Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 14(1) 2021: 10-20
http://ejournal.undip.ac.id/index.php/media_statistika

GEOGRAPHICALLY WEIGHTED PANEL REGRESSION WITH FIXED EFFECT FOR MODELING THE NUMBER OF INFANT MORTALITY IN CENTRAL JAVA, INDONESIA

Agus Rusgiyono, Alan Prahutama
Department of Statistics, Faculty of Science and Mathematics, Diponegoro University

e-mail: agus.rusgi@gmail.com

DOI: 10.14710/medstat.14.1.10-20

Article Info:
Received: 19 November 2019
Accepted: 28 April 2021
Available Online: 30 June 2021

Keywords:
Geographically Weighted Panel Regression; Fix Effect; Infant Mortality
.

Abstract: One of the regression methods used to model by region is Geographically Weighted Regression (GWR). The GWR model developed to model panel data is Geographically Weighted Panel Regression (GWPR). Panel data has several advantages compared to cross-section or time-series data. The development of the GWPR model in this study uses the Fixed Effect model. It is used to model the number of infant mortality in Central Java. In this study, the weighting used by the fixed bisquare kernel resulted in a significant variable percentage of clean and healthy households. The value of R-square is 67.6%. Also in this paper completed by spread map base on GWPR model.

1. INTRODUCTION

Sustainable Development Goal's (SDG's) contains 17 goals and 169 development goals expected to answer the underdevelopment of countries worldwide, and both developed and developing countries. The health sector in SDG's includes no hunger, good health, gender equality, clean water and sanitation. In the aim of good health, it means that it guarantees a healthy life and promotes well-being for all people of all ages. One of the targets is to end infant and toddler deaths that can be prevented by reducing the neonatal mortality rate to 12 per 1000 Live Birth (LB) and under-five mortality rates by 25 per 1000 Live Birth (LB) (Prahutama et al., 2018).

Infant Mortality Rate (IMR) is the number of deaths of infants under one year of age per 1000 live births in a given year. IMR in Central Java continues to decline, and it was inseparable from the government's work program in reducing IMR. Besides, the number of infant mortality in 2016 reached 5485 cases, which continued to be suppressed until 2018, the number of infant deaths decreased to 4481 cases (Prahutama et al., 2018). IMR analysis based on the time factors was useful in providing information about changes in IMR. Therefore an assessment of the IMR analysis in Central Java is needed.

Infant Mortality Rate in Central Java has a data structure which is panel data where the data contains cross-section data (between units) and time-series data (between times). The method used to model cross-section data and time-series data is panel data regression. Panel data regression combines cross-section data and time-series data, where the same cross-section units are measured at different times (Greene, 2002). In example, panel regression can be used to model the economic productivity in outside of Java island based on the infrastructure, but it can't handle the correlation between each location (Sitorus & Yuliana, 2018). It's mean that panel regression cannot overcome the spatial effects. The spatial effect can contribute more analysis about dependency between location and its factors. For example, spatial regression was used to model poverty based on the unemployment rate in Indonesia. It concluded that there was an impact between poverty and unemployment rate, and there was dependency between locations (Rita, Diah, 2015). Moreover, spatial regression can be used to overcome the presence of spatial effects. Spatial panel regression has been done for modeling Gross Domestic Product in Jambi province based on areas (Diputra et al., 2012). Besides, the spatial modeling poverty indicator in Central Java uses Geographically Weighted Regression (GWR). It resulted in the contribution of factors only 68.64% to poverty indicator (Slamet et al., 2018). In the other hand, modeling infant mortality in China using GWR has been done. The results showed that three significant variables were per capita income of rural residents, Engel's coefficient of rural residents, and the proportion of government health expenditure (S. Wang & Wu, 2020).

Modelling for panel types and spatial effects can be developed through spatial panel regression analysis by combining Geographically Weighted Regression (GWR) models with panel data regression models to form a Geographically Weighted Panel Regression (GWPR) model (Fotheringham et al., 2002). Geographically Weighted Regression (GWR) is a technique that investigates heterogeneity in cross-space data relationships (Lu et al., 2014). (Soemartojo et al., 2018) uses GWR model that applied in inpatients claim data. This paper models infant mortality in Central Java using the GWPR model. Also (Siswantining et al., 2020) applied spatial analysis for modelling tuberculosis disease in Jakarta city, Indonesia. The dependent variable is the infant mortality rate in Central Java in 2015 up to 2017, while the independent variables used include the percentage of pregnant women who visit K1; Percentage of pregnant women who received Fe3 tablets; Percentage of deliveries assisted by health workers; The percentage of households that are clean and healthy. The novelty of this research is to apply GWPR to model infant mortality in central Java.

2. Materials

2.1. Model of Fixed Effect Geographically Weighted Panel Regression

Geographically Weighted Panel Regression (GWPR) is a model development that combines the Geographically Weighted Regression (GWR) model with panel data regression. GWPR has the same idea as the GWR cross-sectional analysis, which combines the entire location (cross-section) and observation (Chotimah et al., 2019). According to (J. Wang et al., 2020). The equation of the Fixed Effect Geographically Weighted Panel Regression model is as follows:

$$y_{it} = \sum_{k=1}^{p} \beta_k \left(u_{i,} v_i\right) X_{itk} + \varepsilon_{it} \text{ ; } i = 1, 2, ..., n \text{ and } t = 1, 2, \ldots, T \quad (1)$$

y_{it} is a dependent variable in observation location of i, at the time of t, $\beta_k(u_i, v_i)$ is the coefficient of independent variable regression in observation location of i; (u_i, v_i) are coordinates of the geographical location of the observation location of i; X_{itk} is an independent variable of k in the observation location of i, at the time of t. ε_{it} is observation residual in the observation location of i, at the time of t.

