modeling approaches.

Accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: https://doi.org/10.30812/varian.v8i3.5146

Corresponding Author:

Raupong Raupong,
Department of Statistics, Universitas Hasanuddin,
Email: raupong.stat.uh@gmail.com

How to Cite:
Djalumang, A. M., Raupong, R., Siswanto, S. (2025). Mixed Geographically Weighted Regression Modeling Using the MM-Estimator Method on Data of Poverty. *Jurnal Varian*, *8*(3), 293–306.

A. INTRODUCTION

Geographically Weighted Regression (GWR) is a regression model that accounts for spatial aspects, so that each observation location has different (local) model parameters. However, it is not uncommon for variables to have a global effect or to be constant across research locations. Therefore, the GWR model is still considered inappropriate to use and has been developed into Mixed Geographically Weighted Regression (MGWR) (Cholid, 2023). In the MGWR model, the GWR approach is combined with global linear regression. The parameter estimation results are partly global and partly local. MGWR model parameter estimators use WLS as in the GWR model (Zhang et al., 2019).

The drawback of the GWR and MGWR models is that they cannot handle outliers. Outliers are extreme values that exhibit unique characteristics and differ from other observations (Özdemir & Arslan, 2022). In general, outliers can occur for several reasons, including measurement, system, and other errors. Outliers are also observations that may significantly affect the regression coefficients, resulting in less precise estimates. Therefore, a robust method is required to estimate the MGWR model's parameters (Ibrahim et al., 2022).

Robust regression is a regression method used when the residuals are not normally distributed or some outliers affect the model (Syam et al., 2024). This method can yield results that are robust to the presence of outliers and is a valuable tool for analyzing data tainted by them (Azzahro & Sofro, 2023). One of the significant and popular robust regression techniques is the MM-estimator method. The MM-estimator method was initially introduced by combining the S-estimator and the M-estimator. Its high breakdown point and high efficiency under regular error distribution are two key advantages (Prahutama & Rusgiyono, 2021).

South Sulawesi Province is one of several Indonesian provinces with significant poverty rates. According to data from the Central Statistics Agency, the number of South Sulawesi residents living in poverty increased from 6.5 thousand in September 2022 to 788.85 thousand in March 2023. Compared with September 2022, the percentage of people living in poverty increased by 0.04 percentage points to 8.70 percent in March 2023. The description explains that this research will use the MM-estimator approach to fit a mixed spatially weighted regression model to data on the proportion of people living in poverty in South Sulawesi Province in 2023.

The main difference between this study and previous studies is that most previous studies on MGWR modeling still use estimation methods such as Weighted Least Squares (WLS), which are less robust to outliers, resulting in biased and inaccurate parameter estimates. The novelty of this study lies in applying the MM-Estimator method to MGWR modeling of poverty data, aiming to improve the model's resistance to outliers and to provide more efficient, spatially accurate parameter estimates. The purpose of this study is to estimate MGWR model parameters using the MM-Estimator, a more robust method against outliers, and to identify factors that significantly affect the percentage of poor people in South Sulawesi Province in 2023.

B. RESEARCH METHOD

1. Geographically Weighted Regression Model

The global linear regression model was further developed by incorporating geographically characteristics and spatial variability, leading to the formulation of the GWR model. This advanced model allows regression coefficients to vary across locations, enabling more effective capture of local relationships. Equation (1) displays the GWR model as follows (Comber et al., 2023):

$$y_i = \beta_0 (u_i, v_i) + \sum_{j=1}^{k} \beta_j (u_i, v_i) x_{ij} + \varepsilon_i \quad i = 1, 2, \ldots, n \tag{1}$$

In the GWR model, $\beta_0 (u_i, v_i)$ represents the intercept value, which is location-specific and may vary across observation points (u_i, v_i). Similarly, $\beta_j (u_i, v_i)$ denotes the j-th regression coefficient at the i-th observation location, reflecting the local influence of the j-th explanatory variable on the dependent variable. The error term, ε_i is assumed to be normally distributed with a mean of zero and a constant variance σ^2, capturing the difference between the observed and predicted values at each location. Equation (2) displays the GWR model's parameter estimate as follows:

$$\hat{\beta} (u_i, v_i) = \left[X^t W(u_i, v_i) X\right]^{-1} X^t W(u_i, v_i) Y \tag{2}$$

2. Mixed Geographically Weighted Regression Model

A global regression model and a local regression model are combined in the modeling technique known as MGWR. By keeping global factors with constant parameters but adding local variables whose effects change across space, MGWR provides a more flexible explanation for the spatial (Shen & Tao, 2022). Assuming that the model's intercept is local, the MGWR model with p and q predictor variables that are local may be expressed as follows in Equation (3):

$$y_i = \beta_0 (u_i, v_i) + \sum_{j=1}^{q} \beta_j (u_i, v_i) x_{ij} + \sum_{j=q+1}^{k} \beta_j x_{ij} + \varepsilon_i \quad i = 1, 2, \ldots, n \tag{3}$$

In the MGWR model, $\beta_0 (u_i, v_i)$ represents the intercept value that varies depending on the specific location of the observation. Then term $\beta_j (u_i, v_i)$ refers to the j-th regression coefficient at the i-th observation point, indicating that this coefficient

can changes across different spatial location. In contrast, β_j without location dependence denotes the j-th global regression coefficient, which remains constant and does not vary based on location.
The global model parameter estimates are shown in Equation (4) as follows (Yang et al., 2020):

$$\hat{\beta}_g = \left[X_g^t (I - S_l)^t (I - S_l) X_g\right]^{-1} X_g^t (I - S_l)^t (I - S_l) Y \tag{4}$$

And the local parameter estimates are shown in Equation (5) as follows:

$$\hat{\beta}_l (u_i, v_i) = \left[X_l^t W (u_i, v_i) X_l\right]^{-1} X_l^t W(u_i, v_i)(Y - X_g \hat{\beta}_g) \tag{5}$$

3. Robust Regression

Robust regression methods are applied when the error distribution deviates from normality or when several outliers impact the model results (Begashaw & Yohannes, 2020). Linear regression models require robust parameter estimation if the data used contains outliers (Kalina & Tichavský, 2020). One type of parameter estimation that can be used to estimate robust regression in this study is the Method of Moments (MM-estimator) estimation (Tirink & Önder, 2022). The MM-estimator procedure is to estimate the regression parameters using the S-estimator that minimizes the scale error of the M-estimator and then proceed to calculate the final parameter estimates with the M-estimator. The MM-estimator is defined in Equation (6) as follows (Farouk et al., 2023):

$$\hat{\beta}_{MM} = \min \sum_{i=1}^{n} \rho \left(\frac{e_i}{\hat{\sigma}}\right) = \min \sum_{i=1}^{n} \rho \left(\frac{y_i - \sum_{j=0}^{k} X_i \beta_j}{\hat{\sigma}}\right) \tag{6}$$

The weight function used in robust MM-estimator regression is often the Tukey bisquare weight.

