International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN - 2395-1990
Online ISSN : 2394-4099
OPEN ACCESS
Available Online at :www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET25122143

Factors Influencing Stunting in Indonesia 2018 With Geographically Weighted Regression Analysis

Rizqi Dwi Yuniarsyih R.A[1], Hakim Zoelva Mahesa[1], I Putu Gde Inov Bagus Prasetya[1], Syella Zignora Limba[1], Muhammad Nur Aidi[1], Yekti Widodo[2], Elisa Diana Julianti[2], Fitrah Ernawati[2], Aya Yuriestia Arifin[2]
[1]Departement of Statistics and Data Science, IPB University, Bogor, Indonesia
[2]National Research and Innovation Agency

ARTICLEINFO

Article History:

Accepted : 26 March 2025
Published: 28 March 2025

Publication Issue :
Volume 12, Issue 2
March-April-2025

Page Number :
485-493

ABSTRACT

Stunting is a health issue with long-term effects on human resource quality. This study analyzes factors influencing stunting prevalence in Indonesia using the Geographically Weighted Regression (GWR) approach. Using 2018 provincial-level data, the study examines the impact of social, health, and education factors on stunting rates, categorized by gender. Analysis employs Global Regression to identify key explanatory variables before applying GWR to capture spatial variations. Results reveal significant regional differences, where factors such as Insurance, Participation School, Immunization, Happiness, Water, and Years School exhibit varying influences. The GWR model outperforms global regression, achieving R-squared values of 0.6641 (males) and 0.6589 (females), with AIC values of 506.871 and 499.160, respectively. These findings highlight the importance of localized policies to address stunting effectively.

Keywords: Stunting, Quadran Method, Global Regression, Geographically Weighted Regression.

INTRODUCTION

Stunting is one of public health issues that has global attention, especially in developing countries such as Indonesia. Stunting is defined as a condition of growth failure in children due to chronic nutritional deficiency, repeated infections, and inadequate psychosocial stimulation during the first thousand days of life (UNICEF 2019). According to data from the Central Statistics Agency and the Ministry of Health of the Republic of Indonesia, the prevalence of stunting in Indonesia remains at a rather high level, despite various efforts having been made to reduce these figures (Statistics Indonesia 2022).

The spread of stunting cases in Indonesia is not random; rather, it exhibits specific spatial patterns that can be influenced by various social, economic, and environmental factors (Ministry of Health Indonesia 2021). Therefore, a spatially-based approach is essential to understand how the distribution of stunting and the influencing factors vary across regions. Geographically Weighted Regression (GWR) modeling is one of the methods that can be employed to analyze the relationship between stunting and predictor variables in a spatial context (Fotheringham et al. 2002).

Previous studies have explored various spatial modeling approaches to understand the distribution of stunting. Research (Fadliana and Drajat 2021) applied spatial logistic regression to stunting data in West Java and found that complete basic immunization and food management facilities were significant factors influencing stunting prevalence. Similarly, research (Purwanti and Nurfita 2019) used a spatial regression model and identified household access to proper sanitation as a key determinant. While these studies effectively captured spatial dependencies in stunting cases, they did not incorporate gender as a distinguishing factor in their analyses.

Building on these findings, this study aims to model spatial regression for stunting cases in Indonesia while explicitly considering gender differences. Since gender can influence nutritional status and health outcomes, integrating gender-based analysis into spatial modeling can provide a more nuanced understanding of the spatial variability of stunting. This research applies Geographically Weighted Regression (GWR) to examine how the relationships between stunting prevalence and predictor variables vary across regions for male and female children.

METHODS AND MATERIAL

Geographically Weighted Regression (GWR)

Geographically Weighted Regression (GWR) is a regression method that allows regression parameters to vary at each location based on spatial information around it. Brunsdon et al. (1996) introduced the GWR method to address parametric non-stationarity, which is the condition where the relationship between variables is not constant across geographical areas. In GWR, each parameter is computed based on the observation location, resulting in locally varying parameters (Fotheringham et al. 2002). The GWR model equation can be formulated as follows:

$$y_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + \varepsilon_i$$

where:

- y_i is the value of the dependent variable at location i
- x_{ik} is the value of the independent variable k at location i
- x_{ik} is the local regression coefficient that varies by location
- x_{ik} is the random error.