2.2. Parameter Estimation of Fixed Effect GWPR Model

In estimating the parameters of the GWPR Fixed Effect model, we used the Weighted Least Square approach as estimated in the GWR model by giving different weighting to each location where the data was taken (Cai et al., 2014)

According to Wang et al. (2020), to get an estimate of the GWPR Fixed Effect model at each location (u_i, v_i) done by adding weight $w_{it}(u_i, v_i)$ with *i=1,2,...,n,* and *t= 1,2,..., T*. In equation (1) then minimizing the number of residual squares is then derived and equated with zero. The parameter estimator of the GWPR model for each observation point is:

$$\begin{aligned} \sum_{t=1}^{T}\sum_{i=1}^{n} w_{it}(u_i v_i) y_{it} &= \sum_{t=1}^{T}\sum_{i=1}^{n} w_{it}(u_i v_i) \sum_{k=1}^{p} \beta_k(u_{i,} v_i) X_{itk} \\ &\quad + \sum_{t=1}^{T}\sum_{i=1}^{n} w_{it}(u_i v_i) \varepsilon_{it} \\ \sum_{t=1}^{T}\sum_{i=1}^{n} w_{it}(u_i v_i)\, {\varepsilon_{it}}^2 &= \sum_{t=1}^{T}\sum_{i=1}^{n} w_{it}(u_i v_i)[y_{it} \\ &\quad - \sum_{k=1}^{p} \beta_k((u_{i,} v_i) X_{itk}]^2 \end{aligned} \tag{2}$$

If written in the form of a matrix, the parameter estimator of the GWPR model for each observation point uses weighted least square is as follows:

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

If $\mathbf{X}\boldsymbol{\beta}(\boldsymbol{u_i}, \boldsymbol{v_i}) = \boldsymbol{\beta}^\mathbf{T}(\boldsymbol{u_i}, \boldsymbol{v_i})\mathbf{X}^\mathbf{T}$ then:

$$\begin{aligned} \boldsymbol{\varepsilon}^{\boldsymbol{T}} \boldsymbol{W}(\boldsymbol{u_i}, \boldsymbol{v_i})\boldsymbol{\varepsilon} &= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{W}(\boldsymbol{u_i}, \boldsymbol{v_i})\boldsymbol{y} - \mathbf{2}\boldsymbol{\beta}^\mathbf{T}(\boldsymbol{u_i}, \boldsymbol{v_i})\mathbf{X}^\mathbf{T}\mathbf{W}(\boldsymbol{u_i}, \boldsymbol{v_i})\boldsymbol{y} \\ &\quad + \boldsymbol{\beta}^{\boldsymbol{T}}(\boldsymbol{u_i}, \boldsymbol{v_i})\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{W}(\boldsymbol{u_i}, \boldsymbol{v_i})\mathbf{X}\boldsymbol{\beta}(\boldsymbol{u_i}, \boldsymbol{v_i}) \end{aligned} \tag{3}$$

The number of residual squares will be minimum with the first derivative condition equation (3) to $\boldsymbol{\beta}^\mathbf{T}(\boldsymbol{u_i}, \boldsymbol{v_i})$ equated with zero, and the second derivative is positive. If Equation (3) is lowered against $\boldsymbol{\beta}^\mathbf{T}(\boldsymbol{u_i}, \boldsymbol{v_i})$ and the result is equal to zero then is obtained:

$$\frac{\partial \boldsymbol{\varepsilon}^T \mathbf{W}(\boldsymbol{u_i}, \boldsymbol{v_i})\boldsymbol{\varepsilon}}{\partial \boldsymbol{\beta}^T(u_i, v_i)} = 0$$

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

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

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = [\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)$$

with

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{u_i},\boldsymbol{v_i}) = \begin{bmatrix} \hat{\beta}_1(u_i,v_i) \\ \hat{\beta}_2(u_i,v_i) \\ \vdots \\ \hat{\beta}_p(u_i,v_i) \end{bmatrix} \tag{4}$$

For the second derivative of Equation (4) to $\boldsymbol{\beta}$, obtained:

$$\frac{\boldsymbol{\partial^2 \varepsilon^T W(u_i,v_i)\varepsilon}}{\boldsymbol{\partial^2 \beta^T(u_i,v_i)}} = \boldsymbol{2X^T W(u_i,v_i)X = 2X^T X W(u_i,v_i)}$$

