Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 9, Number 1, 2025
DOI: 10.30829/zero.v9i1.24064
Page: 56-65

Evaluating Robust Estimators in Geographically Weighted Regression for Stunting Analysis at the District-Level Across Java: A Focus on Outlier Handling

[1] **Silfiana Lis Setyowati**
School of Data Science, Mathematics and Informatics, IPB University, Bogor, Indonesia
Ministry of Research, Technology, and Higher Education

[2] **Muhammad Nur Aidi**
School of Data Science, Mathematics and Informatics, IPB University, Bogor, Indonesia

[3] **Utami Dyah Syafitri**
School of Data Science, Mathematics and Informatics, IPB University, Bogor, Indonesia

[4] **Fitrah Ernawati**
Research Center for Public Health and Nutrition, National Research and Innovation Agency of Indonesia

Article Info

Article history:

Accepted, 28 May 2025

Keywords:

District Level Stunting;
M estimator;
MBG Program;
MM estimator;
Regression;
Robust Geographically Weighted;
S estimator.

ABSTRACT

Rate remains high and lags behind neighboring countries such as Vietnam and Thailand. This slow progress underscores the need for region-specific interventions and a deeper understanding of local factors driving stunting to meet the 14% national target. This study applies RGWR, an improvement over GWR for handling outliers. This method uses M, S, and MM estimators applied to the analysis of the prevalence of stunting among children under five the 2018 Riskesdas data across 85 districts in Java. Immunization reduces disease risk, growth monitoring detects stunting early, ARI management mitigates disease impact, parental height influences stunting risk, and working mothers improve family income and healthcare access, all contributing to reduced stunting. Given the regional variation in impact, stunting reduction policies should be spatially tailored, the MBG program should be prioritized in eastern Java regions.

Corresponding Author:

Silfiana Lis Setyowati,
Department of Statistics and Data Science,
IPB University, Indonesia
Email: silfianalissetyowati@gmail.com

1. INTRODUCTION

The government recently launched Free Nutritious Meals Program (Makan Bergizi Gratis - MBG), which is one of the efforts to reduce the high rate of stunting and malnutrition in Indonesia. Stunting is still a national health problem, where the latest stunting rate released by the government is still recorded at 21.5% in 2023, far from the set target of 14% in 2024. Historically, Indonesia's stunting prevalence has shown a downward trend, from 37.2% in 2013 to 30.8% in 2018, 24.4% in 2021, and down again to 21.5% in 2023. Although showing improvement, the rate of decline is still slower than that of some Southeast Asian countries with similar demographic characteristics, such as Vietnam (14.1%) and Thailand (11.0%) in 2022. Based on data from the Indonesian Health Survey 2023, no province in Java Island has achieved the government's target, even though the stunting prevalence rate in Banten is still 24% [1]. A more in-depth analysis of the factors influencing stunting is needed to achieve this target.

Several studies have shown that the factors that influence stunting are local [2], [3]. The effectiveness of the MBG program may vary by region, as a variety of local factors influence it. It will be more effective in areas where stunting is caused by limited access to nutritious food. Despite its great potential, the implementation of MBG faces several crucial challenges. One of the main issues is the very high fiscal burden, with an estimated budget of IDR 450 trillion until 2029, which raises concerns over long-term financial sustainability. Prioritising nutrition budgets for areas with higher impacts identified through Robust Geographically Weighted Regression (RGWR) analysis can increase cost efficiency.

One of the challenges in analyzing stunting data is the presence of outliers that can affect model results. The presence of just one outlier can disrupt the model estimation and affect the prediction of the surrounding areas, making the estimation results inaccurate. A robust model approach is needed to handle outliers well and accommodate this, such us RGWR using M, S, and MM estimators. The purpose of this study is to identify the determinants of stunting with RGWR. By using this approach, it is expected that more stable and accurate estimates will be obtained despite the presence of extreme or outlier data. In the context of the MBG program, the results of RGWR analysis can assist the government in identifying areas that require specific interventions based on local conditions. Thus, the application of RGWR plays an important role in supporting the formulation of MBG policies that are more targeted, efficient, and evidence-based.

2. RESEARCH METHOD

Several recent studies have addressed robust estimation in regression models in recent years. Study about GWPR on sugarcane harvest data in East Java find that show that Estimator M provides more stable predictions than classical panel regression and is effective in reducing outliers, although still sensitive to extreme outliers [4]. Several recent studies have addressed robust estimation in regression models. One of the recent studies proposed three robust estimation methods (M, S, and MM) for panel Poisson regression models, aiming to improve robustness against outliers. The MM estimator is the best, with the smallest MSE value, compared to S, M, and the classical method [5]. Another study about GWPR on sugarcane harvest data in East Java find that show that Estimator M provides more stable predictions than classical panel regression and is effective in reducing outliers, although still sensitive to extreme outliers [6]. Research on the Robust Application of the MM estimator is still rarely applied to spatial regression models, However, it has great potential because it combines the strengths of the M and S estimators simultaneously. Therefore, this study will compare the performance of robust M, S, and MM estimators in the GWR model using data on the prevalence of stunting among toddlers in all districts across Java. This research was conducted following the steps of the research flowchart in Figure 1.