Estimation of parameters in GWR is performed by considering spatial weights. Coefficient estimation in GWR modeling is carried out using Weighted Least Square (WLS) method, which assigns greater weights to nearby observation locations. With this approach, GWR is capable of capturing spatial variations in the relationships among variables more accurately compared to global regression. The equation for parameter estimation in matrix form is expressed as follows:

$$\hat{\boldsymbol{\beta}}_i = [\boldsymbol{X}^T \boldsymbol{W}_i \boldsymbol{X}]^{-1} \boldsymbol{X}^T \boldsymbol{W}_i \boldsymbol{Y}$$

where $\boldsymbol{W}_i$ is the weight matrix that determines the influence of surrounding observations on location i.

Spatial Weights

The selection of spatial weights in Geographically Weighted Regression (GWR) analysis is an important step in determining the relationships between locations within the model. Queen Contiguity is a weighting method that establishes spatial relationships based on the connectivity of sides and corners among regional units (Akolo 2022). The Queen Contiguity method assigns weights to all spatial units that share a side or corner with a specific location, while non-adjacent units receive a weight of zero. The Queen Contiguity weight matrix can be expressed as follows:

$$W_{ij} = \begin{cases} 1, if\ regions\ i\ and\ j\ are\ neighbors \\ 0, if\ not\ neighbors \end{cases}$$

K-Nearest Neighbor (KNN) method forms a weighting matrix based on the proximity to a predetermined number of nearby neighbors. Unlike Queen Contiguity method which only considers administrative boundaries, KNN gives weight based on the distance between points so that it is more flexible in capturing spatial relationships. In the KNN method, the distance between each location is calculated using the Euclidean distance, which is formulated as follows (Agustina et al. 2022):

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

where (u_i, v_i) and (u_i, v_i) is coordinates of location i and j are specifies. After calculating the distance, each location i is connected to k nearest neighbors based on the smallest distance value. The set of nearest neighbor locations is defined as:

$$J_k(i) = \{j[1], j[2], ..., j[k]\}$$

with k as the specified number of neighbors. In the formation of the weight matrix, weights are assigned only to the nearest neighbors according to the following scheme:

$$W_{ij} = \begin{cases} 1, if\ j\ included\ in\ J_k(i) \\ 0, if\ not \end{cases}$$

In addition, weights may also be assigned by considering distance using a kernel function, such as the Gaussian Kernel formulated as follows:

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

with h as the bandwidth that controls the decay rate of weights based on distance.

Spatial Autocorrelation Test with Moran's Index

Moran Index is used to identify the presence of spatial autocorrelation as well as the spatial distribution patterns that are formed, both in the form of clustering and trends in a specific area (Pfeiffer et al. 2008). The equation for Moran Index is as follows:

$$I = \frac{n}{\sum_i \sum_j W_{ij}} \times \frac{\sum_i \sum_j W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

The range of the Moran index is $-1 < I < 1$ with criteria according to Pfeiffer et al. (2008) as follows.

- If I = 0 , there is no spatial autocorrelation.
- If I > 0, positive spatial autocorrelation occurs, meaning that adjacent areas have similar values, resulting in a clustered distribution pattern.
- If I < 0, negative spatial autocorrelation occurs, meaning that adjacent areas have different values, resulting in a dispersed distribution pattern.

Data and Data Sources

The data used is secondary data from provinces in Indonesia in 2018. The data is divided into two categories, namely the number of stunting cases in males and females. The stunting data used in this study was obtained from Prof. Dr. Ir. M. Nur Aidi, MS. And Dr. Yekti Widodo, who have compiled and processed the dataset for research purposes. The variables used can be seen in Table 1.

Table 1 Variables Used

Code	Variable	Source
Stunting	Many cases of stunting	Aidi MN and Widodo Y
Insurance	Percentage of the Population with Health Insurance	Indonesian Health Profile
Participation_School	Pure Participation Rate of High School/Equivalent	Statistics Indonesia
Immunization	Percentage of children aged 12-23 months who receive	Indonesian Health

Code	Variable	Source
	complete basic immunization	Profile
Happiness	Happiness Index	Statistics Indonesia
Water	Percentage of Households with Access to Safe Drinking Water Sources	Statistics Indonesia
Pregnancy	Average age of a mother's first pregnancy	Statistics Indonesia
Years_School	Average years of schooling	Statistics Indonesia

Research Methodology

The research procedure consists of the following stages:

i. Global regression analysis (OLS) to identify factors influencing stunting in male and female groups before applying the spatial model.

ii. Multicollinearity test with Variance Inflation Factor (VIF) to detect high linear relationships among independent variables.

