BAREKENG: Journal of Mathematics and Its Applications
March 2024 Volume 18 Issue 1 Page 0383–0394
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss1pp0383-0394

ESTIMATION OF GEOGRAPHICALLY WEIGHTED PANEL REGRESSION MODEL WITH BISQUARE KERNEL WEIGHTING FUNCTION ON PERCENTAGE OF STUNTING TODDLERS IN INDONESIA

Asnita[1], Sifriyani[2*], Meirinda Fauziyah[3]

[1,2,3]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University
Barong Tongkok Street, Samarinda, 75123, Indonesia

*Corresponding author's e-mail: ****sifriyani@fmipa.unmul.ac.id***

Article History:

Received: 4th September 2023
Revised: 9th December 203
Accepted: 8th January 2024

Keywords:

GWPR;
Panel Regression;
Stunting;
Weighting Function.

ABSTRACT

Stunting is a condition of failure to thrive in children under five years old due to chronic malnutrition. Efforts that can be made to reduce the incidence of stunting in Indonesia include identifying factors that are thought to affect the incidence of stunting in Indonesia. The analysis methods used in this study are the global Fixed Effect Model (FEM) and the local Geographically Weighted Panel Regression (GWPR) model. FEM is a global regression model that assumes that each individual's model has a different intercept value. GWPR is a local regression model from FEM that considers aspects of geographic location by repeating data at each observation location, different times, and using spatial data. The weighting function used in this study is fixed bisquare and adaptive bisquare. This study aims to obtain a GWPR model on the percentage of stunting toddlers in Indonesia from 2019 until 2022 with independent variables, namely the percentage of children receiving exclusive breastfeeding (x_1), the percentage of households that have access to proper sanitation (x_2), the average per capita health expenditure of the population for a month (x_3), the average length of schooling for women (x_4), and the number of poor people (x_5). The variables are obtained from Statistics Indonesia (BPS) and Study of Indonesia's Nutritional Status (SSGI). The results showed that the best weighting function, namely adaptive bisquare with a CV value of 264.80.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.

How to cite this article:

Asnita, Sifriyani and M. Fauziyah., "ESTIMATION OF GEOGRAPHICALLY WEIGHTED PANEL REGRESSION MODEL WITH BISQUARE KERNEL WEIGHTING FUNCTION ON PERCENTAGE OF STUNTING TODDLERS IN INDONESIA," *BAREKENG: J. Math. & App.*, vol. 18, iss. 1, pp. 0383-0394, March, 2024.

Copyright © 2024 Author(s)
Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*
Research Article · Open Access

1. INTRODUCTION

The percentage of stunted children under five in Indonesia is known to change from time to time. Stunting is a condition of failure to thrive in children under five years old as a result of chronic malnutrition that can result in children being too short for their age. Reducing the percentage of stunting in toddlers has become one of the targets of 14% in 2024. Achieving these targets requires hard work from the government and various parties. A significant decrease compared to the percentage of stunting from Basic Health Research in 2018 of 30.80%. However, the results of the 2021 Indonesian Nutritional Status Study still show that the percentage of stunting in Indonesia is 24.41%. According to the World Health Organization (WHO), the number of stunted toddlers is high if the percentage reaches 20% or more, so it becomes an urgent health problem and must be addressed. Data in the health sector generally increases and decreases from year to year, so data in the health sector can be found in the form of panel data.

Panel data can be defined by data obtained from combining cross section data with time series data to obtain an overview of the unit over a period of time. Panel data can be modeled using panel regression. Panel data containing spatial heterogeneity cannot be modeled using panel regression because it is indicated that the parameters of the regression model are influenced by geographical factors. Therefore, a suitable model for spatial panel data is used, namely the Geographically Weighted Panel Regression (GWPR) model. It is a local model of the panel regression model that considers geographical aspects. GWPR model parameters can be estimated using the Weighted Least Square (WLS) approach [1]. The WLS method is a form of development of the Ordinary Least Square (OLS) method by considering spatial weighting at each observation location. Spatial weighting represents the magnitude of the influence of one data with other data [2]. This study uses the panel regression model and GWPR model to be applied to stunting toddler data in Indonesia.

Previous research on the application of spatial regression analysis to determine the factors that cause stunting in West Nusa Tenggara in 2021 [3], the results showed that the factors that affect stunting toddlers are the number of babies who receive Early Initiation of Breastfeeding and babies who receive exclusive breastfeeding. In another study on detecting the spatial influence on Fixed Effects Model (FEM) stunting prevalence [4], the results showed that stunting prevalence was caused by the average achievement of other provincial school years. Research on Geographically Weighted Regression (GWR) modeling on stunting cases in East Nusa Tenggara Province in 2020 [5], the results showed that the percentage of pregnant women at risk of Chronic Energy Deficiency, the percentage of babies receiving exclusive breastfeeding, the percentage of women who graduated from SMA/SMK/MA/Package C, and the percentage of women who had married underage are factors that affect stunting in East Nusa Tenggara Province.

Based on this background, researchers are interested in conducting a study entitled “Estimation of Geographically Weighted Panel Regression Model with Bisquare Kernel Weighting Function on Percentage Data of Stunting Toddlers in Indonesia.” The purpose of this study is to obtain the GWPR model and what factors affect the percentage of stunting toddlers in Indonesia based on the GWPR model.

2. RESEARCH METHODS

2.1 Data and Data Sources

Data and data sources are described in Table 1.