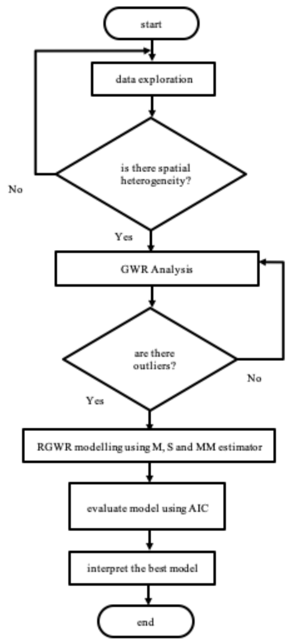

Figure 1 . Research Flowchart

2.1 Data

The data used in this study are the result from Basic Health Research Survey (Riskesdas) 2018. The data used included all toddlers (children under 5 years old) in 85 districts on the island of Java, totaling 17,822 individuals. The response variable is the prevalence of stunting in each district and city. This study uses a total of 12 predictor variables, namely: percentage of children under five who were fully immunized (X_1), percentage of children under five who were growth monitored (X_2), percentage of children under five who were given vitamin A (X_3), percentage of children under five who suffered ISPA in the last 12 months (X_4), percentage of children under five who suffered pulmonary tuberculosis in the last 12 months (X_5); percentage of children under five who suffered from diarrhea in the last month (X_6), average height of fathers (X_7), average height of mothers (X_8), percentage of working fathers (X_9), percentage of working mothers (X_{10}), percentage of low-income families (X_{11}), and percentage of households using an appropriate water source for drinking (X_{12}).

Each variable was aggregated by weighting the number of cases by the number of children under five in each district to calculate the percentage prevalence of each indicator. This follows a commonly used epidemiological formula to express proportions. as shown below:

$$X_{it} = \frac{S_{it}}{R_t} x\ 100\% \tag{1}$$

where X_{it} is the *i-th* variable in the *t-th* district, S_{it} is the number of cases of the *i-th* variable in the *t-th* district, and R_t is the number of children under five in the *t-th* district. Meanwhile, the parental height variable is presented in the form of an average.

$$X_{7t} = \frac{\sum_{i=1}^{N_t} H_{it}}{R_t} \text{ and } X_{8t} = \frac{\sum_{i=1}^{N_t} I_{it}}{R_t} \tag{2}$$

where H_{it} is the father's height for the i-th toddler in the t-th district and I_{it} is the mother's height for the i-th toddler in the t-th district.

2.2 Methods

In GWR analysis, the presence of outlier data is a significant methodological challenge as it can decrease the accuracy of parameter estimates. Outliers refer to extreme values that deviate far from the general pattern of the data; in this study, outliers are identified as extreme residual values from the GWR model results. An outlier can broadly impact a spatial context, as data also influence local parameter estimates in surrounding areas through spatial weighting. To overcome these problems, a robust estimation approach is used, which can reduce the model's sensitivity to the influence of extreme data. Robust models are designed to produce estimates that remain stable despite outliers in the data.

Some robust estimators that can be applied to handle outliers are M, S, and MM estimators. M estimator is a development of the maximum likelihood estimation method that focuses on minimizing the objective function ρ to reduce the influence of outliers in the data. The principle of the M estimator is to provide a balanced weight around the average value, which is expected to reduce the influence of outliers on the estimation results [6]. The M estimator is incorporated with spatial weighting as follows:

$$\hat{\beta}_{M((u_t,v_t)} = \arg\min_{\beta} \sum_{i=1}^{n} w_{it}.\rho\,(y_i - x_i^T\,\beta) \tag{3}$$

w_{it} is the spatial kernel weight that reflects the proximity between location i and t. To solve the minimization problem, the iteratively reweighted least squares (IRLS) method is used

The S estimator utilizes the standard deviation of the residuals, which is a solution to overcome the limitations of the M estimator, which only relies on the median without considering the data distribution [7]. The S estimator is defined as:

$$\hat{\beta}_{S(u_t,v_t)} = \arg\min_{\beta} \sum_{i=1}^{n} w_{it}.\left(\frac{y_i - x_i^T\beta}{\hat{\sigma}}\right)^2 \tag{4}$$

where $\hat{\sigma}$ is a robust estimate of the residual standard deviation.

The MM estimator combines the M and S estimators, where the estimation goes through two stages involving scale estimation and weighted iteration to obtain regression parameters that are more robust to outliers [8]. Mathematically, the MM estimator is defined as follows:

$$\hat{\beta}_{MM} = \arg\min_{\beta} \sum_{i=1}^{n} w_{it}.\rho 1\,(y_i - x_i^T\,\beta) + \lambda \sum_{i=1}^{n} w_{it}.\rho 2(\hat{\sigma}_i) \tag{5}$$

where $\rho 1$ is the robust loss function applied to the residuals $\rho 2$ is another robust function applied to the scale estimate of the residuals, and λ is the regularization parameter that controls the trade-off between the two components in this objective function.