$\boldsymbol{X}^T\boldsymbol{X}$ is a positive definite matrix with all the main diagonal elements in the form of squares. If $\boldsymbol{x}_{it}^T = (x_{it1}, x_{it2}, \dots, x_{nTp})$ is row element of i from matrix $\mathbf{X^T}$ then the estimator of y in observation location (u_i,v_i) is:

$$\widehat{\boldsymbol{y}}_{\boldsymbol{it}} = \boldsymbol{x}_{\boldsymbol{it}}^{\boldsymbol{T}}\widehat{\boldsymbol{\beta}}(\boldsymbol{u_i},\boldsymbol{v_i})$$

$$\widehat{\boldsymbol{y}}_{\boldsymbol{it}} = \boldsymbol{x}_{\boldsymbol{it}}^{\boldsymbol{T}}[\boldsymbol{X^T W(u_i,v_i)X}]^{-1}\boldsymbol{X^T W(u_i,v_i)y}$$

Given $[\boldsymbol{X^T W}(u_i,v_i)\boldsymbol{X}]^{-1}\boldsymbol{X^T W}(u_i,v_i)$ is row element of i from matrix $\mathbf{L}$, so that the estimator of y all observations can be determined as follows:

$\widehat{\boldsymbol{y}} = (\hat{y}_{1t}, \hat{y}_{2t}, \dots, \hat{y}_{nT})^T = \boldsymbol{Ly}$

$\widehat{\boldsymbol{\varepsilon}} = (\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t}, \dots, \hat{\varepsilon}_{nT}) = (\boldsymbol{I} - \boldsymbol{L})\boldsymbol{y},$

with $\mathbf{I}$ is a sized identity matrix ($nT \times nT$), and matrix of $\mathbf{L}$ can be written as:

$$\boldsymbol{L}_{(\boldsymbol{nT}\times\boldsymbol{nT})} = \begin{bmatrix} \boldsymbol{x_{11}^T[X^T W(u_i,v_i)X]^{-1}X^T W(u_i,v_i)} \\ \boldsymbol{x_{21}^T[X^T W(u_i,v_i)X]^{-1}X^T W(u_i,v_i)} \\ \vdots \\ \boldsymbol{x_{n1}^T[X^T W(u_i,v_i)X]^{-1}X^T W(u_i,v_i)} \\ \vdots \\ \boldsymbol{x_{1T}^T[X^T W(u_i,v_i)X]^{-1}X^T W(u_i,v_i)} \\ \boldsymbol{x_{2T}^T[X^T W(u_i,v_i)X]^{-1}X^T W(u_i,v_i)} \\ \vdots \\ \boldsymbol{x_{nT}^T[X^T W(u_i,v_i)X]^{-1}X^T W(u_i,v_i)} \end{bmatrix}$$

2.3. Weighted of Fixed Effect GWPR Model

In giving weight to the GWPR model, it is the same as the weighting in the GWR model, which depends on the distance between the points of observation. Observations in local sampling locations will be given weights based on kernel functions on GWPR as well as in GWR (Cai et al., 2014)

The kernel function gives the weight according to the optimum bandwidth, whose value depends on the condition of the data. The kernel function is used to estimate the parameters in the model if the distance function is a function that is monotone down. Here is the dimensionless weighting matrix ($nT \times nT$) (Ningrum et al., 2020):

$$\mathbf{W}(\boldsymbol{u_i},\boldsymbol{v_i}) = \begin{bmatrix} w_{1t} & 0 & \cdots & 0 \\ 0 & w_{2t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{nT} \end{bmatrix}$$

$\mathbf{W}(u_i,v_i)$ is a weighted matrix in observation location of i with dimension $nT \times nT$. w_{nT} is weighted for the data of n at the time of T in around observation location.

One of the weights formed by using kernel functions is the kernel bisquare function. The kernel bisquare weighting function is (Bai et al., 2020)

$$\text{Fixed Bisquare: } w_{ij} = \begin{cases} \left(1-\left(\frac{d_{ij}}{h}\right)^2\right)^2 & , for\ d_{ij} < h \\ 0, other \end{cases}$$

where

$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}$ is Euclid distance between observation location of i with observation location of j; u is latitude and v is longitude, and h is bandwidth in all location.

Optimum bandwidth selection is important because it will affect the accuracy of the model for data. Several methods can be used to choose the optimum bandwidth. One method that can be used to select optimum bandwidth is using Cross-Validation (Fotheringham et al., 2002).

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

with $\bar{y}_i$ is an average of the dependent variable from time to time in observation location of i and $\hat{\bar{y}}_{\neq i}(b)$ is estimator of $\bar{y}_i$ with *bandwidth* h with observation location (u_i, v_i) which was removed from the estimation process.

2.4. The Testing of Fixed Effect GWPR Model

Testing the Fixed Effect GWPR model includes a model match test and a partial test of parameter significance.

a. Fit Test of the Model

According (J. Wang et al., 2020), the hypothesis as follows:

H_0 : $\beta_k(u_i, v_i) = \beta_k$ for $k = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, n$

(there is no significant difference between the panel data regression model and GWPR)

H_1 : At least there is one of $\beta_k(u_i, v_i) \neq \beta_k$ for $k = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, n$

(there is a significant difference between the panel data regression model and GWPR)

Test Statistic

$$F = \frac{RSS(H_1)/df_1}{RSS(H_0)/df_2}$$

$RSS(H_0) = \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{H})\boldsymbol{y}$ is Residual Sum of Square of Panel regression

where $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T}$; $RSS(H_1) = \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{y}$ is *Residual Sum of Square* of GWPR model

$df_1 = nT - p - 1$; $df_2 = \frac{\delta_1^2}{\delta_2}$, where $\delta_i = tr([(\boldsymbol{I} - \boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{L})]^{\boldsymbol{i}})$, $i = 1,2$