Table 1. Data and Data Sources

Variable	Symbol	Name of Variables	Data Sources	Unit
Dependent	y	Percentage of Stunting Toddlers	Study of Indonesia’s Nutritional Status (SSGI)	Percent (%)
Independent	x_1	Percentage of Children Receiving Exclusive Breastfeeding	Statistics Indonesia (BPS)	Percent (%)
	x_2	Percentage of Households that Have Access to Proper Sanitation	Statistics Indonesia (BPS)	Percent (%)

Variable	Symbol	Name of Variables	Data Sources	Unit
	x_3	Average per Capita Health Expenditure of the Population During the Month	Statistics Indonesia (BPS)	Rupiah (Rp)
	x_4	Average Length of Schooling for Women	Statistics Indonesia (BPS)	Year
	x_5	Number of Poor People	Statistics Indonesia (BPS)	People

2.2 Regression Analysis

Regression analysis is an analysis that measures the relationship between one dependent variable and one or more independent variables with the aim of estimating the value of the dependent variable based on the value of the known independent variable. Regression analysis has mathematical equations that make it possible to perform an estimate of parameter values [6].

Linear regression models can be obtained by estimating the parameters. Linear regression models are divided into simple and multiple linear regression models. A simple linear regression model is a linear regression model where there is one dependent variable and one independent variable. In contrast, multiple linear regression models have one dependent variable and two or more independent variables. A regression model is good if it satisfies the classical assumption that no autocorrelation occurs, no heteroscedasticity occurs, and residual models are normally distributed. The multicollinearity test was carried out on multiple linear regression analysis [7].

2.3 Multicollinearity Detection

Multicollinearity is the existence of linear relationships between independent variables in a multiple linear regression model. Multicollinearity can be detected by looking at the value of the Variance Inflation Factor (VIF) and is said to be detected when the value of VIF is greater than 10. Multicollinearity detection can be done by progressing each independent variable with another independent variable. The value of VIF can be calculated using **Equation (1)**.

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

where VIF_k is the value of the VIF of the k independent variable and R_k^2 is the k coefficient of determination **[8]**.

2.4 Panel Regression Model

Panel regression models are divided into Common Effect Model (CEM), Random Effect Model (REM), and Fixed Effect Model (FEM). FEM is a linear regression model that assumes that each individual's model has a different intercept value. The general form of FEM is given in **Equation (2)**.

$$y_{it} = \beta_{0_i} + \beta_1 x_{it1} + \beta_2 x_{it2} + \cdots + \beta_k x_{itk} + \varepsilon_{it},\ i = 1,2,\ldots,n;\ t = 1,2,\ldots,T \tag{2}$$

β_{0_i} in **Equation (2)** shows that the interception of each cross section unit is different. Estimation on FEM can be done by transforming β_{0_i} through the within estimator method. The within estimator method begins by forming an average model that is based on **Equation (2)** for each $t = 1,2,\ldots,T$ to get the given cross section equation in **Equation (3)**.

$$\bar{y}_i = \beta_{0_i} + \beta_1 \bar{x}_{i1} + \beta_2 \bar{x}_{i2} + \cdots + \beta_k \bar{x}_{ik} + \bar{\varepsilon}_i;\ i = 1,2,\ldots,n \tag{3}$$

where

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \quad \bar{x}_{ik} = \frac{1}{T}\sum_{t=1}^{T} x_{itk}, \quad \bar{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it}\ ; k = 1,2,\ldots,p \tag{4}$$

If **Equation (2)** is subtracted by **Equation (3)**, FEM is obtained with within estimator given in **Equation (5)**.

$$\begin{gathered}(y_{it} - \bar{y}_i) = \left(\beta_{0_i} - \beta_{0_i}\right) + \beta_1(x_{it1} - \bar{x}_{i1}) + \beta_2(x_{it2} - \bar{x}_{i2}) + \cdots + \\ \beta_k(x_{itk} - \bar{x}_{ik}) + (\varepsilon_{it} - \bar{\varepsilon}_i)\end{gathered} \tag{5}$$

or

$$y_{it}^* = \beta_1 x_{it1}^* + \beta_2 x_{it2}^* + \cdots + \beta_k x_{itk}^* + \varepsilon_{it}^* \,; t = 1,2,\dots,T, \tag{6}$$

with

$$y_{it}^* = (y_{it} - \bar{y}_i), x_{itk}^* = (x_{itk} - \bar{x}_{ik}), \text{dan } \varepsilon_{it}^* = (\varepsilon_{it} - \bar{\varepsilon}_i) \tag{7}$$

The data resulting from the transformation in **Equation (7)** is called demean data **[9]**.

2.5 FEM Parameter Significance Test

1) Simultaneous Test

Testing the significance of parameters in FEM is simultaneously carried out to determine how the effect of the independent variable simultaneously on the dependent variable. The hypotheses on simultaneous significance testing are as follows.

H_0: $\beta_1 = \beta_2 = \cdots = \beta_k = 0$

(Simultaneously the independent variable has no effect on the dependent variable)

H_1: There is at least one of $\beta_k \neq 0, k = 1,2,3,\dots,p$

(There is at least one independent variable that affects the dependent variable)

Test statistics testing the significance of parameters on FEM are simultaneously given in **Equation (8)**

$$F_{FEM} = \frac{KTR}{KTG} \tag{8}$$

F_{FEM} value obtained from the Regression Middle Square (KTR) value divided by the Error Middle Square (KTG) value. F_{FEM} test statistics distributed $F_{(p\,;\,nT-n-p)}$ with critical areas H_0 rejected at the level of significance of α if $F_{FEM} > F_{(\alpha\,;\,p\,;\,nT-n-p)}$ or $p_{value} < \alpha$ **[10]**.