The data analysis procedure was carried out in the following stages:

1. Data Pre-Processing
   a. Create response variables and target predictor variables according to equations (1) and (2).
   b. Detect multicollinearity by using the Variance Inflation Factor (VIF) value, correlation analysis and VIF measurement are important for detecting and addressing multicollinearity in regression models because they can cause multicollinearity [9].
2. GWR Modeling and GWR Outlier Detection
   In this study, outliers refer to deviations originating from GWR model residuals, namely residual values that differ significantly from other observed residuals. The presence of one outlier in GWR regression can significantly affect local parameter estimates, as it affects estimates in the surrounding area.
   a. Detect spatial heterogeneity with the Breush Pagan test statistic [10].
   b. If there is spatial heterogeneity, GWR modeling is continued in each district according to equation [11]:

$$Y_t = \beta_0(u_t, v_t) + \sum_{k=1}^{p} \beta_k\,(u_t, v_t) X_{tk} + \varepsilon_t \tag{6}$$

   c. Detecting GWR outliers using Cook's Distance. The Cook's Distance statistical formula is as follows [12]:

$$D_i = \frac{e_i^2 h_i}{d.MSE.(1-h_i)^2} \tag{7}$$

   where h_i is the hat value or leverage for the *i*-th data point, which measures how much the data point contributes to the model prediction. The greater the value h_i, the greater the data point's influence on the model. While *d is* the number of parameters or coefficients in the regression model, MSE is the mean square error or mean square of the residuals in the regression model.

3. RGWR modeling with M, S, and MM estimators in each district according to equations 3, 4, and 5.

4. Evaluation of the Best Model
   Model evaluation is performed using the Akaike Information Criterion (AIC), which is used to select the best model based on the balance between model fit and complexity. The main advantage of AIC over other model selection methods lies in its ability to maintain an optimal balance between modeling accuracy and the number of parameters, making it highly effective in predictive modeling contexts [13]. The lower the AIC value, the better the model balances fit and complexity. Mathematically, AIC is formulated as follows:
   $$AIC = 2d - 2\ln(L) \tag{8}$$
   Where d is the number of parameters or coefficients in the regression model, and $ln(L)$ is the natural logarithm of the estimated model likelihood.
5. Perform model validation using the Leave-One-Out Cross-Validation (LOOCV) technique to assess the model's accuracy.
6. Interpret the results of the best model estimation between M, S, or MM RGWR estimators.

3. RESULT AND ANALYSIS

3.1 Exploration of Stunting Conditions in Java Island Districts

The prevalence of stunting in Java Island districts shows clear variations, with values between 14.70% and 47.90%. These results show that the prevalence of stunting among children under five in 2018 is still far from the government's target, and no district has even reached the target. The average stunting prevalence was 30.14%, while the median value was 30.30%, indicating that most districts had stunting rates comparable to the average.

WHO categorizes the prevalence of stunting into four categories, namely low (less than 20%), medium (20-29%), high (30-39%) and very high (more than 40%). Most (56.5%) districts on Java Island have a high stunting prevalence status, while only 1 district (1.2%) has a low stunting prevalence status, as can be seen in Figure 2. What needs to be a serious concern is the presence of 7 districts with very high status.

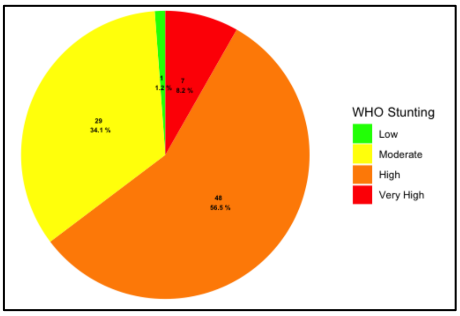

Figure 2. Stunting Condition of Java Island Districts

3.2 Data Pre-Processing

In building a regression model, in addition to correlation, it is necessary to consider the VIF value so that the model avoids multicollinearity, which can affect the estimation results. A VIF value of more than 10 should be removed from the model because it indicates high multicollinearity [14]. The VIF value detailed in Table 1 shows that no variable has a VIF value of more than 10. Based on the correlation and VIF values, multicollinearity is not indicated so that all predictor variables can be retained.

Table 1. VIF Value for Predictor Variables

Variable	VIF Value	Variable	VIF Value	Variable	VIF Value
X_1	1.550	X_5	1.167	X_9	1.548
X_2	2.861	X_6	2.291	X_{10}	1.447
X_3	2.616	X_7	3.395	X_{11}	1.335
X_4	2.186	X_8	2.709	X_{12}	2.035

The boxplot visualization of the twelve predictor variables in figure 3 shows the diversity of data distribution among districts in Java. The variables of complete immunization achievement (X_1), percentage of children under five with diarrhea (X_5), father's height (X_7), mother's height (X_8), and percentage of working fathers (X_9) have narrow distributions, meaning that districts tend to be homogeneous. Other predictor variables, such as growth monitoring (X_2), have a wide distribution, meaning there is variation between districts. Although several outliers were identified in some variables, the data were retained in the analysis as they were considered to reflect valid empirical conditions. The removal of extreme values could potentially remove important information about disparities between regions, which is important to reveal in spatial analysis.