4. Hypothesis Testing

Hypothesis testing in the MGWR model involves conducting two distinct types of tests. The first is simultaneous testing, which evaluates the significance of all global and local variables jointly. The second type is individual testing, often called partial one-on-one testing, which examines each parameter separately to determine its specific contribution within the MGWR framework.

a. Simultaneous Testing of MGWR Model Parameters

There are two types of hypothesis tests conducted in the MGWR model. The first is a simultaneous hypothesis test, which evaluates the overall significance of the global predictor variable parameters, as outlined below (Hermalia & Rini, 2023):

H_0 : $\beta_3 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$
H_1 : at least one $\beta_j \neq 0$

The statistical test is shown in Equation (7) as follows:

$$F_1 = \frac{Y^t[(I - S_l)^t (I - S_l) - (I - S)^t (I - S)]^{Y}/_{r_1}}{Y^t (I - S)^t (I - S)^{Y}/_{u_1}} \sim F_{(\alpha; df_1; df_2)} \tag{7}$$

Furthermore, the second simultaneous hypothesis test is carried out on the parameters of the local predictor variables with the following hypothesis:

H_0 : $\beta_1 (u_i, v_i) = \beta_2 (u_i, v_i) = \beta_4 (u_i, v_i) = \beta_9 (u_i, v_i) = 0$
H_1 : at least one $\beta_j (u_i, v_i) \neq 0$

The statistical test is shown in Equation (8) as follows:

$$F_2 = \frac{Y^t[(I - S_g)^t (I - S_g) - (I - S)^t (I - S)]^{Y}/_{c_1}}{Y^t (I - S)^t (I - S)^{Y}/_{u_1}} \sim F_{(\alpha; df_1; df_2)} \tag{8}$$

b. Individual Testing (Partial One-on-One) MGWR Model Parameters

There are two types of hypothesis tests carried out in the MGWR model. The first is a simultaneous hypothesis test, which examines the collective significance of the global predictor variable parameters using the following hypothesis (Hermalia & Rini, 2023):

H_0 : $\beta_j = 0, for j = 3, 5, 6, 7, 8$
H_1 : at least one $\beta_j \neq 0$

The statistical test is shown in Equation (9) as follows:

$$t_{g_{hitung}} = \frac{\hat{\beta}_j}{\hat{\sigma}\sqrt{g_{jj}}}\, t_{(\alpha/2, df)} \tag{9}$$

Furthermore, the second individual hypothesis test is conducted on the parameters of the local predictor variables with the following hypothesis:

H_0 : $\beta_j\left(u_i, v_i\right) = 0$, for $j = 1, 2, 4, 9$
H_1 : at least one $\beta_j\left(u_i, v_i\right) \neq 0$

The statistical test is shown in Equation (10) as follows:

$$t_{l_{hitung}} = \frac{\hat{\beta}_j(u_i, v_i)}{\hat{\sigma}\sqrt{m_{jj}}}\, t_{(\alpha/2, df)} \tag{10}$$

5. Data

This analysis used secondary data released by the South Sulawesi Province's Central Statistics Agency in 2023. Astronomical location data, including the latitude and longitude of each district and city in South Sulawesi Province, are also included in this study as a geographic weighting element. The response variable used is the percentage of the poor population, while the predictor variables include the percentage of the population (X_1), the open unemployment rate (X_2), the percentage of the poor population aged 15 and over who did not complete elementary school (X_3), the expected years of schooling (X_4), the percentage of households using non-PLN electricity sources (X_5), the percentage of monthly per capita food expenditure (X_6), the poverty depth index (X_7), the underemployment rate (X_8), and the percentage of families that have access to sources of clean drinking water (X_9). The flowchart analysis and specific steps of the data processing procedure used in this study are shown in Figure 1.

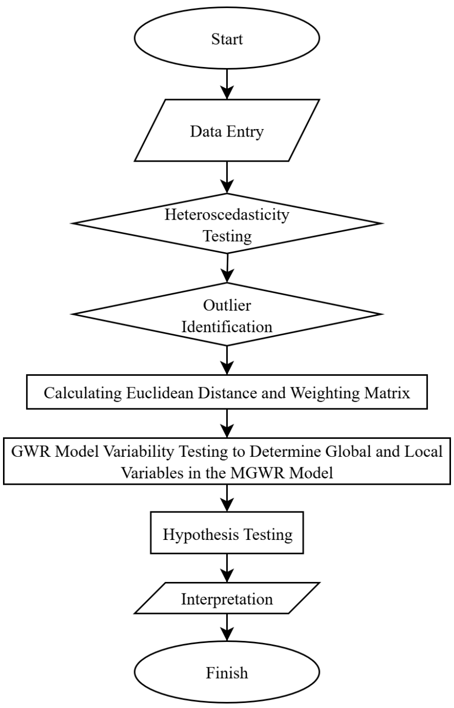

Figure 1. Research Flowchart

DOI: https://doi.org/10.30812/varian.v8i3.5146

The research implementation flow in the data processing section to address the outlier issue that arises in the data using the MM-estimator is depicted in Figure 1. A more thorough description of the flow chart above may be found below:

1) Preliminary analysis:
   a. Preliminary analysis:
   The Breusch-Pagan (BP) test is used to assess spatial heterogeneity, with the statistics calculated using Equation (11). If the BP statistic value exceeds $\chi^2_{(\alpha,\ k)}$ or the p-value is less than α, where k is the number of predictor variables, this test will reject H_0 (Faradiba & Dhuhri, 2024).

$$BP = \left(\frac{1}{2}\right) f^t Z \left(Z^t Z\right)^{-1} Z^t f \sim \chi^2_{(\alpha,k)} \tag{11}$$

   b. Identify outliers by looking at the leverage using Equation (12) (Kannan & Manoj, 2015),

$$h_{ii} = x_i' \left(X'X\right)^{-1} x_i \tag{12}$$

   and the DFFITS value using Equation (13) (Kannan & Manoj, 2015).

$$DFFITS = t_i \left(\frac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}} \tag{13}$$

2) Calculate the Euclidean distance (d_{ij}) between the i-th location and the j-th location located at coordinates (u_i, v_i) and (u_j, v_j) with Equation (14) (Edayu & Syerrina, 2018).