2) Partial Test

Testing the significance of parameters in FEM is partially carried out to determine how each independent variable affects the dependent variable. The hypotheses on partial significance testing are as follows.

H_0: $\beta_k = 0, k = 1,2,3,\dots,p$

(There is no effect of the independent variable on the dependent variable)

H_1: $\beta_k \neq 0, k = 1,2,3,\dots,p$

(There is has effect of the independent variable on the dependent variable)

Test statistics testing the significance of parameters on FEM are partially given in **Equation (9)**.

$$T_{FEM} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \tag{9}$$

with $SE(\hat{\beta}_k)$ is the standard deviation of $\hat{\beta}_k$. Critical areas of partial testing where H_0 rejected at the level of significance α if $|T_{FEM}| \geq t_{(\frac{\alpha}{2}\,;\,nT-p)}$ or $p_{value} < \alpha$ **[10]**.

2.6 Homoscedasticity Test

The variance of the error in the panel regression model is assumed to be constant because if the error variance is not constant it will cause the estimation of the parameters of the model obtained to be inefficient so that the conclusions obtained do not represent the actual situation. Homoscedasticity in regression models can be tested using the Glejser Test. The hypotheses are as follows.

H_0: $\sigma_{1,1}^2 = \sigma_{2,1}^2 = \cdots = \sigma_{n,T}^2 = \sigma^2$

(Error variance is constant)

H_1: At least one $\sigma_{i,t}^2 \neq \sigma^2, i = 1,2, \dots, n\,; t = 1,2, \dots, T$

(Error variance is not constant)

The statistics of the Glejser Test are given in **Equation (10)**.

$$F_{Glejser} = \frac{\left(\widehat{\boldsymbol{\varphi}}^{\mathbf{T}}\mathbf{X}^{*\mathbf{T}}\boldsymbol{\varepsilon}^{*} - n(\bar{\varepsilon}^{*})^2\right)\Big/p}{\left(\boldsymbol{\varepsilon}^{*\mathbf{T}}\boldsymbol{\varepsilon}^{*} - \widehat{\boldsymbol{\varphi}}^{\mathbf{T}}\mathbf{X}^{*\mathbf{T}}\boldsymbol{\varepsilon}^{*}\right)\Big/(nT-n-p)} \tag{10}$$

The statistics of the $F_{Glejser}$ test follow the distribution of $F_{(p\,;\,nT-n-p)}$ where n is the number of observation locations, T the number of observation times and p the number of independent variables. The critical area of the Glejser Test where H_0 is denied at the level of significance α if $F_{Glejser} > F_{(\alpha\,;\,p\,;\,nT-n-p)}$ or if $p_{value} < \alpha$ **[11]**.

2.7 Spatial Weighting Function

There are several methods that can be used to calculate spatial weighting, including kernel functions that are divided into fixed kernel functions and adaptive kernel functions **[12]**,**[13]**. The fixed kernel function generates a constant bandwidth value for each observation location. In contrast, the kernel adaptive function produces different bandwidth values for each observation location.

Fixed bisquare kernel functions are given in **Equation (11)**.

$$w_{ij} = \begin{cases} \left(1 - \left(\dfrac{d_{ij}}{b}\right)^2\right)^2 & \text{, if } d_{ij} \leq b \\ 0 & \text{, for other } d_{ij} \end{cases} \tag{11}$$

The adaptive bisquare kernel function is given in **Equation (12)**.

$$w_{ij} = \begin{cases} \left(1 - \left(\dfrac{d_{ij}}{b_i}\right)^2\right)^2 & \text{, if } d_{ij} \leq b_i \\ 0 & \text{, for other } d_{ij} \end{cases} \tag{12}$$

where w_{ij} is a weighting function between location *i* and location *j*, and d_{ij} is the distance between location *i* and location *j* obtained from the Euclidean distance that can be calculated using **Equation (13)**.

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{13}$$

One way that can be done to select the optimum bandwidth is to use the Cross Validation (CV) method given in **Equation (14)**.

$$CV = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(b)]^2 \tag{14}$$

with $\hat{y}_{\neq i}(b)$ is the estimated value of y_i where the observation of the *i* location is omitted from the estimation process **[14]**.

2.8 Geographically Weighted Panel Regression (GWPR) Model

GWPR is a modification of the regression model which is a combination of GWR and panel data. The GWPR model is a local regression model of FEM, with repeating data at each observation location, different times, and spatial data **[15]**. The coordinates at each observation location are known with the coordinates of the *i* observation location being (u_i, v_i) where u_i state the location of latitude and v_i state the location of longitude. Based on FEM with within estimator, GWPR models at *i* and *t-time* observation locations **[16]**, are given in **Equation (15)**.

$$\begin{aligned} y_{it}^{*} = \; & \beta_1(u_i, v_i)x_{it1}^{*} + \beta_2(u_i, v_i)x_{it2}^{*} + \cdots + \\ & \beta_k(u_i, v_i)x_{itk}^{*} + \varepsilon_{it}^{*},\; i = 1,2, \dots, n;\; t = 1,2, \dots, T \end{aligned} \tag{15}$$