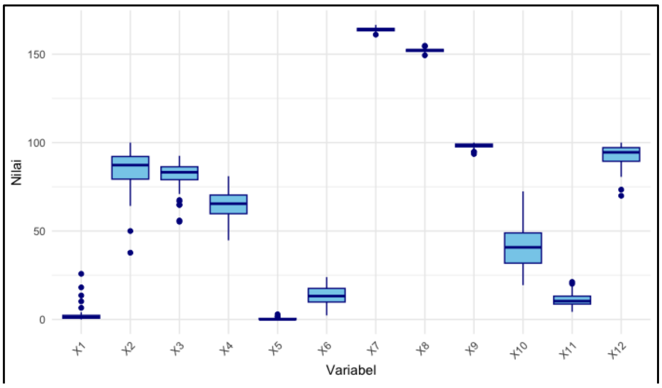

Figure 3. Boxplot of Predictor Variables

3.3 GWR Modeling and Outlier Detection

We found spatial heterogeneity in the stunting prevalence data of districts on Java Island. This is indicated by the Breush-Pagan test statistic value of 27.286 with a p-value of 0.007027, which rejects H_0. This means the GWR model is more suitable than the classical regression model [15].

To obtain local parameter estimates, we use equation (6). As an example, to calculate the estimate for Blitar City (59-th observation), we calculate:

$$\hat{\boldsymbol{\beta}}(u_{59}, v_{59}) = \left(\begin{bmatrix} 1 & x_{1,1} & \cdots & x_{1,12} \\ 1 & x_{2,1} & \cdots & x_{2,12} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{85,1} & \cdots & x_{85,12} \end{bmatrix}' \cdot \begin{bmatrix} w_1^{(59)} & 0 & \cdots & 0 \\ 0 & w_2^{(59)} & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{85}^{(59)} \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} 1 & x_{1,1} & \cdots & x_{1,12} \\ 1 & x_{2,1} & \cdots & x_{2,12} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{85,1} & \cdots & x_{85,12} \end{bmatrix}'$$

$$\cdot \begin{bmatrix} w_1^{(59)} & 0 & \cdots & 0 \\ 0 & w_2^{(59)} & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{85}^{(59)} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{85} \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta}_{0(u_{59},v_{59})} \\ \hat{\beta}_{1(u_{59},v_{59})} \\ \vdots \\ \hat{\beta}_{12(u_{59},v_{59})} \end{bmatrix} = \begin{bmatrix} 305.904 \\ -0.301 \\ \vdots \\ 0.009 \end{bmatrix}$$

In this study, we used Cook's Distance test statistic to detect the presence of outliers in the residuals of the GWR model. Although other methods may be superior in some cases, Cook's Distance remains one of the practical tools for identifying influential observations in spatial models. If the Cook's Distance value is> $\frac{4}{n}$, where n is the number of observations, then the observation is considered an outlier. For example in Trenggalek District, a residual $e = 13.606$, leverage $h = 0.306$, total of parameters $d = 13$, and MSE of GWR model $MSE = 14.351$ we compute

$$D_i = \frac{13.606^2 \ x \ 0.306}{13 \ x \ 14.351 \ x \ (1-0.306)^2} = 0.633$$

Since the cutoff value $\frac{4}{85}$ of 0.047, Trenggalek is confirmed as an outlieras. Thre are 12 districts are classified as outliers based on Cook's Distance. The areas detected as outliers and their respective Cook's Distance values are presented in Table 2.

Table 2 . Regions detected as outliers

No.	District	Cook's Distance Value	No.	District	Cook's Distance Value
1	Purwakara	0.056	7	Pekalongan	0.169
2	Sampang	0.080	8	Rembang	0.244
3	Pamekasan	0.097	9	Wonogiri	0.273
4	Ngawi	0.101	10	Gunung Kidul	0.476
5	Demak	0.109	11	Jepara	0.527
6	Sleman	0.121	12	Trenggalek	0.633

Figure 4 shows the mapping of outlier areas in the central and eastern parts of Java. These areas require further attention, as outliers may indicate the presence of local factors that significantly influence stunting prevalence.