I is an identity matrix measuring $nT \times nT$, **L** is a projection matrix of the GWPR model. If the level of significance is given α, then reject H_0 $F > F_{1-\alpha, df1, df2}$ or p-value $< \alpha$. It means that there are significant differences between the panel data regression model and GWPR.

b. Test of Significant Parameter Model

This test is conducted to find out which parameters significantly influence the dependent variable. The following is the testing hypothesis (J. Wang et al., 2020)

H_0: $\beta_k(u_i, v_i) = 0$, for k = 1, 2, …, p and i = 1, 2, …, n

(The variable of X_k is not significant to y_i)

H_1: at least one of $\beta_k(u_i, v_i) \neq 0$ for k = 1, 2, …, p and i = 1, 2, …, n

(The variable of X_k is significant to y_i)

Parameter estimate of $\hat{\beta}_k(u_i, v_i)$ will follow the normal distribution with mean $\beta_k(u_i, v_i)$ and variance-covariance matrix is $\boldsymbol{C_i C_i^T}\sigma^2$, with C = $(\boldsymbol{X^T W}(u_i, v_i)\boldsymbol{X})^{-1}\boldsymbol{X^T W}(u_i, v_i)$ so that it gets

$$\frac{\hat{\beta}_k(u_i, v_i) - \beta_k(u_i, v_i)}{\sigma\sqrt{C_{kk}}}$$

with $\boldsymbol{C}_{kk}$ is a diagonal element of k from the matrix $\boldsymbol{C_i C_i^T}$ and $\hat{\sigma} = \sqrt{\frac{RSS(H_1)}{\delta_1}}$

$$RSS(H_0) = \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{H})\boldsymbol{y} \text{ with } \boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T}$$

$$RSS(H_1) = \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{y}$$

$$\delta_1 = tr((\boldsymbol{I} - \boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{L}))$$

$$\delta_2 = tr((\boldsymbol{I} - \boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{L}))^2$$

So the test statistic used is:

$$T_{value} = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma}\sqrt{C_{kk}}}$$

T_{value} will follow t distribution with degree of freedom df = $\frac{\delta_1^2}{\delta_2}$. If the level of significance α, then reject H_0 $|T_{value}| > t_{(\frac{\alpha}{2}, df)}$ or p-value < α.

3. METHODS

The independent variable used in this study is the Infant Mortality Rate in Central Java (in each district/city in Central Java). The dependent variable used in this study is the percentage of pregnant women who visited K1 (First visited) (X_1); Percentage of pregnant women who received Fe3 tablets (X_2); Percentage of labour assisted by health personnel (X_3); Percentage of households that are clean and healthy (X_4). The year used is 2015-2017. The Panel Data Structure that was used in this analysis can be seen in Table 1. In this analysis, uses *R* software to build the syntax. Following are the steps in the analysis of GWPR modelling as follows:

1. Calculate the Euclidean distance between the location of i and the location of j, which has a coordinate (u_i, v_i).
2. Calculate the optimum bandwidth with a minimum CV method.

3. Calculate fixed weighting bisquare matrices using optimum bandwidth.
4. Calculate parameter estimates of the GWPR Fixed Effect model using the bisquare fixed weighting matrix.
5. Test hypothesis of GWPR's Fixed Effect model.

Also in this analyzed, made the spread of map the number infant mortality based on GWPR model.

Table 1. The Panel Data Structure

District	Years	Y_{it}	X_{1it}	X_{2it}	$\cdots$	X_{4it}	U_i	V_i
1		$Y_{1,1}$	$X_{1,1,1}$	$X_{2,1,1}$	$\cdots$	$X_{4,1,1}$	U_1	V_1
2		$Y_{2,1}$	$X_{1,2,1}$	$X_{2,2,1}$	$\cdots$	$X_{4,2,1}$	U_2	V_2
3	2015	$Y_{3,1}$	$X_{1,3,1}$	$X_{2,3,1}$	$\cdots$	$X_{4,3,1}$	U_3	V_3
…		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
35		$Y_{35,1}$	$X_{1,35,1}$	$X_{2,35,1}$	$\cdots$	$X_{4,35,1}$	U_{35}	V_{35}
1		$Y_{1,2}$	$X_{1,1,2}$	$X_{2,1,2}$	$\cdots$	$X_{4,1,2}$	U_1	V_1
2	2016	$Y_{2,2}$	$X_{1,2,2}$	$X_{2,2,2}$	$\cdots$	$X_{4,2,2}$	U_2	V_2
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
35		$Y_{35,2}$	$X_{1,35,2}$	$X_{2,35,2}$	$\cdots$	$X_{4,35,2}$	U_{35}	V_{35}
1		$Y_{1,3}$	$X_{1,1,3}$	$X_{2,1,3}$	$\cdots$	$X_{4,1,3}$	U_1	V_1
2		$Y_{2,3}$	$X_{1,2,3}$	$X_{2,2,3}$	$\cdots$	$X_{4,2,3}$	U_2	V_2
3	2017	$Y_{3,3}$	$X_{1,3,3}$	$X_{2,3,3}$	$\cdots$	$X_{4,3,3}$	U_3	V_3
…		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
35		$Y_{35,3}$	$X_{1,35,3}$	$X_{2,35,3}$	$\cdots$	$X_{4,35,3}$	U_{35}	V_{35}

4. RESULTS AND DISCUSSION

4.1. The GWPR Fixed Effect Model of Infant Mortality in Central Java

In GWPR fixed effect modelling, the first step is to determine the geographical location of each village located in the province of Central Java. Then, the average for the dependent variable and the independent variable is calculated for the entire time in each location to obtain the bandwidth value and weighting value.