2.9 Estimation of Geographically Weighted Panel Regression Model

Estimation of GWPR model parameters can be done using the Weighted Least Square (WLS) approach, which is a form of developing the Ordinary Least Square (OLS) method by considering spatial weighting at each observation location. Based on the WLS method [17], the GWPR model parameter estimator is obtained by minimizing the sum of squares errors of **Equation (16)** with spatial weighting so that **Equation (16)** is obtained.

$$\sum_{t=1}^{T}\sum_{i=1}^{n} w_{it}(u_i,v_i)\varepsilon_{it}^2 = \sum_{t=1}^{T}\sum_{i=1}^{n} w_{it}(u_i,v_i)[y_{it}^* - \beta_1(u_i,v_i)x_{111}^* + \beta_2(u_i,v_i)x_{212}^* + \cdots + \beta_k(u_i,v_i)x_{itk}^*] \tag{16}$$

Equation (16) can be denoted in matrix form so that **Equation (17)** is obtained.

$$\begin{aligned}\boldsymbol{\varepsilon}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\boldsymbol{\varepsilon}^* &= [\mathbf{y}^* - \mathbf{X}^*\boldsymbol{\beta}(u_i,v_i)]^{\mathrm{T}}\mathbf{W}(u_i,v_i)[\mathbf{y}^* - \mathbf{X}^*\boldsymbol{\beta}(u_i,v_i)] \\ &= \mathbf{y}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* - \mathbf{y}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{X}^*\boldsymbol{\beta}(u_i,v_i) \\ &\quad - \boldsymbol{\beta}^{\mathrm{T}}(u_i,v_i)\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* + \boldsymbol{\beta}^{\mathrm{T}}(u_i,v_i)\mathbf{X}^{*\mathrm{T}} \\ &\quad \mathbf{W}(u_i,v_i)\mathbf{X}^*\boldsymbol{\beta}(u_i,v_i) \qquad (17) \\ &= \mathbf{y}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* - 2\boldsymbol{\beta}^{\mathrm{T}}(u_i,v_i)\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* + \\ &\quad \boldsymbol{\beta}^{\mathrm{T}}(u_i,v_i)\mathbf{X}^{*\mathrm{T}}\,\mathbf{W}(u_i,v_i)\mathbf{X}^*\boldsymbol{\beta}(u_i,v_i) \qquad (18)\end{aligned}$$

with

$$\boldsymbol{\beta}(u_i,v_i) = [\beta_1(u_i,v_i)\ \ \beta_2(u_i,v_i)\ \ \ldots\ \ \beta_p(u_i,v_i)]^T$$

and

$$\mathbf{W}(u_i,v_i) = diag[w_{i11}, w_{i21}, \ldots, w_{in1}, w_{i12}, w_{i22}, \ldots, w_{in2}, \ldots, w_{i1T}, w_{i2T}, \ldots, w_{inT}]$$

A decrease in the first-order **Equation (18)** to β^{T} $\boldsymbol{\beta}^{\mathrm{T}}(u_i,v_i)$ and equated to zero will yield the following equations.

$$-2\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* + 2\mathbf{X}^{*\mathrm{T}}\,\mathbf{W}(u_i,v_i)\mathbf{X}^*\boldsymbol{\beta}(u_i,v_i) = 0 \tag{19}$$

$$2\mathbf{X}^{*\mathrm{T}}\,\mathbf{W}(u_i,v_i)\mathbf{X}^*\boldsymbol{\beta}(u_i,v_i) = 2\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* \tag{20}$$

$$\mathbf{X}^{*\mathrm{T}}\,\mathbf{W}(u_i,v_i)\mathbf{X}^*\boldsymbol{\beta}(u_i,v_i) = \mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* \tag{21}$$

The two fields in **Equation (21)** are multiplied by $(\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{X}^*)^{-\mathbf{1}}$ to obtain $\widehat{\boldsymbol{\beta}}^*(u_i,v_i)$.

$$(\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{X}^*)^{-\mathbf{1}}\mathbf{X}^{*\mathrm{T}}\,\mathbf{W}(u_i,v_i)\mathbf{X}^*\boldsymbol{\beta}(u_i,v_i) = (\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{X}^*)^{-\mathbf{1}}\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{y}^* \tag{22}$$

So that **Equation (23)** is obtained.

$$\widehat{\boldsymbol{\beta}}^*(u_i,v_i) = (\mathbf{X}^{*\mathrm{T}}\mathbf{W}(u_i,v_i)\mathbf{X}^*)^{-\mathbf{1}}\mathbf{X}^{*\mathrm{T}}\,\mathbf{W}(u_i,v_i)\mathbf{y}^* \tag{23}$$

with $\widehat{\boldsymbol{\beta}}^*(u_i,v_i)$ GWPR model parameter estimator and

$\mathbf{X}_{it}^{*T} = \begin{bmatrix} x_{i11}^* & x_{i12}^* & \cdots & x_{i1p}^* \\ x_{i21}^* & x_{i22}^* & \cdots & x_{i2p}^* \\ \vdots & \vdots & \ddots & \vdots \\ x_{iT1}^* & x_{iT2}^* & \cdots & x_{iTp}^* \end{bmatrix}$ is the matrix for the i observation of each unit of time of the $\mathbf{X}^*$ [18].

2.10 Stunting

Stunting is a disorder of growth and development of children due to chronic malnutrition in toddlers (infants under the age of 5 years). Poor health levels from an early age will result in an increased tendency to non-communicable diseases in adulthood [19]. Malnutrition, with its various forms, especially stunting in early childhood, is a threat to child development that occurs in various countries. In 2016, about 23 percent of children under the age of five were stunted globally. Stunting refers to the condition of a child who is too short for his age. This condition describes chronic malnutrition that occurs due to poor nutritional quality since in the womb, in childhood, and/or caused by infection or disease [20].