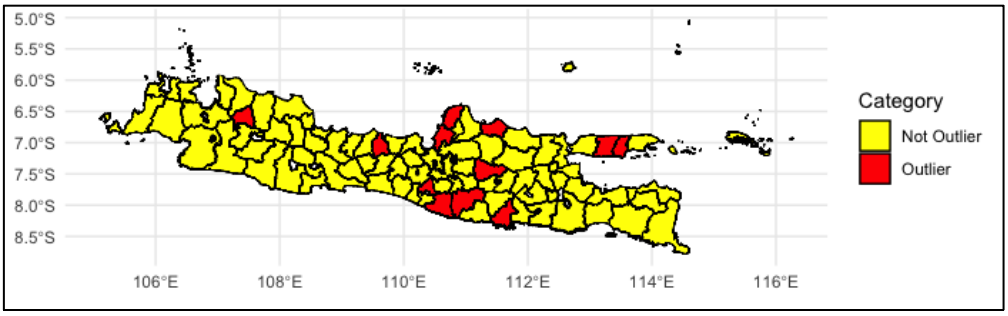

Figure 4. Outlier Map Based on Cook's Distance

3.4 RGWR Modeling of M, S, and MM Estimators

To reduce the influence of outliers on the results of local GWR parameter estimation, this study applies the RGWR approach using M, S, and MM estimators. The estimation of RGWRM model parameters is based on equation (3), which is solved numerically using the IRLS method. The Tuckey loss function is used as the weight function multiplied by the spatial kernel weights and iterated until convergence of $\hat{\beta}_{M(u_t,v_t)}$ is achieved. A similar procedure is also performed to obtain $\hat{\beta}_{s(u_t,v_t)}$ as defined in equation (4) and $\hat{\beta}_{MM(u_t,v_t)}$ as defined in equation (5).

The results of RGWR modeling with three types of estimators, namely M, S, and MM estimators, show variations in the performance of parameter estimation in each method. For an example we calculate the p-value for observation 59 on X_1 with degrees of freedom 72:

$$p - value_{59} = 2\,.\,(1 - F_{72}(|t|)$$
$$p - value_{59} = 2\,.\,(1 - F_{72}(|8.64|) = 1{,}18\ x\ 10^{-3}$$

Since the p-value is less than 0.05, variable X1 is considered statistically significant at the 95% confidence level. The summary of variables that are statistically significant across all districts or in some districts only, based on each estimation method, is presented in Table 3.

Tabel 3. Significant variables identified by RGWR estimators

Method	Variables significant at all districs	Variables significant at some distics only
M Estimation	X_1, X_2	X_4, X_5, X_6, X_9, X_{12}
S Estimation	X_1, X_2, X_3, X_4, X_5,X_6, X_8,X_{10}, X_{12}	
MM Estimation	X_1, X_2, X_3, X_4, X_5,X_6, X_7, X_8,X_9, X_{10}, X_{11}, X_{12}	

The M estimator adopts the principle of maximum likelihood estimation, which focuses on reducing the influence of outliers by giving lower weights to data far from the median. The analysis found that RGWRM is still less robust or affected by outliers, so it captures less significant variables than other methods. Some significant variables in other methods not captured significantly in this model are X_3, X_7, X_8, X_{10} and X_{11}. This is to previous research, which states that although the M estimator is designed to reduce the influence of outliers, it is still sensitive to very extreme outliers that can affect the estimation results [16].

The S estimator was developed to improve the weakness of the M estimator by incorporating the residual standard deviation scale in the estimation. The advantage of the S estimator is that it can provide more robust estimates of outliers, especially on variables with a broader distribution. The superiority of the S estimator in handling data variation and outliers positively impacts the number of significant variables in the model. By improving the stability of the estimates and reducing the influence of extreme data, RGWRS produces more variables that show significant influence than RGWRM.

The MM estimator is an estimation technique that combines the advantages of the M-estimator and S-estimator to achieve high efficiency and breakdown values [17]. RGWRMM model can produce optimal parameter signification. The MM estimator detected previously insignificant variables, such as X_{17} (average of father's height), X_{11} (percentage of working father), and X_{11} (proverty rate), which became statistically significant in this model, indicating that the MM estimator has a higher ability to identify relevant determinants, even though the data contains outliers.

3.5 Model Evaluation

The performance of the three models was evaluated using the AIC criterion. A model with a lower AIC value is considered better, as it indicates that it provides a better fit without adding complexity. Table 4 compares the performance of the three models. RGWRMM shows the best performance with the lowest AIC value and optimal number of significant parameters. RGWRS is superior to RGWRM due to its smaller AIC value and a larger number of significant parameters. These results indicate that RGWRMM is the most robust model in handling outliers, with the most significant variables than the other models.

Table 4 . AIC Value

Model	AIC	Number of Significant Parameters
RGWRM	476.177	7
RGWRS	288.033	9
RGWRMM	262.472	12

Model validation was performed using the LOOCV method, resulting in an RGWRMM LOOCV RMSE value of 4.726967. When converted relatively, the average RGWRMM model prediction deviated by approximately 14.72% from the actual value, indicating a relatively good level of prediction accuracy.