Then the optimum bandwidth value is searched using cross-validation (CV) criteria. Table 2 shows the bandwidth value with fixed weight Bisquare kernel *h*

Table 2. The Value of Bandwidth

Bandwidth	Cross Validation (CV)
0.1870044	0.03670715
0.1155982	0.17199960
0.2311359	0.03834037
0.1597297	0.03894728
0.2038612	0.03703698
0.1765864	0.03712706
0.1934431	0.03672399

The optimum bandwidth value obtained based on Table 2 is 0.1870044 with a CV value of 0.03670715 because it has the smallest CV value.

After getting the optimum bandwidth value, the next step is to find a weighting matrix using the fixed Bisquare kernel weighting function. The location weighting matrix (u_i, v_i) is obtained using Euclidean distance (d_{ij}). The weighted value that has been obtained will be used to estimate the parameters of the GWPR fixed effect for each location. Because each location has a different weighting value, it allows the parameter estimation values to have different values. The results of the GWPR fixed effect modeling parameter estimation with fixed Bisquare kernel weighted in 35 districts / cities in Central Java can be seen in Table 3.

Table 3. The Parameter Estimate of Fixed Effect GWPR Model

District/City	β_1	β_2	β_3	β_4	District/City	β_1	β_2	β_3	β_4
Cilacap	-0.0021	-2.61	-0.009	-2.189	Pati	-0.0023	-0.933	-0.043	-0.706
Banyumas	-0.0345	-2.37	-0.098	-1.907	Kudus	-0.0064	-0.324	-0.054	-0.447
Purbalingga	-0.0698	-2.58	-0.076	-2.102	Jepara	-0.0311	-1.154	-0.001	-0.872
Banjarnegara	-0.0898	-1.92	-0.008	-1.474	Demak	-0.0044	-1.272	-0.001	-0.944
Kebumen	-0.0009	-1.83	-0.065	-1.384	Semarang	-0.0002	-1.390	-0.009	-1.023
Purworejo	-0.0032	-1.64	-0.007	-1.195	Temanggung	-0.0043	-1.280	-0.001	-0.923
Wonosobo	-0.0067	-1.74	-0.008	-1.316	Kendal	-0.0041	-1.741	-0.003	-1.322
Magelang	-0.0006	-1.49	-0.010	-1.072	Batang	-0.0055	-1.840	-0.034	-1.397
Boyolali	-0.0557	-1.25	-0.002	-0.870	Pekalongan	-0.0001	-1.970	-0.008	-1.511
Klaten	-0.0122	-1.21	-0.007	-0.817	Pemalang	-0.0045	-2.130	-0.081	-1.637
Sukoharjo	-0.0043	-1.37	-0.007	-0.955	Tegal	-0.0012	-2.410	-0.008	-1.885
Wonogiri	-0.0066	-0.94	-0.013	-0.572	Brebes	-0.0061	-2.740	-0.009	-2.116
Karanganyar	-0.0001	-1.01	-0.033	-0.667	Magelang City	-0.0067	-1.510	-0.008	-1.108
Sragen	-0.0056	-1.00	-0.005	-0.680	Surakarta City	-0.0051	-1.130	-0.065	-0.768
Grobogan	-0.0066	-1.48	-0.002	-1.130	Salatiga City	-0.0065	-1.440	-0.067	-0.964
Blora	-0.0051	-0.60	-0.021	-0.489	Semarang City	-0.0089	-1.976	-0.006	-1.080
Rembang	-0.0113	-0.73	-0.002	-0.580	Pekalongan City	-0.0114	-2.510	-0.098	-1.507
					Tegal City	-0.0221	-0.320	-0.006	-1.937

4.2. The Testing of Fixed Effect GWPR Model

Testing the GWPR with fixed effect model includes the model match test and the significance test of the model parameters as follows:

a. The Testing of Fit Modelling

This test is conducted to find out whether there is a difference between the fixed effect model of panel data regression and the GWPR fixed effect with the hypothesis:

H_0 : $\beta_k(ui,vi) = \beta_k$ for $k = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, n$

(there is no significant difference between the panel data regression model and GWPR)

H_1 : at least one of $\beta_k(ui,vi) \neq \beta_k$ for $k = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, n$

(there is significant difference between the panel data regression model and GWPR)

Test the suitability of the model using the F test is 3.43706 with p-value 0.000. Reject H_0 if $F > F(\alpha, df_1, df_2)$ or p-value $< \alpha$. In significance level $\alpha = 5\%$ with fixed bisquare kernel weighted obtained the value of $F = 3.43706 > F_{(0,05;86;106,0869)} = 1.3991$and p-value= $0.0000 < \alpha$ (0.05). It show that reject H_0, means there is significant difference between the panel data regression model and GWPR.