3. RESULTS AND DISCUSSION

3.1 Data Description

The description of the research variables consists of the average, minimum, and maximum values. The calculation results are presented in **Table 2**.

Table 2. Description of Research Data

Variable	Year	Average	Minimum	Maximum
Percentage of Stunting Toddlers (y)	2019	27.77	14.30	43.70
	2020	26.80	13.68	42.99
	2021	25.21	10.90	37.80
	2022	23.29	8.00	35.30
Percentage of Children Receiving Exclusive Breastfeeding (x_1)	2019	64.37	39.64	79.05
	2020	66.49	52.98	78.93
	2021	68.88	52.75	81.46
	2022	69.21	53.60	79.69
Percentage of Households that Have Access to Proper Sanitation (x_2)	2019	77.34	38.27	94.67
	2020	79.81	40.31	96.96
	2021	80.97	40.81	97.12
	2022	81.00	40.34	96.21
Average per Capita Health Expenditure of the Population During a Month (x_3)	2019	26,267.00	9,849.00	50,306.00
	2020	27,414.00	10,290.00	57,065.00
	2021	29,523.00	10,813.00	76,550.00
	2022	27,468.00	10,135.00	73,191.00
Average Length of Schooling for Women (x_4)	2019	8.17	5.79	10.65
	2020	8.31	5.88	10.76
	2021	8.40	6.00	10.83
	2022	8.58	6.29	10.97
Number of Poor People (x_5)	2019	729.00	48.61	4,056.00
	2020	810.30	52.70	4,586.00
	2021	779.52	49.49	4,259.60
	2022	775.39	50.58	4,236.51

Based on **Table 2**, for the dependent variable (y), it is known that the average percentage of stunting toddlers in 34 provinces in Indonesia in 2019 was 27.77%. The lowest percentage of stunting toddlers occurred in Bali Province at 14.30% and the highest occurred in East Nusa Tenggara Province at 43.70%. The average percentage of stunting toddlers in 34 provinces in Indonesia in 2020 decreased from 2019 to 26.80%. The lowest percentage of stunting toddlers occurred in Bali Province at 13.68% and the highest occurred in East Nusa Tenggara Province at 42.99%. The average percentage of stunting toddlers in 34 provinces in Indonesia in 2021 decreased from 2020 to 25.21%. The lowest percentage of stunting toddlers occurred in Bali Province at 10.90% and the highest occurred in East Nusa Tenggara Province at 37.80%. The average percentage of stunting toddlers in 34 provinces in Indonesia in 2022 has decreased from 2021 to 23.29%. The lowest percentage of stunting toddlers occurred in Bali Province at 8.00% and the highest occurred in East Nusa Tenggara Province at 35.30%.

3.2 Multicollinearity Detection

Multicollinearity detection is done by looking at the value of VIF calculated using **Equation (1)**. The results of the calculation of the VIF value can be seen in **Table 3**.

Table 3. Independent Variable VIF value

x_k	VIF_k	Detection Results
x_1	1.2350	No multicollinearity
x_2	1.7431	No multicollinearity
x_3	1.9074	No multicollinearity
x_4	1.6956	No multicollinearity
x_5	1.2656	No multicollinearity

Based on Table 3, the VIF value of each independent variable is less than 10, so the result is that there is no multicollinearity between independent variables in the regression model. Because there is no multicollinearity between independent variables, it can proceed to the next step of analysis.

3.3 Panel Regression Test

The panel regression model is divided into CEM, REM, and FEM. From the three panel regression models, the best panel regression model will be selected through two tests. The first test is to compare between CEM and FEM using the Chow test. Then the second test compares FEM with REM using the Hausman test. The results of the two tests are presented inTable 4.

Table 4. Panel Regression Test

Testing	*p-value*	Selected Model
Chow Test	2.2×10^{-16}	FEM
Hausman Test	0.00048	FEM

Based on Table 4, it can be seen that the Chow test and Hausman test have *p-value* < 0.05 so that the best panel regression model used is FEM.

3.4 FEM Panel Regression

A general FEM with within estimators for stunting toddler percentage data with 5 independent variables is given to Equation (24).

$$y_{it}^* = \beta_1 x_{it1}^* + \beta_2 x_{it2}^* + \beta_3 x_{it3}^* + \beta_4 x_{it4}^* + \beta_5 x_{it5}^* + \varepsilon_{it}^* ;\quad i = 1,2, \dots 34 ; t = 1,2,3,4 \tag{24}$$

The estimation of FEM parameters in Equation (24) based on Equation (15) and the estimated results can be seen in Table 5.

Table 5. Parameter Estimation Value in FEM

Parameter	Estimation Value
β_1	-0.0556
β_2	-0.4645
β_3	-0.00004
β_4	-5.9649
β_5	0.0047

Based on the estimated parameter values in Table 5, the FEM formed is

$$\hat{y}_{it}^* = -0.0556 x_{it1}^* - 0.4645 x_{it2}^* - 0.00004 x_{it3}^* - 5.9649 x_{it4}^* + 0.0047 x_{it5}^* ;\quad i = 1,2, \dots 34 ; t = 1,2,3,4 \tag{25}$$

Furthermore, simultaneous testing of the significance of FEM parameters was carried out to determine the effect of the independent variable on the bound variable simultaneously. The hypotheses of testing the significance of FEM parameters simultaneously is as follows.