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{14}$$

3) Using Equation (15) and the Cross-Validation technique, the optimum bandwidth value is obtained by minimizing the Cross-Validation score (Rahmah et al., 2020).

$$CV\left(h\right) = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq i}\left(h\right)\right)^2 \tag{15}$$

4) Determine the weighted matrix to find the closeness between areas. Equation (16) uses the fixed tricube kernel function as the weight matrix in this study (Edayu & Syerrina, 2018).

$$w_j\left(u_i, v_i\right) = \begin{cases} \left[1 - \left(\frac{d_{ij}}{h}\right)^3\right], & \text{for } d_{ij} < h \\ 0, & \text{more} \end{cases} \tag{16}$$

5) To identify global and local variables in the MGWR model using Equation (17), do a GWR model variability test.

$$F = \frac{V_j^2 / tr\left(\frac{1}{k} B_j^t \left[I - \frac{1}{k} J\right] B_j\right)}{Y^t \left(I-S\right)^t \left(I-S\right) Y / b_1} \sim F_{(\alpha; df_1; df_2)} \tag{17}$$

6) Use the MM-estimator method to find the MGWR model's parameter estimation. Here is a description of the MGWR process that use the MM-estimator method:
   a. Calculating the error value $e_i = y_i - \hat{y}$ from OLS.
   b. Calculating the value of $\hat{\sigma}_S = \sqrt{\frac{n\sum_{i=0}^{n}\left(e_i^2\right) - \left(\sum_{i=0}^{n} e_i\right)^2}{n(n-1)}}$
   c. Calculating the value of $u_i = \frac{e_i}{\hat{\sigma}_S}$
   d. Calculate the weights for the S-estimator as follows Equation (18):

$$w_i = \begin{cases} \left[1 - \left(\frac{u_i}{1.547}\right)^2\right]^2 & \text{for } |u_i| \le 1.547 \\ 0 & \text{for } |u_i| > 1.547 \end{cases} \tag{18}$$

   e. Estimate the coefficients $\beta_g^{(1)}$ and $\beta_l\left(u_i, v_i\right)^{(1)}$ by WLS method using weight ω_i so as to obtain $e_i = y_i - \hat{y}$.

DOI: https://doi.org/10.30812/varian.v8i3.5146

f. The error $e_i^{(1)}$ in the first step is used to calculate the scale error of the M-estimator,

$$\hat{\sigma}_M = \frac{median\,|e_i - median\,(e_i)|}{0.6745} \tag{19}$$

g. Calculate the value of $u_i = \frac{e_i}{\hat{\sigma}_S}$ then calculate the initial weight $\omega_i^{(1)}$.

$$\omega_i^{(1)} = \begin{cases} \left[1 - \left(\frac{u_i}{4.685}\right)^2\right]^2 & \text{for } |u_i| \leq 4.685 \\ 0 & \text{for } |u_i| > 4.685 \end{cases} \tag{20}$$

h. The first iteration of the WLS technique uses the scale estimate $\hat{\sigma}_M$ from step f and the error $e_i^{(1)}$ to compute the regression coefficient $\hat{\beta}_{MM}$ using $\omega_i^{(1)}$.
i. Repeating steps f, g, and h until a convergent estimator is obtained.
j. Using the MM-estimator to calculate global and local parameters in the MGWR model.
k. Conducting hypothesis testing of the MGWR model, namely simultaneous tests and individual tests (partial one-on-one).
l. Making interpretations and conclusions.

C. RESULT AND DISCUSSION

1. Spatial Heterogeneity Testing

To assess the variation across locations, spatial heterogeneity testing was conducted. The H_0 decision is denied because the Breusch-Pagan statistic value of 19.306 is higher than the $\chi^2_{(0.05,9)}$ value of 16.92. This indicates that the data on the proportion of the population living in poverty in South Sulawesi Province in 2023 exhibits geographical heterogeneity. A local method is required to develop the model due to data variability, indicating that the districts and cities in South Sulawesi Province have distinct features. The GWR model is one of the models that may be applied.

2. Outlier Identification

Outliers in a given observation can be effectively identified using specific statistical criteria, namely leverage values and the DFFITS method. These techniques help to assess the influence and impact of individual data points on the overall model. The criteria and procedures for detecting outliers based on leverage and DFFITS are described as follows.

a. Leverage Value

Leverage values that exceed the value of the cut-off point, namely $(2k-1)/n = 17/24 = 0.7083$, are identified as outliers when the number of predictor variables is k and the quantity of observational data is n.

Table 1. Leverage Value

Data to-	Leverage Value	Description
1	0.9003	Outliers
22	0.8477	

Based on the leverage value in Table 1, the 1st data and 22nd data have values more than 0.7083, indicating that they are outlier data that have an impact on the regression model.

b. DFFITS Method

This method is used to determine which outlier data points affect the regression model. Data is said to be an outlier if the |DFFITS| $> 2\sqrt{\frac{k}{n}}$. In this study, the number of predictor variables is nine and consists of 24 observations, so the value of $2\sqrt{\frac{9}{24}} = 1.2247$ is obtained.

Table 2. DFFITS Value

Data to-	|DFFITS|	Description
1	-6.5185	Outliers
9	1.6492	
11	2.4169	

DOI: https://doi.org/10.30812/varian.v8i3.5146

Data to-	|DFFITS|	Description
18	1.6014	
20	-1.4250	
22	-3.2217	

According to Table 2, the DFFITS values for the 1st, 9th, 11th, 18th, 20th, and 22nd data points are greater than 1.2247, indicating that these data points are outliers that affect the regression model. Because outliers are detected in the data, the MM-estimator method is used to address them.