b. The Testing of Parameter Significance Modelling

This test is used to find out which independent variables influence the dependent variable in the GWPR fixed effect model with fixed bisquare kernel weighted with the following hypothesis:

H_0: $\beta_k(u_i, v_i) = 0$, for k = 1, 2, …, p and i = 1, 2, …, n
(The variable of X_k is not significant to y_i)
H_1: at least one of $\beta_k(u_i, v_i) \neq 0$ for k = 1, 2, …, p and i = 1, 2, …, n
(The variable of X_k is significant to y_i)

The results of individual parameter testing for each observation location with fixed bisquare kernel weighted can be seen in Appendix 1. Based on Appendix 1, it can be seen that the significant variable is X_4, which is the percentage of households living a clean and healthy life. While variables X_1, X_2, and X_3 are not significant. The independent variable is significant if the t-statistic value is greater than the value of the t-table (0.025; 106.0869) of 1.9826, or the p-value is less than 0.05. The R-square value generated from the model reached 67.6%.

4.3. Spatial Characteristics Analyzed from GWPR Model

Figures 1, 2, and 3 show the spread map of infant mortality in 2015, 2016, and 2017 in Central Java used the GWPR model. The GWPR model that was applied in the figures used a fixed-effect model with Bisquare Kernel weighted. Each figure was categorized by three levels, such as low, middle, and high. Each level from 2015 to 2017 has different values. The figures show the highest number in 2015 hit same greater than 530, details Brebes and Banjarnegara. However, in 2016 decreased gently, which was the same greater than 505, among others Brebes, Banyumas, Magelang, and Sragen. On the other hand, in 2017 decreased significantly, which was the same greater than 207 as follows Brebes, Pemalang, Banjarnegara, Magelang, and Semarang. Banjarnegara hit the highest score of number mortality based on GWPR model from 2015 until 2017. From year to year, there were changes in levels. For example, in 2016, Magelang and Banyumas increased smoothly one level, while Sragen increased two levels significantly. And also, there were decrease levels; for example, in 2017, Sragen dan Banyumas declined one level.

On the other hand, Appendix 2 shows the correlation significant from variables X_3 (Labor assisted by Health workers) and X_4 (Clean and health Lifestyle), which was plotted by the number of infant mortality in 2015, 2016, and 2017. In those plots, the number of infant mortality was divided into six levels, among others the lowest, pretty low, middle, pretty high, high, and highest. In addition, for Y-axis was X_3, and X-axis was X_4, consist of three levels of details, the lowest, medium, and the highest.

Figure 4 shows the lowest values of X_3 and X_4 variables that impact the high number of infant mortality were Brebes, Pemalang, Kudus, and Pati in 2015 by using the GWPR model. In details for Brebes was the highest while Pemalang, kudus, and Pati categorized in pretty high. Meanwhile, the highest value of X_3 and X_4 variables that impact the low number of infant mortality were Wonogiri, Klaten, Sragen, and Grobogan. For Wonogiri, Klaten and Sragen were categorized in lowest value, while for Grobogan were pretty low values.

Figure 5 shows the spread map the number of infant mortality by GWPR method in 2016. Based on the graph, the lowest value of clean and healthy lifestyle and labor assisted by health workers impact to the high of the number infant mortality were Brebes, Kendal, dan Banjarnegara. Banjarnegara hit the highest values for the number of infant mortality, while Brebes categorized in high score, and Kendal was pretty high. However the low of number infant moratlity in 2017 was Grobogan, Klaten, Wonogiri and Semarang city. Grobogan and Klaten were classified in the lowest score, while Klaten was ranked pretty low. However, in Semarang city, the value was pretty high, but the variables clean and healthy lifestyle and labor assisted by health worker were at the highest level.

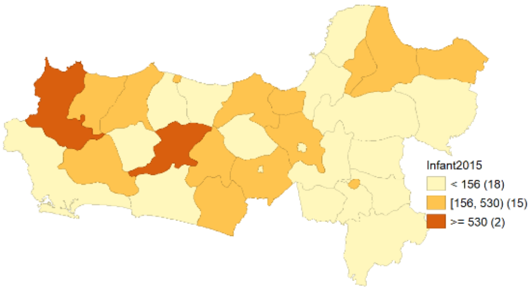

Figure 1. The Spread Map of Infant Mortality used GWPR in 2015

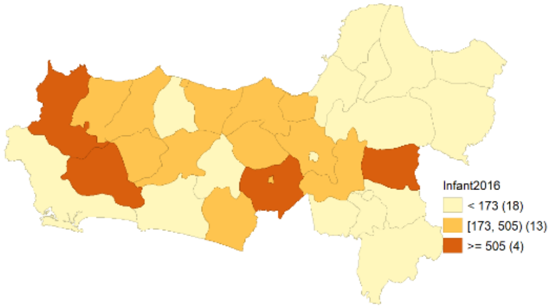

Figure 2. The Spread Map of Infant Mortality used GWPR in 2016

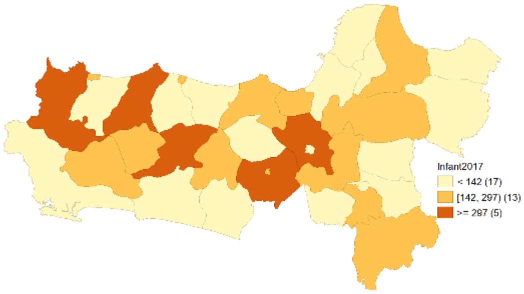

Figure 3. The Spread Map of Infant Mortality used GWPR in 2017

On the other hand, Figure 6 shows the highest score the number of infant mortality in 2017 was Brebes, Banyumas, Pemalang, and Purbalingga. They were in high value, while Boyolali and Pati were pretty high value. Meanwhile, the low value of the number of infant mortality based on X_3 and X_4 predictor variables was Wonogiri, Klaten, and Karanganyar. For Wonogiri got the lowest value of the number of infant mortality, while Klaten and Karanganyar hit the pretty low value.