H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$
(Simultaneously, the variables Percentage of Children Receiving Exclusive Breastfeeding, Percentage of Households that have Access to Proper Sanitation, Average per Capita Health Expenditure of the Population during a Month, Average Length of Schooling for Women and Number of Poor People do not affect the Percentage of Stunting Toddlers in 34 provinces in Indonesia)

H_1: At least one of $\beta_k \neq 0, k = 1,2, \dots ,5$
(Simultaneously, there is at least one variable from the Percentage of Children Receiving Exclusive Breastfeeding, Percentage of Households that have Access to Proper Sanitation, Average per Capita Health Expenditure of the Population during a Month, Average Length of Schooling for Women, and Number of Poor People affect the Percentage of Stunting Toddlers in 34 provinces in Indonesia)

The statistical values of F_{FEM} and p_{value} tests can be seen in **Table 6**.

Table 6. FEM Simultaneous Test

F_{FEM}	$F_{(0,05;5;97)}$	p_{value}	Decision
19.6597	2.3082	0.17651×10^{-13}	H_0 rejected

Based on **Table 6**, it is concluded that simultaneously, there is at least one variable from the Percentage of Children Receiving Exclusive Breastfeeding, Percentage of Households that have Access to Proper Sanitation, Average per Capita Health Expenditure of the Population during a Month, Average Length of Schooling for Women and Number of Poor People affect the Percentage of Stunting Toddlers in 34 provinces in Indonesia.

Then a partial test of the significance of FEM parameters was carried out. The hypothesis of testing the significance of FEM parameters partially is as follows.

H_0: $\beta_k = 0 ; k = 1,2, \dots 5$
(There is no effect of x_k variables on the Percentage of Stunting Toddlers in 34 provinces in Indonesia)

H_1: $\beta_k \neq 0, k = 1,2, \dots ,5$
(There is effect at least one of x_k variables on the Percentage of Stunting Toddlers in 34 provinces in Indonesia)

The results of partial parameter significance test calculations can be seen in **Table 7**.

Table 7. FEM Partial Test

x_k	$\lvert T_{FEM}\rvert$	P_{value}	Decision
x_1	0.8057	0.4224	H_0 accepted
x_2	3.6604	0.0004	H_0 rejected
x_3	0.7361	0.4634	H_0 accepted
x_4	3.9841	0.0001	H_0 rejected
x_5	1.8936	0.0613	H_0 accepted

Based on **Table 7**, H_0 rejected with a significance level of 0.05 was obtained for the variables Percentage of Households that Have Access to Adequate Sanitation and Average Length of Schooling for Women as indicated by test statistical scores $|T_{FEM}|$ from each of these variables is greater than the value of $t_{(0.025;131)} = 2.2675$ and the $p_{value} < \alpha = 0.05$ so that it can be concluded that the variables Percentage of Households that have access to Proper Sanitation and Average Length of Schooling for Women affect the Percentage of Stunting Toddlers in 34 provinces in Indonesia.

3.5 FEM Homoscedasticity Test

A homoscedasticity test is performed to determine whether the error variance of the entire observation site is constant. Homoscedasticity uses the Glejser test with the following hypotheses.

H_0: $\sigma_{1,1}^2 = \sigma_{2,1}^2 = \cdots = \sigma_{n,T}^2 = \sigma^2$
(Error variance is constant)

H_1: At least one $\sigma_{i,t}^2 \neq \sigma^2, i = 1,2, \dots , n ; t = 1,2, \dots , T$
(Error variance is not constant)

The calculation results of FEM homoscedasticity test based on R software output can be seen in **Table 8**.

Table 8. FEM Glejser Test

$F_{Glejser}$	$F_{(0.05;5;97)}$	P_{value}	Decision
2.7329	2.2675	0.0236	H_0 rejected

Based on **Table 8**, decided H_0 rejected with a significance level of 0.05 is obtained as indicated by the statistical value of the test $F_{Glejser} = 2.7329 > F_{(0,05;5;97)} = 2.2675$ and the $p_{value} = 0.0236 < \alpha = 0.05$. It was concluded that the error variance is not constant across observation locations. Thus, the assumption of

FEM homoscedasticity is not fulfilled due to spatial effects on the model. This makes the resulting FEM less suitable for modeling the Percentage of Stunting Toddlers in 34 provinces in Indonesia so that it can use a local model, namely GWPR.

3.6 Spatial Weighting Function

Based **Equation (13)**, the CV for kernel fixed bisquare and kernel adaptive bisquare can be seen in **Table 9**.

Table 9. Cross Validation (CV) Values

Weighting Function	CV
Fixed Bisquare	365.42
Adaptive Bisquare	264.80

Based on the CV values in **Table 9**, it can be concluded that the best model is with the kernel adaptive bisquare function with a CV value is 264.80.