iii. Spatial assumption tests:
- Moran's Index test to examine spatial autocorrelation in the residuals of the OLS model.
- Lagrange Multiplier test (LM-Lag and LM-Error) to determine whether a spatial modeling approach is required.
- Breusch-Pagan test to detect heteroscedasticity in the residuals of the model.

iv. Selection of the GWR model if heteroscedasticity is present in the OLS model

v. Determination of geographical coordinates with longitude and latitude for spatial analysis.

vi. Selection of the optimal bandwidth using the Cross Validation (CV) method, choosing the bandwidth with the lowest CV value.

vii. Parameter estimation of the GWR model for each location using the selected optimal bandwidth.

viii. Model evaluation by comparing R-square and Akaike Information Criterion (AIC) between the OLS and GWR models to determine the best model.

ix. Mapping of significant variables in the GWR model to identify factors affecting stunting differently across regions.

RESULTS AND DISCUSSION

Data Exploration

Stunting is a condition of growth failure in children caused by chronic malnutrition, repeated infections, and inadequate psychosocial stimulation during the first 1,000 days of life (Black et al. 2013). According to WHO, a stunting prevalence of ≥20% indicates a serious public health issue (de Onis et al. 2018). With a prevalence rate of 30.8% in 2018 (Riskesdas 2018), Indonesia continues to face significant challenges in reducing stunting rates in line with global targets. Boxplot diagram illustrating number of stunted individuals by province for both males and females in 2018 is available in Figures 1 and 2

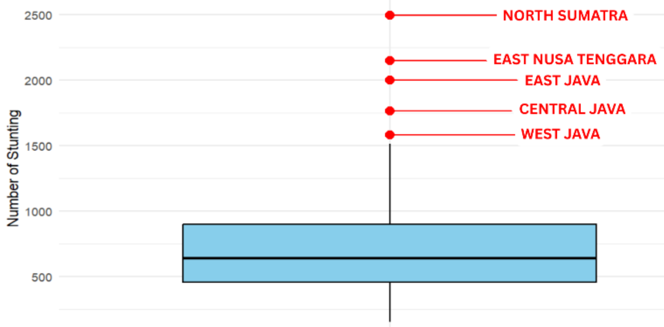

Figure 1 Boxplot of stunting number for males

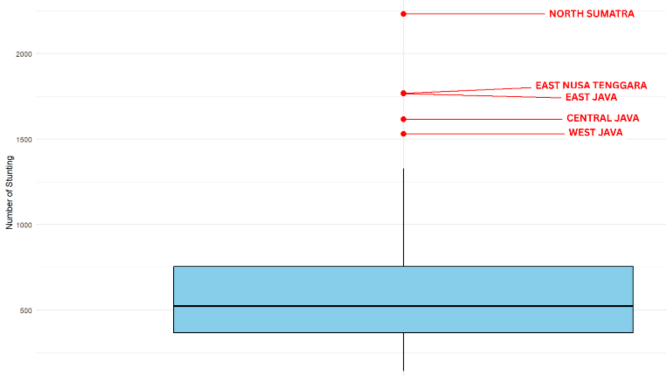

Figure 2 Boxplot of stunting number for females

Box plot diagrams in Figures 1 and 2 illustrate that for both males and females, there are 5 provinces that are outliers, namely North Sumatra, East Nusa Tenggara, East Java, Central Java, and West Java. Overall, the number of stunting cases in males is higher compared to females. This is consistent with study by Wamani et al. (2007) which revealed that in several developing countries, boys are more susceptible to stunting compared to girls.

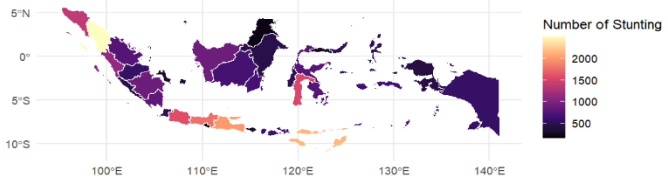

Figure 3 Distribution number of stunted males in Indonesia 2018

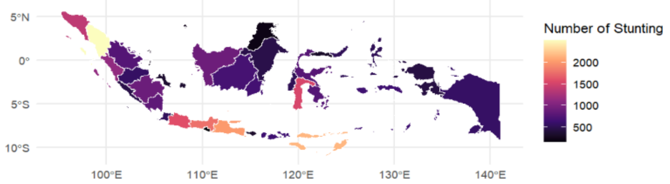

Figure 4 Distribution number of stunted females in Indonesia 2018

Figures 3 and 4 indicate that number of stunting individuals for both males and females displays a relatively similar distribution pattern. Province with the highest number of stunting cases in Indonesia for both males and females is North Sumatra. In contrast, Province with the lowest number of stunting cases for both males and females is Bali.