3. Geographically Weighted Regression

The geographic latitude and longitude of every district and city in South Sulawesi Province were ascertained by GWR modelling. The next stage is to use the Cross-Validation (CV) approach to determine the optimal bandwidth.

Table 3. Cross Validation (CV) Value Bandwidth

Weighting Function	Bandwidth	CV Value
Adaptive Gaussian Kernel	0.3648	176.9096
Adaptive Bi-Square Kernel	0.9999	161.0458
Adaptive Tricube Kernel	362.1721	159.7131
Fixed Gaussian Kernel	138.0854	172.1874
Fixed Bi-Square Kernel	0.9999	160.7669
Fixed Tricube Kernel	**378.6171**	**159.4366**

Based on Table 3, the weight matrix at the i-th observation point, $W(u_i, v_i)$ is then determined using the fixed tricube kernel weight function after the optimal bandwidth value is obtained. The first step to get the weight matrix is to calculate the Euclidean distance (d_{ij}) at each observation location. The weight matrix at the location (u_i, v_i) is a diagonal matrix $W(u_i, v_i)$, so 24 weight matrices are obtained for the poverty percentage data in South Sulawesi Province. The weight matrix is calculated using Equation (15) as shown in Table 4.

Table 4. Weighting Matrix

Region	Kep. Selayar	Bulukumba	Bantaeng	...	Palopo
Kep.Selayar	1	0.9815	0.9675	...	0.0115
Bulukumba	0.9815	1	0.9991	...	0.1937
Bantaeng	0.9675	1.597.131	1	...	0.1907
⋮	⋮	⋮	⋮	⋱	⋮
Palopo	0.0115	0.1937	0.1907	...	1

4. GWR Variability Test

The variability test is conducted to determine global variables and local variables. Based on Table 5, five variables are globally influential, namely X_3, X_5, X_6, X_7, and X_8. While four variables that have a local effect are X_1, X_2, X_4, and X_9.

Table 5. GWR Variability Test

Variable	$p-value$	Test Decision
X_1	0.0361	Reject H_0
X_2	0.0331	Reject H_0
X_3	0.9998	Failure to Reject H_0
X_4	0.0004	Reject H_0
X_5	0.0681	Failure to Reject H_0
X_6	0.9059	Failure to Reject H_0
X_7	0.7954	Failure to Reject H_0
X_8	0.2908	Failure to Reject H_0
X_9	0.0477	Reject H_0

5. MGWR Model Parameter Estimation using MM-estimator

Equation (21) below is a description of the MGWR model that is thought to have outliers:

$$y_i = \beta_0 (u_i, v_i) + \sum_{j=1}^{q} \beta_j (u_i, v_i) x_{ij} + \sum_{j=q+1}^{k} \beta_j x_{ij} + \varepsilon_i \tag{21}$$

or Equation (22)

$$\varepsilon_i = y_i - \beta_0 (u_i, v_i) - \sum_{j=1}^{q} \beta_j (u_i, v_i) x_{ij} - \sum_{j=q+1}^{k} \beta_j x_{ij} \tag{22}$$

a. MGWR Model Global Parameter Estimation Process

Assuming that the i-th row of the matrix X_l is structured in the form $x_{li}^t = (1, x_{1i}, x_{2i}, \ldots, x_{qi})$, the predicted value $\widetilde{y}$ for the i-th observation can be computed accordingly based on this specific arrangement of the explanatory variables. This formulation allows the incorporation of an intercept term along with multiple predictors in the prediction process.

$\widehat{\widetilde{y}}_i = x_{li}^t \hat{\beta}_l(u_i, v_i)$

$\widehat{\widetilde{y}}_i = x_{li}^t((X_l^t W(u_i, v_i) \omega X_l)^{-1} X_l^t W(u_i,\ v_i) \omega \widetilde{Y})$

So that for all observations can be written:

$\widehat{\widetilde{Y}} = \left(\widehat{\widetilde{y}}_1, \widehat{\widetilde{y}}_2, \ldots, \widehat{\widetilde{y}}_n\right) = S_l \widetilde{Y}$

Next, substitute the elements of $\hat{\beta}_l(u_i, v_i)$ into the MGWR model as follows:

$Y = X_l \beta_l (u_i, v_i) + X_g \beta_g + \varepsilon$
$Y = S_l \widetilde{Y} + X_g \beta_g + \varepsilon$
$Y = S_l (Y - X_g \beta_g) + X_g \beta_g + \varepsilon$
$\varepsilon = (I - S_l) Y - (I - S_l) X_g \beta_g$

Furthermore, to find the global parameters of the MGWR model, the error value obtained is substituted in Equation (6), so that it is obtained:

$$\sum_{i-1}^{n} \rho \left(\frac{(I - S_{li}) y_i - (I - S_{li})(\sum_{j=q+1}^{k} \beta_j x_{ij})}{\hat{\sigma}} \right) \tag{23}$$

To minimize the residual function $\rho(u_i)$ in Equation (23), the first partial derivative of $\rho(u_i)$ against β_j is equalized to zero, resulting in the following equation:

$$\sum_{i=1}^{n} x_{ij} (1 - S_{li}) \psi \left(\frac{(1 - S_{li}) y_i - (1 - S_{li}) + \sum_{j=q+1}^{k} \beta_j x_{ij}}{\hat{\sigma}} \right) = 0 \tag{24}$$

Provide a solution by defining a weighting function ω_i is shown in Equation (25) as follows:

$$\omega_i = \omega (u_i) = \frac{\psi(u_i)}{u_i} \tag{25}$$

with

$$u_i = \frac{\varepsilon_i}{\hat{\sigma}} = \left(\frac{(1 - S_{li}) y_i - (1 - S_{li})(\sum_{j=q+1}^{k} \beta_j x_{ij})}{\hat{\sigma}} \right) \tag{26}$$

Based on Equation (25), Equation (24) can be changed into:

$$\omega_i = \frac{\psi \left(\frac{(1-S_{li})y_i-(1-S_{li})(\sum_{j=q+1}^{k} \beta_j x_{ij})}{\hat{\sigma}} \right)}{\left(\frac{(1-S_{li})y_i-(1-S_{li})(\sum_{j=q+1}^{k} \beta_j x_{ij})}{\hat{\sigma}} \right)} \tag{27}$$