3.6 Determinants Of Stunting

Based on the RGWRMM model, the factors that contribute to reducing stunting in Java Island districts are complete immunization (X_1), growth monitoring (X_2), ARI treatment (X_4), average father's height (X_7), average mother's height (X_8), and percentage of working mothers (X_{10}). Increases in vitamin A coverage (X_3), percentage of working fathers (X_9), and proper sanitation (X_{12}) have not been able to reduce the prevalence of stunting in Java Island districts. Meanwhile, an increase in toddlers suffering from pulmonary tuberculosis (X_5) and diarrhea (X_6) contributed to increasing the prevalence of stunting.

Complete immunization prevents infectious diseases that can interfere with child growth. A study of 16 developing countries in the world showed that children with incomplete vaccination schedules are more prone to stunting, underweight, and wasting, which are indicators of chronic malnutrition [18]. In addition to immunization, growth monitoring efforts through routine height and weight measurements also play an important role in reducing stunting [19,20]. This study found that the percentage of toddlers suffering from ARI can contribute to reducing stunting. This is associated with effective ARI management, where areas of high ARI prevalence in children under five tend to carry out interventions to treat ISPA that simultaneously reduce the prevalence of stunting, such as increasing the coverage of complete basic immunization and improving nutritional status in children under five [21]. For example, Tegal City, which has the highest ARI prevalence, reduced stunting prevalence from 37.1% in 2013 to 30.6% in 2018. Genetic factors such as parental height, both father and mother, significantly influence the incidence of stunting prevalence. The highest risk of stunting was found in children born to parents with the shortest height [22]. Meanwhile, an increase in the number of working mothers was also found to reduce the prevalence of stunting. This was associated with maternal education and improved family economics. A previous study revealed that the incidence of stunting in Indonesia is closely related to economic status, where families with low economic status are more at risk of having children with stunted status [23].

Spatial mapping of the determinants of stunting based on the estimated coefficients of the RGWRMM model is presented in Figure 5. Each map shows the distribution of the quartile values of the regression coefficients for each predictor variable at the district/city level in Java, classified according to the direction and strength of their association with stunting prevalence. Yellow to blue indicates negative quartile values (Q1 to Q3), indicating variables negatively correlated with stunting. In contrast, pink to purple colours indicate positive quartile values (Q1 to Q3), indicating a positive correlation or increased risk of stunting.

Variables with positive estimated coefficients, indicating a tendency toward increased stunting prevalence, also exhibit interesting spatial patterns. The percentage of infants suffering from pulmonary tuberculosis has a stronger positive influence in the western part of Java Island, suggesting that the burden of infectious diseases in this region significantly contributes to increased stunting rates. On the other hand, the prevalence of diarrhea and poverty levels show an increasing influence on the rise in stunting in the eastern part of Java Island. In western Java, such as Jakarta, Banten, and Bogor, high population density and crowded living conditions can increase the transmission of respiratory diseases such as pulmonary tuberculosis, especially in vulnerable communities with poor ventilation and limited access to health services. Conversely, eastern Java—such as Banyuwangi, Bondowoso, and Sampang—has higher poverty rates and limited access to clean water and sanitation. These conditions fuel the spread of waterborne diseases like diarrhea, particularly among infants. Meanwhile, increases in vitamin A coverage, the percentage of working fathers, and improved access to clean water have not been able to reduce stunting prevalence in the districts of Java Island significantly. These findings emphasize the importance of area-based approaches in policy planning, as stunting determinants are not geographically homogeneous.