Quadrant Method

Quadrant method in spatial analysis is used to identify the distribution patterns of a phenomenon within a geographic area. This technique is often applied in epidemiological, ecological, and geographical studies to understand the spatial distribution of a specific variable (Boots and Getis 1988). In the context of stunting, this method divides the area into four categories based on the relationship between the number of stunting cases in males and females. Distribution pattern of stunting using quadrant method can be observed in Figure 5.

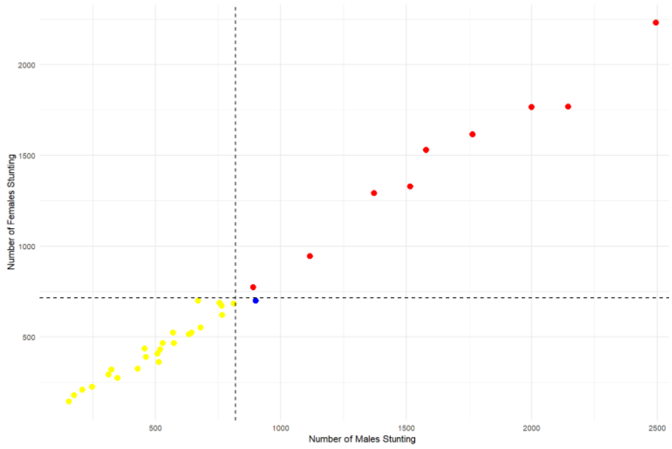

Figure 5 Distribution Pattern of Stunting using Quadrant Method

Figure 5 illustrates distribution pattern of stunting using Quadrant method, where number of male stunting cases is plotted on X-axis and number of female stunting cases is plotted on Y-axis. The graph is divided into four quadrants based on the median or average values, marked by dashed horizontal and vertical lines. The first quadrant (Red Point) indicates areas with a high number of both male and female stunting. Provinces that fall into quadrant 1 include Aceh, North Sumatra, West Sumatra, South Sumatra, West Java, Central Java, East Java, East Nusa Tenggara, and South Sulawesi. These areas can be regarded as stunting hotspots, indicating regions that require special attention in health and nutrition policy interventions.

Quadrant 2 (Blue Dot) indicates an area with a high number of male stunting but a low number of female stunting. The area that falls within Quadrant 2 is West Kalimantan. Quadrant 3 indicates an area with a low number of male stunting but a high number of female stunting; in this case, there are provinces that fall into Quadrant 3. Quadrant 4 (Yellow Dot) represents areas with low cases of stunting for both males and females. The majority of areas are in Quadrant 4, which means that most areas have relatively low stunting rates for both males and females.

Factors Influencing Stunting

Regression is a statistical technique used to measure relationship between independent variables (causal factors) and dependent variables (the number of stunting cases). In this context, Ordinary Least Squares (OLS) is used as the initial global regression model before applying the Geographically Weighted Regression (GWR) method to capture spatial variations in the relationships between variables (Brunsdon et al. 1996). Results of the global regression model for both male and female groups are presented in Table 2.

	Males		Females	
	Estimate	P- *Value*	Estimate	P- *Value*
Intercept	14547.437	0.0362	14156.381	0.0254
Insurance	-6.181	0.5133	-3.648	0.6696
Participation_School	31.123	0.1601	30.294	0.1326
Immunization	-11.665	0.0597	-9.576	0.0862
Happiness	-190.183	0.0221	-185.027	0.0148
Water	13.362	0.3055	11.791	0.3162
Pregnancy	-37.775	0.8134	-61.213	0.6735
Years_School	-136.150	0.3260	-99.132	0.4286
R-*square*	0.3627		0.3484	

Tabel 2 Result of Global Regression Model

In the male group, global regression results show an R-square value of 0.3627, meaning that only 36.27% of the variation in stunting is explained by the model with significant variables, namely Happiness and Immunization. Meanwhile, in female group, R-square value is 0.3484, indicating that only 34.84% of variation in stunting is explained by model with the significant variable is Happiness. Global regression analysis is less capable of capturing spatial heterogeneity as seen from the R-square value and number of insignificant variables. Due to these limitations, GWR was applied to account for spatial variations by allowing regression coefficients to vary across locations, making it more effective for analyzing spatially dependent phenomena (Fotheringham et al. 2002)

First stage in GWR is testing of spatial assumptions, including Moran's Index test for spatial autocorrelation, LM test for spatial dependence, and Breusch-Pagan test for heteroskedasticity. This testing ensures that relationships between variables are not only global in nature but also take into account the spatial variations that may occur at each location (Lesage and Pace 2009). Results of spatial assumption tests are presented in Table 3.