Then the two segments in Equation (27) are multiplied by $\frac{(1-S_{li})y_i-(1-S_{li})(\sum_{j=q+1}^{k}\beta_j x_{ij})}{\hat{\sigma}}$ to obtain:

$$\omega_i \frac{(1-S_{li})\,y_i-(1-S_{li})(\sum_{j=q+1}^{k}\beta_j x_{ij})}{\hat{\sigma}} = \psi\left(\frac{(1-S_{li})\,y_i-(1-S_{li})(\sum_{j=q+1}^{k}\beta_j x_{ij})}{\hat{\sigma}}\right)=0 \tag{28}$$

Then Equation (28) is substituted into Equation (24) to obtain:

$$\sum_{i=1}^{n} x_{ij}\,(1-S_{li})\,\omega_i\left(\frac{(1-S_{li})\,y_i-(1-S_{li})\left(\sum_{j=q+1}^{k}\beta_j x_{ij}\right)}{\hat{\sigma}}\right)=0 \tag{29}$$

If Equation (29) is denoted into matrix form, Equation (30) is obtained:

$$\hat{\beta}_g=\left(X_g^t\,(I-S_l)^t\,\omega\,(I-S_l)\,X_g\right)^{-1}X_g^t\,(I-S_l)^t\,\omega(I-S_l)Y \tag{30}$$

Equation (31) displays the initial estimate of $\hat{\beta}_g^{(0)}$ derived from the estimation of the MGWR model using OLS based on Equation (4). The weight $\omega_i^{(0)}$ will be calculated using the estimate of $\hat{\beta}_g^{(0)}$.

$$\omega_i^{(0)}=\frac{\psi\left(\frac{(1-S_{li})y_i-(1-S_{li})(\sum_{j=q+1}^{k}\beta_j x_{ij})}{\hat{\sigma}}\right)}{\left(\frac{(1-S_{li})y_i-(1-S_{li})\left(\sum_{j=q+1}^{k}\beta_j x_{ij}\right)}{\hat{\sigma}}\right)} \tag{31}$$

So that the value of $\omega_i^{(0)}$ can be used to determine $\hat{\beta}_g^{(1)}$ is shown in Equation (32) as follows:

$$\hat{\beta}_g^{(1)}=\left(X_g^t\omega^{(0)}\,(I-S_l)\,X_g\right)^{-1}X_g^t\,(I-S_l)^t\,\omega^{(0)}\,(I-S_l)\,Y \tag{32}$$

The calculation of $\omega_i^{(1)}$ is done by utilizing the estimate of $\hat{\beta}_g^{(1)}$, and then updated iteratively in each iteration. The final result of this process is the global parameter estimates of the MGWR model with the MM-estimator, as listed in Equation (33).

$$\hat{\beta}_{gMM}^{(m)}=\left(X_g^t\,(I-S_l)^t\,\omega^{(m-1)}\,(I-S_l)\,X_g\right)^{-1}X_g^t\,(I-S_l)^t\,\omega^{(m-1)}\,(I-S_l)\,Y \tag{33}$$

b. MGWR Model Local Parameter Estimation Process

To estimate the local parameters of the MGWR model, Equation (22) is transformed into Equation (34) as shown below. This conversion facilitates the calculation of location-specific regression coefficients within the model framework:

$$\varepsilon_i=\widetilde{y}_i-(\beta_0\,(u_i,v_i)+\sum_{j=1}^{q}\beta_j\,(u_i,v_i)\,x_{ij} \tag{34}$$

Estimation of the local parameters of the MGWR model with the MM-estimator is done by minimizing the inclusion function $\rho(u_i)$ by adding a spatial weighting function $W(u_i,v_i)$ at each observation location so that Equation (35) is obtained as follows:

$$MM=\sum_{i=1}^{n}W\,(u_i,v_i)\,\rho\left(\frac{\widetilde{y}_i-(\beta_0\,(u_i,v_i)+\sum_{j=1}^{q}\beta_j\,(u_i,v_i)\,x_{ij}}{\hat{\sigma}}\right) \tag{35}$$

To minimize the inclusion function $\rho(u_i)$ against $\beta_j\,(u_i,v_i)$ is equalized to zero, so the following equation is obtained is shown in Equation (36):

$$\sum_{i=1}^{n}x_{ij}W\,(u_i,v_i)\,\psi\left(\frac{\widetilde{y}_i-(\beta_0\,(u_i,v_i)+\sum_{j=1}^{q}\beta_j\,(u_i,v_i)\,x_{ij}}{\hat{\sigma}}\right)=0 \tag{36}$$

Provide a solution by defining a weighting function ω_i is shown in Equation (37) as follows:

$$\omega_i=\omega\,(u_i)=\frac{\psi(u_i)}{u_i} \tag{37}$$

with

$$u_i = \frac{\varepsilon_i}{\hat{\sigma}} = \frac{\widetilde{y}_i - \left(\beta_0 \left(u_i, v_i\right) + \sum_{j=1}^{q} \beta_j \left(u_i, v_i\right) x_{ij}\right.}{\hat{\sigma}} \tag{38}$$

Based on Equation (38), Equation (37) can be changed into:

$$\omega_i = \frac{\psi\left(\frac{\widetilde{y}_i - (\beta_0(u_i,v_i) + \sum_{j=1}^{q} \beta_j(u_i,v_i)x_{ij}}{\hat{\sigma}}\right)}{\frac{\widetilde{y}_i - (\beta_0(u_i,v_i) + \sum_{j=1}^{q} \beta_j(u_i,v_i)x_{ij}}{\hat{\sigma}}} \tag{39}$$

Equation (40) is then obtained by multiplying the two segments in Equation (39) by $\frac{\widetilde{y}_i - (\beta_0(u_i,v_i) + \sum_{j=1}^{q} \beta_j(u_i,v_i)x_{ij}}{\hat{\sigma}}$.