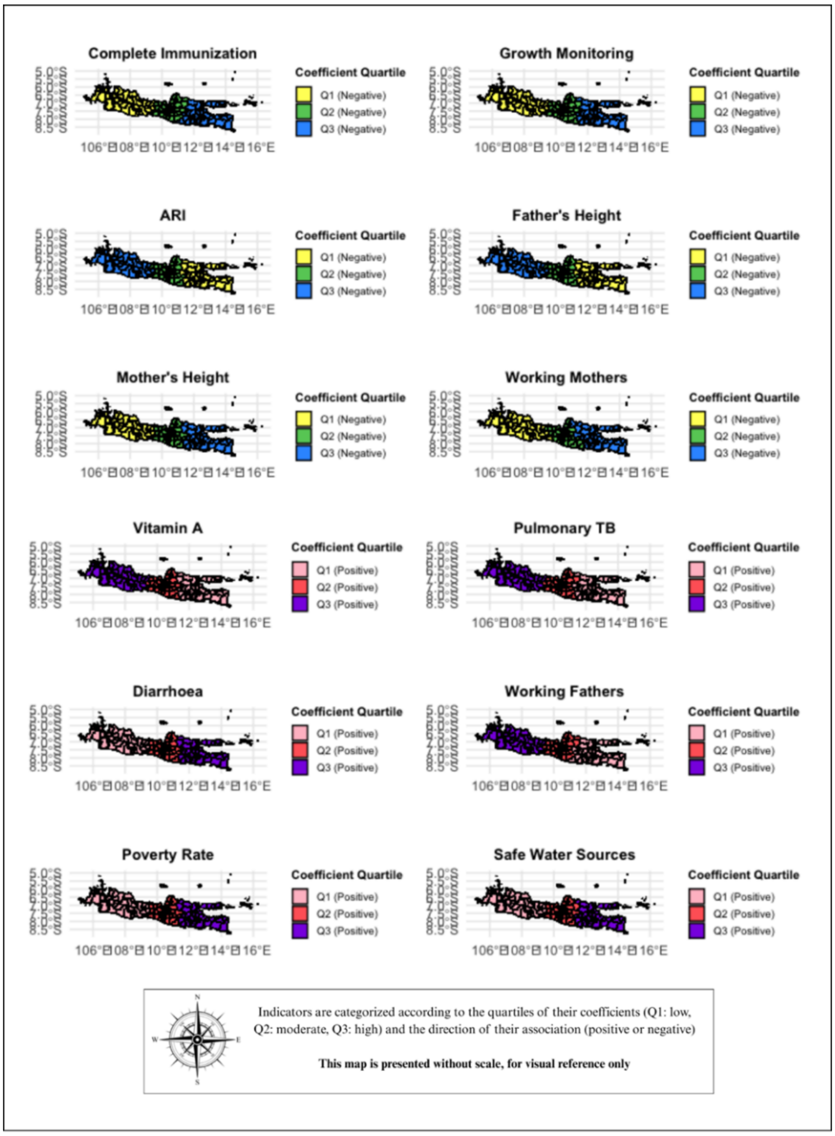

Figure 5. Mapping the Determinants of Stunting

4. CONCLUSION

Based on the results obtained, the RGWRMM model is the best model for predicting stunting prevalence in districts on Java Island. The RGWRMM model outperformed the other models by achieving the lowest AIC value and the optimal number of statistically significant predictor variables, demonstrating its superior ability to model the spatial variation of stunting prevalence in the districts of Java Island.

Factors that contribute to reducing stunting in Java Island districts are complete immunization (X1), growth monitoring (X2), ARI treatment (X4), average father's height (X7), average mother's height (X8), and percentage of working mothers (X10). Complete immunization and growth monitoring help prevent stunting by reducing the risk of infectious diseases and enabling early detection of growth issues. Effective ARI management also supports better growth outcomes. In addition, parental height influences stunting risk, with shorter parents more likely to have stunted children. Working mothers tend to improve household income and access to healthcare, which positively affects child nutrition and reduces stunting prevalence. Based on these findings, intervention policies that need to be implemented to reduce the prevalence of stunting in Java Island can be tailored to the characteristics of each region. In this context, the MBG program should be prioritized in eastern Java regions where poverty levels are high, leading to the prevalence of stunting.

Further research is recommended to integrate more sophisticated spatial clustering techniques to identify areas with similar sociodemographic and environmental characteristics or to apply Geographically Weighted Interaction Regression (GWIR) to analyze the interaction effects between predictor variables on stunting. Additionally, relevant additional predictors such as the quality and accessibility of healthcare services, the affordability of nutritious food, and environmental factors like water quality and exposure to pollution are recommended to enhance the accuracy of the model.

5. REFERENCES

[1] Ministry of Health, "Indonesia Health Survey (IHS) 2023 In Figures: Accurate Data for the Right Policy," Jakarta: BPKP, 2023.

[2] G. Nduwayezu, C. Kagoyire, P. Zhao, L. Eklund, P. Pilesjo, J. Bizimana, and A. Mansourian, "Spatial Machine Learning for Exploring the Variability in Low Height-For-Age From Socioeconomic, Agroecological, and Climate Features in the Northern Province of Rwanda," *GeoHealth*, vol. 8, 2024, doi: 10.1029/2024GH001027.

[3] S. Rahardiantoro, A. Juhanda, A. Kurnia, A. Aswi, B. Sartono, D. Handayani, A. Soleh, Y. Yanti, and S. Cramb, "Spatio-temporal modeling to identify factors associated with stunting in Indonesia using a Modified Generalized Lasso," *Spatial and Spatio-temporal Epidemiology*, vol. 51, 2024, p. 100694, doi: 10.1016/j.sste.2024.100694.

[4] E. Ahmed, M. Abonazel, M. Al-Ghamdi, H. Elshamy, and I. Khattab, "Proposed robust estimators for the Poisson panel regression model: application to COVID-19 deaths in Europe," *Communications in Mathematical Biology and Neuroscience*, 2024, doi: 10.28919/cmbn/8795.

[5] E. Ahmed, M. Abonazel, M. Al-Ghamdi, H. Elshamy, and I. Khattab, "Proposed robust estimators for the Poisson panel regression model: application to COVID-19 deaths in Europe," *Communications in Mathematical Biology and Neuroscience*, 2024, doi: 10.28919/cmbn/8795.

[6] Quarta, Y., Mondiana, Pramoedyo, H., & Iriany, A. (2024). Applied fixed effect of Geographically Weighted Panel Regression (GWPR) with M- Estimator approach to estimate sugarcane yield data in East Java. *Journal of Applied and Natural Science*. doi: 10.31018/jans.v16i2.5443.