	Moran's Indeks	LM		Breusch-Pagan
		LMlag	LMerr	
Males	0.3854	0.7515	0.9672	0.02122
Females	0.3377	0.5842	0.8712	0.02343

Tabel 3 Results of spatial assumption tests

Table 3 shows that both in male group and female group, results of the Moran's Index and LM test are not significant, which means that there is no spatial autocorrelation in OLS model residuals (Anselin 1995) and there is no strong evidence that model requires additional spatial approaches (Lesage and Pace 2009). However, Breusch-Pagan test is significant, indicating the presence of heteroskedasticity in model, meaning that variability of errors is not constant across regions. Results of this assumption test indicate that GWR model is necessary to address spatial heterogeneity. GWR method will yield significantly different parameter values for each region (Fotheringham et al. 2002). Results of the GWR analysis can be seen in Figures 6 and 7.

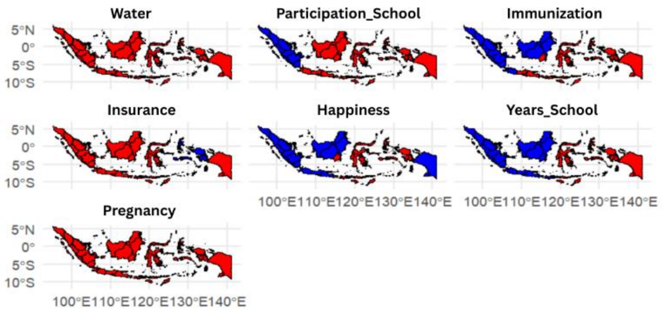

Figure 6 Significance Map of Each Variable in GWR Model for Male Group

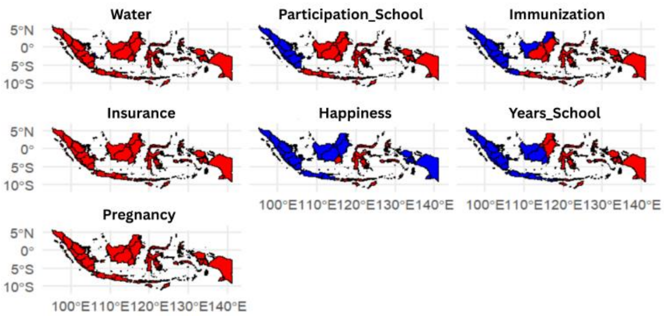

Figure 7 Significance Map of Each Variable in GWR Model for Female Group

Figures 6 and 7 illustrate that there are provinces with differing factors influencing the prevalence of stunting between male and female groups, indicated by the color blue. The disparities between provinces in Indonesia are caused by differences in access to healthcare services, infrastructure, and food consumption patterns (Sarkar et al. 2014). The areas where the causes of stunting differ between males and females include Bangka Belitung (Water), North Maluku and Maluku (Insurance), West Papua (Insurance and Happiness), Central Kalimantan (Immunization) and East Kalimantan (Immunization and Years_School), North Kalimantan (Years_School). In general, the variables that most significantly affect stunting cases are Happiness, Years_School, and Immunization, while Pregnancy is a variable that does not influence stunting across all regions of Indonesia.

Model	r-*square*		AIC	
	Males	Females	Males	Females
Global Regression	0.3627	0.3484	532.343	525.666
GWR	0.6641	0.6589	506.871	499.160

Table 4 Comparison of model goodness of fit

Table 4 presents a comparison of global regression model and GWR based on the r-square and AIC values. R-square value in the GWR model is higher than that in OLS, indicating a better predictive capability. Meanwhile, the lower AIC value of GWR compared to global regression indicates that model is more efficient and better suited for the data, leading to the conclusion that GWR method is superior in modeling stunting cases in Indonesia 2018 compared to the global regression model. This is consistent with research conducted by Nakaya et al. (2005), in which GWR has been used to evaluate spatial patterns in health issues, such as malnutrition and infectious diseases, with results indicating that this approach is more accurate than global regression models in capturing inter-regional differences.