$$\omega_i \left(\frac{\widetilde{y}_i - \left(\beta_0 \left(u_i, v_i\right) + \sum_{j=1}^{q} \beta_j \left(u_i, v_i\right) x_{ij}\right.}{\hat{\sigma}}\right) = \psi \frac{\widetilde{y}_i - \left(\beta_0 \left(u_i, v_i\right) + \sum_{j=1}^{q} \beta_j \left(u_i, v_i\right) x_{ij}\right.}{\hat{\sigma}} \tag{40}$$

Then Equation (40) is substituted into Equation (36) to obtain:

$$\sum_{i=1}^{n} x_{ij} W\left(u_i, v_i\right) \omega_i \left(\frac{\widetilde{y}_i - \left(\beta_0 \left(u_i, v_i\right) + \sum_{j=1}^{q} \beta_j \left(u_i, v_i\right) x_{ij}\right.}{\hat{\sigma}}\right) = 0 \tag{41}$$

If Equation (41) is denoted into matrix form then, Equation (42) is obtained:

$$\hat{\beta}_l \left(u_i, v_i\right) = \left(X_l^t W \left(u_i, v_i\right) \omega X_l\right)^{-1} X_l^t W(u_i, v_i)\omega \widetilde{Y} \tag{42}$$

Equation (43) displays the local parameter estimations of the MGWR model in the following manner:

$$\hat{\beta}_l \left(u_i, v_i\right) = (X_l^t W \left(u_i, v_i\right) \omega X_l)^{-1} X_l^t W \left(u_i, v_i\right) \omega (Y - X_g \hat{\beta}_g) \tag{43}$$

Equation (44), which uses the estimate of $\hat{\beta}_l \left(u_i, v_i\right)^{(0)}$ to determine the weight $\omega_i^{(0)}$, displays the initial estimate of $\hat{\beta}_l \left(u_i, v_i\right)^{(0)}$ based on the estimation of the MGWR model using WLS based on Equation (5).

$$\omega_i^{(0)} = \frac{\psi\left(\frac{\widetilde{y}_i - (\beta_0^{(0)}(u_i,v_i) + \sum_{j=1}^{q} \beta_j^{(0)}(u_i,v_i)x_{ij}}{\hat{\sigma}}\right)}{\frac{\widetilde{y}_i - (\beta_0^{(0)}(u_i,v_i) + \sum_{j=1}^{q} \beta_j^{(0)}(u_i,v_i)x_{ij}}{\hat{\sigma}}} \tag{44}$$

Using $\omega^{(0)}$, Equation (45) presents the calculation steps to obtain $\hat{\beta}_l \left(u_i, v_i\right)^{(1)}$.

$$\omega_i^{(0)} = \frac{\psi\left(\frac{\widetilde{y}_i - (\beta_0^{(0)}(u_i,v_i) + \sum_{j=1}^{q} \beta_j^{(0)}(u_i,v_i)x_{ij}}{\hat{\sigma}}\right)}{\frac{\widetilde{y}_i - (\beta_0^{(0)}(u_i,v_i) + \sum_{j=1}^{q} \beta_j^{(0)}(u_i,v_i)x_{ij}}{\hat{\sigma}}} \tag{45}$$

Equation (46) displays the local parameter estimates of the MGWR model with the MM-estimator at the i-th observation location. The weight value $\omega_i^{(1)}$ is calculated using the value $\hat{\beta}_l \left(u_i, v_i\right)^{(1)}$, and so on. The value of ω will change with each iteration process.

$$\hat{\beta}_{lMM} \left(u_i, v_i\right)^{(m)} = \left(X_l^t W \left(u_i, v_i\right) \omega^{(m-1)} X_l\right)^{-1} X_l^t W \left(u_i, v_i\right) \omega^{(m-1)} (Y - X_g \hat{\beta}_g) \tag{46}$$

6. MGWR Modeling with MM-estimator

The calculation results for global parameter estimation $\hat{\beta}_g$ obtained $X_3 = -0.0794$, $X_5 = -0.1173$, $X_6 = 0.0052$, $X_7 = 0.1213$, $X_8 = 3.1994$. As for local parameter estimation, the following in Table 6 is an example for the Selayar Islands Region:

Table 6. Parameter Estimation of Selayar Islands Local Variables

Variable	$\hat{\beta}_l(u_1, v_1)$
Intercept	0.2606
X_1	0.0579

Variable	$\hat{\beta}_l(u_1, v_1)$
X_2	0.1604
X_4	0.1058
X_5	0.0998

Based on the parameter estimation results obtained, the MGWR model using the MM-estimator method on the percentage of poor population data is different for each observation location.

7. MGWR Model Hypothesis Testing

a. Concurrent Testing of MGWR Model Parameters

Simultaneous assessment of the global and local variable settings makes up this test. With a significance level of $\alpha = 5\%$, the test statistic value $F_{count} = 5.5985 > F_{table} = 2.1371$ was obtained from the simultaneous testing of the MGWR model parameters on global variables. The test decision then rejects H_0, indicating that the global variables have a significant impact on the response variable at the same time. At a significance level of $\alpha = 5\%$, the test statistic value $F_{count} = 4.0041 > F_{table} = 2.1667$ indicates that the MGWR model parameters were tested simultaneously on local variables. The test decision rejects H_0, indicating that local variables have a significant impact on the response variable at the same time.

b. Individual (Partial One-on-One) Testing of MGWR Model Parameters

In the MGWR model, this test identifies the local and global factors that significantly affect the response variable. The poverty depth index (X_7), which is shown in Table 7, is the global predictor variable that significantly affects.

Table 7. Individual Test of Global Parameters

Variabel	t_{g_count}	t_{table}	Description
X_3	0.4148		Not Significant
X_5	0.7389		Not Significant
X_6	0.6163	2.4231	Not Significant
X_7	5.4093		Significant
X_8	0.0477		Not Significant

Next, testing the significance of local variable parameters. The following is an example of testing parameters individually (partial one-on-one) in the Selayar Islands:

Table 8. Individual Test of Global Parameters

Region	Variable	t_{l_count}	t_{table}	Description
Selayar Islands	X_1	2.6698	2.4231	Significant
	X_2	1.2769		Not Significant
	X_4	10.6474		Significant
	X_9	0.0371		Not Significant

According to Table 8, the population percentage (X_1) and projected years of education (X_4) are the local factors that significantly impact the percentage of the Selayar Islands population living in poverty. Each observation location produces MGWR models that vary, influenced by the coefficient values of global and local variables that are significant to the response variable. The grouping of districts based on significant global and local variables in the MGWR model is shown in Table 9.