[7] T. Fayose, K. Ayinde, O. Alabi, and A. Bello, "Robust weighted ridge regression based on S-estimator," *African Scientific Reports*, 2023, doi: 10.46481/asr.2023.2.3.126.

[8] K. Khotimah, K. Sadik, and A. Kurnia, "The MM estimator method for robust regression," *Journal of Physics: Conference Series*, vol. 1863, no. 1, p. 012033, 2020, doi: 10.1088/1742-6596/1863/1/012033.

[9] C. Garcia, R. Gómez, and C. García, "Choice of the ridge factor from the correlation matrix determinant," *Journal of Statistical Computation and Simulation*, vol. 89, pp. 211–231, 2018, doi: 10.1080/00949655.2018.1543423.

[10] D. Wirwana, M. N. Aidi, and A. Fitrianto, "Analysis of factors affecting district/city GRDP in Kalimantan Island," *International Journal of Science: Basic and Applied Research (IJSBAR)*, vol. 64, no. 1, pp. 131-148, 2022.

[11] M. N. Aidi, F. Ernawati, E. Efriwati, N. Nurjanah, R. Rachmawati, E. D. Julianti, D. Sundari, F. Retiaty, A. Fitrianto, K. Nurfadilah, and A. Y. Arifin, "Spatial distribution and identifying biochemical factors affecting hemoglobin levels among women of reproductive age for each province in Indonesia: A geospatial analysis," *Geospatial Health*, vol. 17, no. 2, p. 1118, 2022, doi: 10.4081/gh.2022.1118.

[12] J. A. Padrón Hidalgo, A. Pérez-Suay, F. Nar, and G. Camps-Valls, "Nonlinear Cook Distance for Anomalous Change Detection," *arXiv preprint arXiv:2012.12307*, 2020.

[13] J. Zhang, Y. Yang, and J. Ding, "Information criteria for model selection," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 15, 2023, doi: 10.1002/wics.1607.

[14] A. Ghosh and N. Varadharajan, "The association between neighborhood characteristics and depression: was the regression model satisfactory?" *The British Journal of Psychiatry*, vol. 216, p. 235, 2020, doi: 10.1192/bjp.2020.19.

[15] S. Sifriyani, I. Mandang, F. Amijaya, and R. Ruslan, "Developing Geographically Weighted Panel Regression Model for Spatio-Temporal Analysis of COVID-19 Positive Cases in Kalimantan, Indonesia," *Journal of Southwest Jiaotong University*, vol. 57, no. 3, 2022, doi: 10.35741/issn.0258-2724.57.3.10.

[16] D. Khan, M. Ali, Z. Ahmad, S. Manzoor, and S. Hussain, "A new efficient redescending M-estimator for robust fitting of linear regression models in the presence of outliers," *Mathematical Problems in Engineering*, 2021, doi: 10.1155/2021/3090537.

[17] A. Prahutama and A. Rusgiyono, "Robust regression with MM-estimator for modeling the number of maternal mortality of pregnancy in Central Java, Indonesia," *Journal of Physics: Conference Series*, vol. 1943, 2021, doi: 10.1088/1742-6596/1943/1/012148.

[18] M. Solis-Soto, D. Paudel, and F. Nicoli, "Relationship between vaccination and nutritional status in children: Analysis of recent Demographic and Health Surveys," *Demographic Research*, vol. 42, 2020, doi: 10.4054/demres.2020.42.1.

[19] C. Wright, F. Petermann-Rocha, R. Bland, P. Ashorn, S. Zaman, and F. Ho, "Weight velocity in addition to latest weight does not improve the identification of wasting or the prediction of stunting and mortality: A longitudinal analysis using data from Malawi, South Africa, and Pakistan," *The Journal of Nutrition*, vol. 154, pp. 2583-2589, 2024, doi: 10.1016/j.tjnut.2024.06.011.

[20] Himmah, E. F., Kaestria, R., & Riana. (2025). Mathematical modeling of stunting with the influence of nutritional intervention. *JTAM (Jurnal Teori dan Aplikasi Matematika)*, *9*(1), 54–67. https://doi.org/10.31764/jtam.v9i1.26817

[21] V. Sutriana, M. Sitaresmi, and A. Wahab, "Risk factors for childhood pneumonia: A case-control study in a high prevalence area in Indonesia," *Clinical and Experimental Pediatrics*, vol. 64, pp. 588-595, 2021, doi: 10.3345/cep.2020.00339.
[22] H. Wu, C. Yang, and B. Xi, "Association of parental height with offspring stunting in 14 low- and middle-income countries," *Frontiers in Nutrition*, vol. 8, 2021, doi: 10.3389/fnut.2021.650976.
[23] F. Ernawati, Y. Octaria, and Y. Widodo, "Economic Status, Stunting, and Diet Quality as Important Determinants of Anaemia among Indonesian Children aged 6-35 Months Old: A SEANUTS Study," *Malaysian Journal of Medicine and Health Sciences*, vol. 16, Supp 13, 2020, doi: 10.2636/2636-9346.