CONCLUSION

This research analyzes factors influencing stunting prevalence in Indonesia using Geographically Weighted Regression (GWR) approach. Results indicate that GWR model is superior in capturing spatial variation compared to global regression. This advantage is evident from the higher R-square value and lower AIC. Variables such as Participation School, Immunization, Happiness, Water, and Years School have varying influences across regions, making it necessary to tailor stunting mitigation policies to the specific conditions of each area. Spatially-based approaches like GWR can serve as a more effective analytical tool for government in designing more targeted interventions. Future research is recommended to include additional variables such as food consumption patterns and environmental factors to enhance the model's accuracy in understanding the factors contributing to stunting in Indonesia.

REFERENCES

[1]. Aidi MN and Widodo Y. 2018. Stunting Dataset from Basic Health Research [Unpublished dataset]. Ministry of Health Indonesia.

[2]. Agustina, M., Abapihi, B., Wibawa, G. N. A., and Yahya, I. 2022. Modeling the Factors Influencing the Open Unemployment Rate in Indonesia Using a Spatial Regression Approach. In Seminar Nasional Sains dan Terapan VI. 56-70.

[3]. Akolo, I. R. (2022). Comparison of the rook and queen contiguity weight matrices in spatial autoregressive model and spatial error model.. Jambura Journal of Probability and Statistics. 3(1):11-18.

[4]. Anselin, L. (1995). Local Indicators of Spatial Association—LISA. Geographical Analysis, 27(2): 93-115.

[5]. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, Onis MD, Ezzati M, Grantham-Mcgregor S, Katz J, Martorell R, Uauy R. 2013. Maternal and child undernutrition and overweight in low-income and middle-income countries. The Lancet, 382(9890), 427-451

[6]. Boots, B. N. & Getis A. (1988). Point pattern analysis. SAGE Publications.

[7]. Breusch, T. S., and Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica, 47(5): 1287-1294.

[8]. Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. Geographical analysis, 28(4), 281-298.

[9]. de Onis M, Borghi E, Arimond M, Webb P, Croft T, Saha K, De-Regil LM, Thuita F, Heidkamp R, Krasevec J, Hayashi C, Flores-Ayala R. Prevalence thresholds for wasting, overweight and stunting in children under 5 years. 2019. Public Health Nutr. 22(1):175-179.

[10]. Fadliana, A. & Drajat, P.P., 2021. Mapping of stunting risk factors based on geographic information systems using the Geographically Weighted Regression method. J IKRAITH INFORMATIKA. 5(3):91-102.

[11]. Fotheringham, A. S. (2002). Geographically weighted regression: the analysis of spatially varying relationships. John Wily and Sons Ltd.

[12]. Lesage JP and Pace RK. (2009). Introduction to Spatial Econometrics. Chapman and Hall/CRC.

[13]. Ministry of Health Indonesia. (2019). National Basic Health Research Report 2018. Kemenkes RI.

[14]. Ministry of Health Indonesia. (2021). National Basic Health Research Report 2021. Kemenkes RI.

[15]. Nakaya T, Fotheringham AS, Brunsdon C, and Charlton M. (2005). Geographically weighted Poisson regression for disease association mapping. Statistics in Medicine, 24(17), 2695-2717.

[16]. Pfeiffer, D. U., Robinson, T. P., Stevenson, M., Stevens, K. B., Rogers, D. J., & Clements, A. C.

2008. Spatial analysis in epidemiology (Vol. 142, No. 10.1093). Oxford: Oxford university press.

[17]. Statistics Indonesia. (2022). Profile of Indonesian Children 2022. BPS RI.

[18]. Sarkar, Debaprasad and Haldar, Sushil Kumar. 2014. Socioeconomic Determinants of Child Malnutrition in India: Evidence from NFHS-III. Economic & Political Weekly. 49(44):73-77.

[19]. Purwanti R and Nurfita D. 2019. Review literatur: Analysis of the socio-demographic determinants of stunting occurrences in young children across various developing countries. Health Research Bulletin. 47(3):153-164.

[20]. UNICEF. (2019). The State of the World's Children: Children, food and nutrition. United Nations Children's Fund.

[21]. UNICEF, WHO, & World Bank. (2021). Levels and Trends in Child Malnutrition. Geneva: World Health Organization.

[22]. Wamani H, Astrøm AN, Peterson S, Tumwine JK, Tylleskär T. Boys are more stunted than girls in Sub-Saharan Africa: A meta-analysis of 16 demographic and health surveys. BMC Pediatrics, 7(1): 17.

[23]. WHO. (2006). WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height, and body mass index-for-age: Methods and development. Geneva: World Health Organization.