Table 9. District/City Grouping Based on Significant Variables

District/City	Significant Variable	
	Global	Local
Selayar Islands	X_7	X_1, X_4
Bantaeng, Barru, Bone, Bulukumba, Gowa, Jeneponto, Maros, Makassar, Pangkep, Pare-pare, Sinjai, Soppeng, Takalar	X_7	X_4
Enrekang, Pinrang, Sidrap, Wajo	X_7	–
Luwu, Palopo, Tana Toraja, Toraja Utara	X_7	X_2
Luwu Utara, Luwu Timur	X_7	X_1, X_2, X_4

The variable X_7 had a large global impact in every region, according to the findings of the MGWR study using the MM-estimator on the data of the percentage of the population living in poverty in South Sulawesi Province in 2023. Nonetheless, each area or city has a different level of local variable effect. X_1 and X_4 are the variables that have a local influence in the Selayar Islands. Additionally, X_4 is a significant local variable in the majority of the districts/cities in the southern region, the west coast, and several central regions, including the districts of Bulukumba, Bantaeng, Jeneponto, Takalar, Gowa, Sinjai, Maros, Pangkep, Barru, Bone, Soppeng, Makassar city, and Pare-pare city. Additionally, there are no significant local variables in the districts of Wajo, Sidrap, Pinrang, and Enrekang, suggesting that global factors are more important in explaining poverty rate variation in these areas. Additionally, X_2 is a local variable that has a major impact on the districts of Luwu, Tana Toraja, Toraja, and Palopo city. Additionally, the variables X_1, X_2, and X_4 are the local factors that significantly affect the districts of Luwu Utara and Luwu Timur, demonstrating the complexity of the causes of poverty in these areas. Figure 2 displays the mapping of the findings.

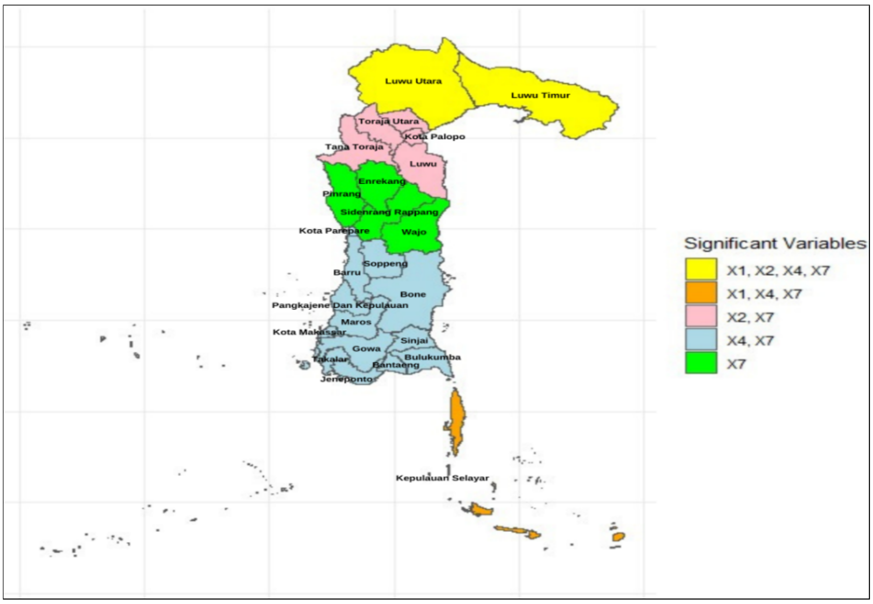

Figure 2. District/City Grouping Based on Significant Variables

The analysis results show that Mixed Geographically Weighted Regression (MGWR) modeling using the MM-Estimator method produces more stable and accurate parameter estimates than conventional estimation methods. This model has proven effective in handling outliers in poverty data in South Sulawesi Province and in accurately describing spatial variations across regions. This differs from the research conducted by Almetwally & Almongy (2018), which compared the S-estimator, M-estimator, and MM-estimator methods, where the results showed that the MM-estimator method was more efficient in dealing with outlier data, but has not been applied to spatial models. Similarly, research by Singgih & Fauzan (2022) compared the M-estimator, S-estimator, and MM-estimator methods, with results showing that the S-estimator method is more efficient in handling data with outliers but has not been applied to spatial models, as well as research by Shabrina et al. (2021) comparing the Geographically Weighted Regression (GWR) and Mixed Geographically Weighted Regression (MGWR) models with a focus on the distribution of Airbnb and urban tourism elements, showing that MGWR is the best model because it can provide smoother, more realistic, and informative results. Still, the study has not addressed the issue of outliers that can affect model parameter estimation.

This study combines the strengths of both approaches, namely the application of the robust MM-Estimator method in MGWR modeling, resulting in a model that not only captures spatial differences across regions but is also resistant to the influence of outliers. Thus, the MGWR with the MM-Estimator model developed in this study can yield more efficient and representative estimates in the context of spatial poverty analysis. A limitation of our study is that we have not explored the effect of variations in the weighting function (kernel function) on the sensitivity of the estimation results. In addition, this study has not focused on comparing the performance of various other robust methods, such as the Huber estimator, which may also be relevant for spatial modeling. In this study, we used 2023 poverty data from South Sulawesi Province, which included nine explanatory variables to capture spatial conditions and differences across districts/cities. We used global and local significance tests to assess model consistency, as has been done in previous spatial studies.

D. CONCLUSION AND SUGGESTION

According to the results of MGWR modeling using the MM-estimator method on data on the percentage of the population living in poverty in South Sulawesi Province in 2023, the poverty depth index for 2023 is the factor with a globally significant effect. Additionally, the percentage of the population, the open unemployment rate, and the anticipated years of education all have a substantial local impact on the proportion of the people living in poverty in South Sulawesi Province in 2023. To develop the model and identify the variables that significantly impact the proportion of people living in poverty in South Sulawesi Province, future studies should include more predictor variables.

ACKNOWLEDGEMENT

The authors would like to express their deepest gratitude to all those who have contributed to the completion of this research.

DECLARATIONS

AUTHOR CONTIBUTION

All authors contributed to the writing of this article.