5. CONCLUSION

Based on the results and discussion, there are differences significantly between the panel data regression model and the GWPR model, so it can be concluded that modeling the number of infant mortality in Central Java using GWPR. However, this paper didn't explain the panel data regression modeling. Additionally, the assumptions were not tested from the GWPR model. There is no guarantee that the GWPR model is better than the panel data regression model for other cases. In the future, it can be developed by a random effect model based on spatial analysis.

REFERENCES

Bai, X., Zhang, T., & Tian, S. (2020). Evaluating Fertilizer Use Efficiency and Spatial Correlation of its Determinants in China: A Geographically Weighted Regression Approach. *International Journal of Environmental Research and Public Health*, *17*(23), 1–23. https://doi.org/10.3390/ijerph17238830

Cai, R., Yu, D., & Oppenheimer, M. (2014). Estimating the Spatially Varying Responses of Corn Yields to Weather Variations Using Geographically Weighted Panel Regression. *Journal of Agricultural and Resource Economics*, *39*(2), 230–252. https://researchwith.montclair.edu/en/publications/estimating-the-spatially-varying-responses-of-corn-yields-to-weat

Chotimah, C., Sutikno, & Setiawan. (2019). Modelling of Income Inequality in East Java Using Geographically Weighted Panel Regression. *IOP Conference Series: Materials Science and Engineering*, *546*(5). https://doi.org/10.1088/1757-899X/546/5/052019

Diputra, T. F., Sadik, K., & Angraini, Y. (2012). Pemodelan Data Panel Spasial dengan Dimensi Ruang dan Waktu. *Forum Statistika Dan Komputasi: Indonesian Journal of Statistics*, *17*(1), 6–14. journal.ipb.ac.id/index.php/statistika

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relathionships*. John Wiley & Sons, Inc.

Greene, william H. (2002). *Econometric Analysis* (5th ed.). Prentice Hall.

Ningrum, A. S., Rusgiyono, A., & Prahutama, A. (2020). Village Classification Index Prediction Using Geographically Weighted Panel Regression. *Journal of Physics: Conference Series*, *1524*(1). https://doi.org/10.1088/1742-6596/1524/1/012040

Prahutama, A., Warsito, B., & Mukid, M. (2018). Modelling Infant and Maternal Mortality Rate based on slum household uses Bivariate Spline Regression Method. *E3S Web of Conferences*, *73*. https://doi.org/10.1051/e3sconf/20187312002

Rita, Diah, O. (2015). Analisis Spasial Pengaruh Tingkat Pengangguran Terhadap Kemiskinan di Indonesia (Studi Kasus: Provinsi Jawa Tengah). *Media Statistika*, *8*(1), 23–30.

Siswantining, T., Dewi Purwandani, N. P. C., Dewi Susilowati, M. H., & Wibowo, A. (2020). Geoinformatics of tuberculosis (TB) disease in Jakarta city Indonesia. *International Journal of GEOMATE*, *19*(72), 35–42. https://doi.org/10.21660/2020.72.5599

Sitorus, Y. M., & Yuliana, L. (2018). Penerapan Regresi Data Panel Pada Analisis Pengaruh Infrastruktur Terhadap Produktifitas Ekonomi Provinsi-Provinsi Di Luar Pulau Jawa Tahun 2010-2014. *Media Statistika*, *11*(1), 1–15. https://doi.org/10.14710/medstat.11.1.1-15

Slamet, I., Nugroho, N. F. T. A., & Muslich. (2018). Geographically Weighted Regression Model on Poverty Indicator. *Journal of Physics: Conference Series*, *943*(1). https://doi.org/10.1088/1742-6596/943/1/012009

Soemartojo, S. M., Ghaisani, R. D., Siswantining, T., Shahab, M. R., & Ariyanto, M. M. (2018). Parameter Estimation of Geographically Weighted Regression (GWR) Model Using Weighted Least Square and its Application. *AIP Conference Proceedings*, *2014*(September). https://doi.org/10.1063/1.5054485

Wang, J., Chen, K., & Song, X. (2020). Differences Among Influencing Factors of China's Provincial Energy Intensity: Empirical Analysis from a Geographically Weighted Regression Model. *Polish Journal of Environmental Studies*, *29*(4), 2901–2916. https://doi.org/10.15244/pjoes/113097

Wang, S., & Wu, J. (2020). Spatial Heterogeneity of the Associations of Economic and Health Care Factors with Infant Mortality in China ising Geographically Weighted Regression and Spatial Clustering. *Social Science and Medicine*, *263*, 113287.