GWPR 3.7Model

The general model of GWPR at the *i* location at the *t* time for data on the Percentage of Stunting Toddlers with 5 independent variables are given in the **Equation (26)**.

$$y_{it}^{*} = \beta_1(u_i,v_i)x_{it1}^{*} + \beta_2(u_i,v_i)x_{it2}^{*} + \beta_3(u_i,v_i)x_{it3}^{*} + \beta_4(u_i,v_i)x_{it4}^{*} + \beta_{5(u_i,v_i)}x_{it5}^{*} + \varepsilon_{it}^{*}\ ; \quad i = 1,2,\dots 34\ ; t = 1,2,3,4 \tag{26}$$

The process of estimating parameters begins with calculating Euclidean distances. The determination of the optimum bandwidth with the best spatial weighting function is adaptive bisquare. The GWPR model formed as many as 34 models with one of the estimated GWPR model in East Kalimantan Province is

$$\hat{y}_{23t}^{*} = -0.0344x_{25t1}^{*} - 0.4425x_{25t2}^{*} - 0.00004x_{25t3}^{*} - 6.7947x_{25t4}^{*} + 0.0047x_{25t5}^{*}\ ; \ t = 1,2,3,4 \tag{27}$$

Based on the GWPR model in **Equation (27)**, the variables obtained are the Percentage of Households that Have Access to Proper Sanitation, the Average Length of Schooling for Women, and the Number of Poor People that affect the Percentage of Stunting Toddlers. The regression coefficient for the variable Percentage of households that have access to Proper Sanitation in East Kalimantan Province of -0.4425 states that every 1% increase in households that have access to Proper Sanitation will reduce the Percentage of Stunting Toddlers in East Kalimantan Province by 0.4425. The regression coefficient for the variable Average Length of Schooling for Women in East Kalimantan Province of -6.7947 states that every 1-year increase in the Average Length of Schooling for Women will reduce the Percentage of Stunting Toddlers in East Kalimantan Province by 6.7947. The regression coefficient for the variable Number of Poor People in East Kalimantan Province of 0.0047 states that every increase of 1 poor person will increase the Percentage of Stunting Toddlers in East Kalimantan Province by 0.0047.

Provincial grouping based on the factors that influence the Percentage of Stunting Toddlers in Indonesia can be illustrated through the distribution map presented in **Figure 1**.

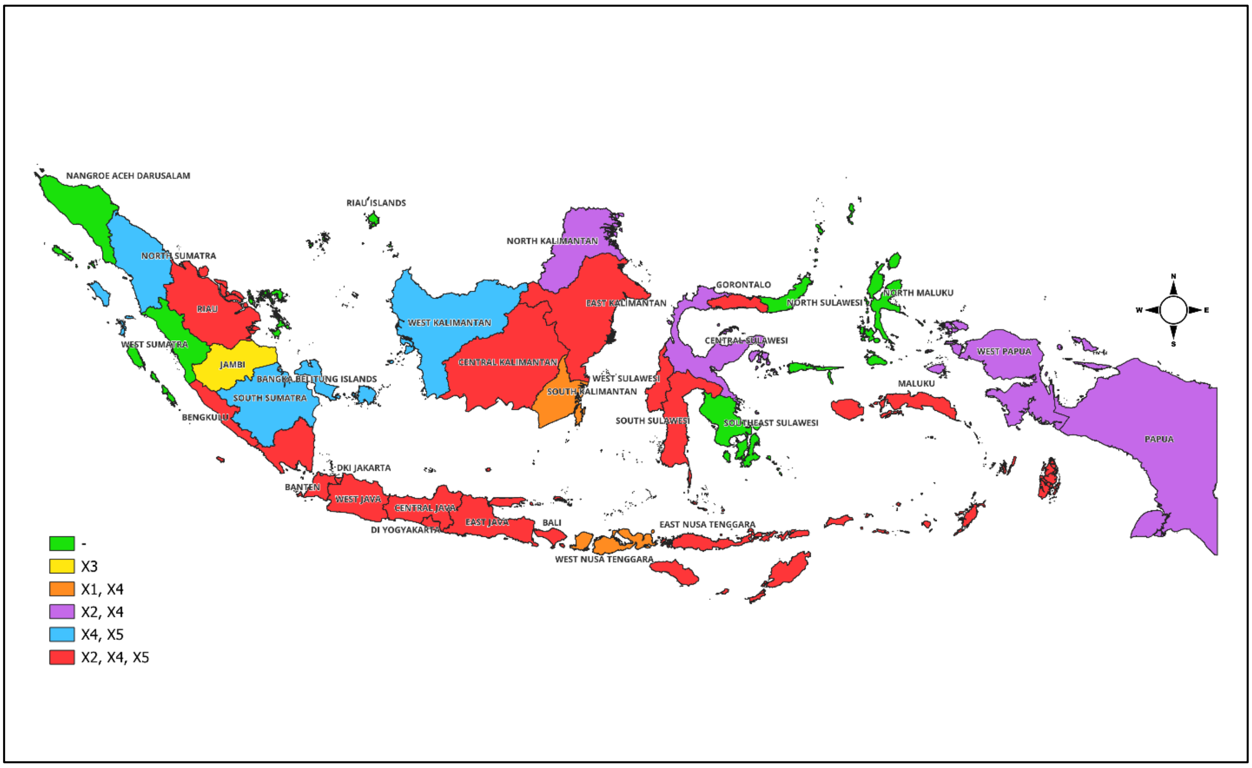

Figure 1. Provincial Grouping Based on the Factors that Influence the Percentage of Stunting Toddlers in Indonesia

Based on **Figure 1**, it can be seen that the Percentage of Households that Have Access to Proper Sanitation, the Average Length of Schooling for Women, and the Number of Poor People affect the Percentage of Stunting Toddlers the most provinces in Indonesia, which are indicated by the red color.

4. CONCLUSIONS

The conclusions obtained from this study are as follows.