FUNDING STATEMENT

This study was conducted without any financial support.

COMPETING INTEREST

The author declares that there is no conflict of interest in publishing this article.

REFERENCES

Almetwally, E., & Almongy, H. (2018). Comparison Between M-estimation, S-estimation, And MM Estimation Methods of Robust Estimation with Application and Simulation. *International Journal of Mathematical Archive*, *9*(11), 55–63.

Azzahro, I. A., & Sofro, A. (2023). Regresi Robust untuk Pemodelan Deforestasi di Indonesia. *MATHunesa: Jurnal Ilmiah Matematika*, *11*(3), 496–507. https://doi.org/10.26740/mathunesa.v11n3.p496-507

Begashaw, G. B., & Yohannes, Y. B. (2020). Review of Outlier Detection and Identifying Using Robust Regression Model. *International Journal of Systems Science and Applied Mathematics*, *5*(1), 4. https://doi.org/10.11648/j.ijssam.20200501.12

Cholid, F. (2023). Perbandingan Geographically Weighted Regression dengan Mixed Geographically Weighted Regression: Studi Kasus Prevalensi Stunting di Indonesia. *Statistika*, *23*(2), 96–109. https://doi.org/10.29313/statistika.v23i2.1700

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*(1), 155–178. https://doi.org/10.1111/gean.12316

Edayu, Z. N., & Syerrina, Z. (2018). A statistical analysis for geographical weighted regression. *IOP Conference Series: Earth and Environmental Science*, *169*, 012105. https://doi.org/10.1088/1755-1315/169/1/012105

Faradiba, H. P., & Dhuhri, A. S. (2024). Using the Geographically Weighted Regression (GWR) Method to Find Out the Factors Causing Poverty in North Sumatra Province. *International Journal of Science and Mathematics Education*, *1*(1), 18–28. https://doi.org/10.62951/ijsme.v1i1.13

Farouk, A. U., Lasisi, K. E., Mohammed, A., & Lawan, A. M. (2023). The Impact of Using Robust MM-Estimator in Regression Models: An Application of Nigerian Inflation on the Economic Growth. *Journal of Science and Applied Research*, *5*(2), 66–74.

Hermalia, H., & Rini, D. S. (2023). Pemodelan Persentase Balita Stunting di Indonesia Tahun 2021 dengan Metode Mixed Geographically Weighted Regression (MGWR). *Journal of Scientech Research and Development*, *5*(2), 780–790. https://doi.org/10.56670/jsrd.v5i2.249

Ibrahim, A., Dike, I. J., & Badawaire, A. B. (2022). Comparative Study of Some Estimators of Linear Regression Models in the Presence of Outliers. *FUDMA Journal of Sciences*, *6*(1), 368–376. https://doi.org/10.33003/fjs-2022-0601-908

Kalina, J., & Tichavský, J. (2020). On Robust Estimation of Error Variance in (Highly) Robust Regression. *Measurement Science Review*, *20*(1), 6–14. https://doi.org/10.2478/msr-2020-0002

Kannan, K. S., & Manoj, K. (2015). Outlier detection in multivariate data. *Applied Mathematical Sciences*, *9*(47), 2317–2324. https://doi.org/10.12988/ams.2015.53213

Özdemir, Ş., & Arslan, O. (2022). Combining empirical likelihood and robust estimation methods for linear regression models. *Communications in Statistics - Simulation and Computation*, *51*(3), 941–954. https://doi.org/10.1080/03610918.2019.1659968

Prahutama, A., & Rusgiyono, A. (2021). Robust regression with MM-estimator for modelling the number maternal mortality of pregnancy in Central Java, Indonesia. *Journal of Physics: Conference Series*, *1943*(1), 012148. https://doi.org/10.1088/1742-6596/1943/1/012148

Rahmah, A. A., Sulvianti, I. D., Suhaeni, C., & Djaafara, B. A. (2020). Pemodelan Faktor Risiko Penyakit Campak pada Balita di Provinsi DKI Jakarta: Pemodelan Faktor Risiko Penyakit Campak pada Balita di Provinsi DKI Jakarta. *Xplore: Journal of Statistics*, *9*(1), 1–11. https://doi.org/10.29244/xplore.v9i1.158

Shabrina, Z., Buyuklieva, B., & Ng, M. K. M. (2021). Short-Term Rental Platform in the Urban Tourism Context: A Geographically Weighted Regression (GWR) and a Multiscale GWR (MGWR) Approaches. *Geographical Analysis*, *53*(4), 686–707. https://doi.org/10.1111/gean.12259

Shen, Y., & Tao, Y. (2022). Associations between spatial access to medical facilities and health-seeking behaviors: A mixed geographically weighted regression analysis in Shanghai, China. *Applied Geography*, *139*, 102644. https://doi.org/10.1016/j.apgeog.2022.102644

Singgih, M. N. A., & Fauzan, A. (2022). Comparison of M Estimation, S Estimation, with MM Estimation to Get the Best Estimation of Robust Regression in Criminal Cases in Indonesia. *Jurnal Matematika, Statistika dan Komputasi*, *18*(2), 251–260. https://doi.org/10.20956/j.v18i2.18630

Syam, U. A., Siswanto, S., & Sunusi, N. (2024). Robust spatial Durbin modelling on tuberculosis data using the MM-estimator method. *Statistics in Transition new series*, *25*(2), 23–38. https://doi.org/10.59170/stattrans-2024-013

Tirink, C., & Önder, H. (2022). Comparison of M, MM and LTS estimators in linear regression in the presence of outlier. *Turkish Journal of Veterinary & Animal Sciences*, *46*(3), 420–428. https://doi.org/10.55730/1300-0128.4212

Yang, H., Xu, T., Chen, D., Yang, H., & Pu, L. (2020). Direct modeling of subway ridership at the station level: A study based on mixed geographically weighted regression. *Canadian Journal of Civil Engineering*, *47*(5), 534–545. https://doi.org/10.1139/cjce-2018-0727

Zhang, S., Wang, L., & Lu, F. (2019). Exploring Housing Rent by Mixed Geographically Weighted Regression: A Case Study in Nanjing. *ISPRS International Journal of Geo-Information*, *8*(10), 431. https://doi.org/10.3390/ijgi8100431