1. Based on the GWPR model, obtained 34 models of the Percentage of Stunting Toddlers with $t = 1,2,3,4$ for each t are 2019, 2020, 2021, and 2022 with one of the estimated GWPR model in East Kalimantan Province is
$$\hat{y}^*_{23t} = -0.0344x^*_{25t1} - 0.4425x^*_{25t2} - 0.00004x^*_{25t3} - 6.7947x^*_{25t4} + 0.0047x^*_{25t5}$$
2. The factors that affect the Percentage of Stunting Toddlers in Indonesia based on the GWPR model divided into 6 groups
   - First group, there is no factor that affect the Percentage of Stunting Toddlers.
   - Second group, the Average per Capita Health Expenditure of the Population during a Month (x_3) affect the Percentage of Stunting Toddlers.
   - Third group, the Percentage of Children Receiving Exclusive Breastfeeding (x_1) and the Average Length of Schooling for Women (x_4) affect the Percentage of Stunting Toddlers.
   - Fourth group, the Percentage of Households that have Access to Proper Sanitation (x_2) and the Average Length of Schooling for Women (x_4) affect the Percentage of Stunting Toddlers.
   - Fifth group, the Average Length of Schooling for Women (x_4) and the Number of Poor People (x_5) affect the Percentage of Stunting Toddlers.
   - Sixth group, the Percentage of Households that have Access to Proper Sanitation (x_2), Average Length of Schooling for Women (x_4) and the Number of Poor People (x_5) affect the Percentage of Stunting Toddlers.

REFERENCES

[1] Sifriyani, *Kapita Selekta Statistika Spasial Spatio-Temporal Models*. Indonesia: Mulawarman University Press, 2023.

[2] Z. Mar'ah, and Sifriyani, "GEOGRAPHICALLY WEIGHTED PANEL REGRESSION (GWPR) FOR COVID-19 CASE IN INDONESIA," *BAREKENG*: *J. Math. & App.*, vol. 17, iss. 2, pp. 0879–0886, June, 2023.

[3] V. S. Puspita and T. S. Yanti, "Penerapan Analisis Regresi Spasial untuk Menentukan Faktor-Faktor Penyebab Stunting di Nusa Tenggara Barat Tahun 2021," *Bandung Conference Series: Statistics*, vol. 2, no. 2, pp. 245–253, 2022.

[4] A. Murwiati, "Deteksi Spatial Dependence terhadap Fixed Effects Model Prevalensi Stunting," *J. Ekon. Pembang. Unila*, vol. 1, no. 1, pp. 21–30, 2021.

[5] M. G. L. Bele, E. M. P. Hermanto and F. Fitriani, "Pemodelan Geographically Weighted Regression pada Kasus Stunting di Provinsi Nusa Tenggara Timur Tahun 2020," *J. Stat. dan Apl.*, vol. 6, no. 2, pp. 179–191, 2022.

[6] A. Widarjono, *Ekonometrika Teori dan Aplikasi*. Yogyakarta: Ekonisia, 2007.

[7] D. Rosadi, *Analisis Ekonometrika dan Runtun Waktu Terapan dengan R*. Yogyakarta: Andi, 2011.

[8] D. N. Gujarati, *Ekonometri Dasar. Terjemahan: Sumarno Zain*. Jakarta: Erlangga, 2003.

[9] J. M. Wooldridge, *Econometrics Analysis of Cross-Section and Panel Data*. London: MIT Press, 2002.

[10] Setiawan and D.E. Kusrini, *Ekonometrika*. Yogyakarta: Andi, 2010.

[11] D. N. Gujarati, *Ekonometika Dasar*. Jakarta: Erlangga, 2006.

[12] Sifriyani *et al.*, "A Spatio-Temporal Description of Covid-19 Cases in East Borneo using Improved Geographically and Temporally Weighted Regression (I-GTWR)," *Commun. Math. Biol. Neurosci*., vol. 2022, pp. 1–22, 2022.

[13] Sifriyani, M. Rasjid, D. Rosadi, S. Anwar, R. D. Wahyuni, and S. Jalaluddin, "Spatial-Temporal Epidemiology of COVID-19 Using a Geographically and Temporally Weighted Regression Model," *Symmetry (Basel).*, vol. 14, no. 4, pp. 742, 2022, doi: https://doi.org/10.3390/sym14040742.

[14] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley, 2002.

[15] S. Sifriyani, I. Mandang, F. D. T. Amijaya, and R. Ruslan, "Developing Geographically Weighted Panel Regression Model for Spatio-Temporal Analysis Of Covid-19 Positive Cases In Kalimantan, Indonesia," *J. Southwest Jiaotong Univ.*, vol. 57, no. 3, 2022.

[16] F. Bruna and D. Yu, "Geographically Weighted Panel Regression," *XI Galician Congr. Stat. Oper. Res.*, pp. 1–23, 2013.

[17] S. Sifriyani, I. Mandang, and F. D. T. Amijaya, "Geographically temporally weighted regression model for GIS mapping of influence COVID-19 in East Kalimantan," *in AIP Conf. Proc.*, vol. 2668, pp. 070012, 2022.

[18] R. Cai, D. Yu, and M. Oppenheimer, "Estimating the Spatially Varying Responses of Corn Yields to Weather Variations Using Geographically Weighted Panel Regression," *J. Agric. Resour. Econ.*, vol. 39, no. 2, 2014.

[19] H. A. D. Boo and J. E. Harding, "The developmental origins of adult disease," *Aust. New Zeal. J. Od Obs. Gyneacology*, vol. 46, pp. 4–14, 2006.

[20] Badan Pusat Statistik (BPS), *Profil Kesehatan Ibu dan Anak*. Jakarta: Badan Pusat Statistik, 2022.