Appendix 1. Test the Significance of the Parameters of GWPR Weighting Fixed Bisquare Kernel

Districs/Cities	t-statistic				p-value			
	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
Cilacap	0.939	0.692	-1.550	9.4319	0.349	0.490	0.0124	0.00
Banyumas	0.872	0.683	-1.510	9.5209	0.385	0.495	0.0134	0.00
Purbalingga	0.976	0.677	-1.580	10.0293	0.331	0.499	0.0116	0.00
Banjarnegara	0.576	0.752	-1.330	8.5723	0.565	0.453	0.0185	0.00
Kebumen	0.493	0.750	-1.270	9.5702	0.623	0.436	0.0205	0.00
Purworejo	0.330	0.780	-1.150	6.2213	0.742	0.430	0.0249	0.00
Wonosobo	0.434	0.792	-1.230	9.7922	0.665	0.410	0.0220	0.00
Magelang	0.211	0.827	-1.050	10.4008	0.833	0.385	0.0920	0.00
Boyolali	0.027	0.871	-0.880	10.3185	0.978	0.390	0.0380	0.00
Klaten	-0.004	0.861	-0.850	8.0554	0.996	0.404	0.0940	0.00
Sukoharjo	0.113	0.837	-0.970	10.0932	0.9103	0.377	0.0330	0.00
Wonogiri	-0.189	0.887	-0.640	9.1854	0.8502	0.371	0.0530	0.00
Karanganyar	-0.139	0.899	-0.690	7.6429	0.8899	0.361	0.0486	0.00
Sragen	-0.145	0.926	-0.680	4.3191	0.8852	0.390	0.0496	0.00
Grobogan	0.229	0.862	-1.050	8.2666	0.8914	0.323	0.0294	0.00
Blora	-0.360	0.993	-0.370	7.3156	0.7914	0.325	0.0712	0.00
Rembang	-0.281	0.988	-0.460	9.9839	0.7789	0.337	0.0644	0.00
Pati	-0.159	0.963	-0.620	9.7578	0.8741	0.307	0.0538	0.00
Kudus	-0.498	1.025	-0.180	10.1011	0.6194	0.350	0.0857	0.00
Jepara	0.006	0.937	-0.780	10.1421	0.9945	0.365	0.0433	0.00
Demak	0.073	0.908	-0.890	3.6054	0.9419	0.383	0.0376	0.00
Semarang	0.152	0.875	-0.980	8.1708	0.8795	0.379	0.0196	0.00
Temanggung	0.063	0.881	-0.910	10.2194	0.9504	0.417	0.0173	0.00
Kendal	0.451	0.815	-1.240	10.0596	0.6527	0.428	0.0147	0.00
Batang	0.538	0.796	-1.290	10.2072	0.5994	0.449	0.0119	0.00
Pekalongan	0.624	0.760	-1.370	8.1606	0.5338	0.464	0.0090	0.00
Pemalang	0.746	0.734	-1.450	9.7581	0.457	0.493	0.0429	0.00
Tegal	0.926	0.687	-1.570	10.0936	0.3565	0.509	0.0345	0.00
Brebes	1.080	0.661	-1.680	8.5444	0.2823	0.408	0.0305	0.00
Magelang City	0.240	0.829	-1.070	10.3399	0.8103	0.378	0.0170	0.00
Surakarta City	-0.057	0.885	-0.790	6.5766	0.9546	0.385	0.0110	0.00
Salatiga City	0.101	0.871	-0.940	8.6874	0.9191	0.385	0.0119	0.00
Semarang City	0.206	0.872	-1.028	9.0872	0.8365	0.444	0.0173	0.00
Pekalongan City	0.633	0.767	-1.370	6.0892	0.5281	0.498	0.0654	0.00
Tegal City	0.978	0.679	-1.611	5.0765	0.3305	0.350	0.0234	0.00

Appendix 2. Spread Map of The Number Infant Mortality uses GWPR Model

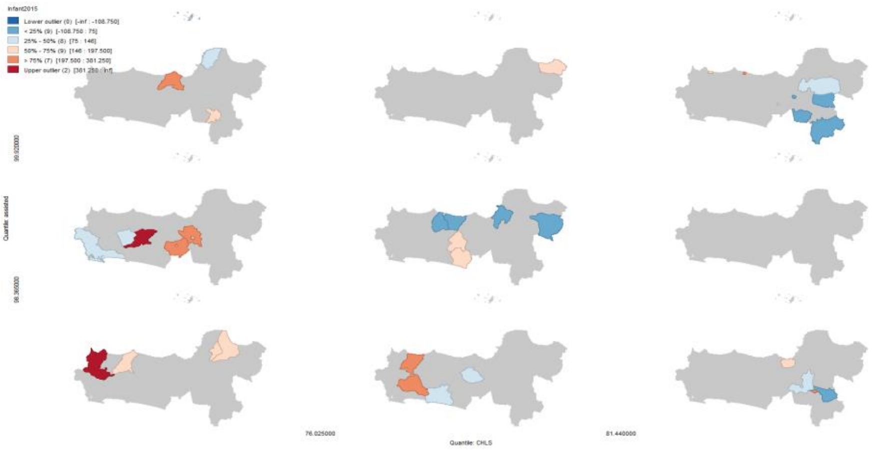

Figure 4. Spread map between independent variables significantly in GWPR model in 2015

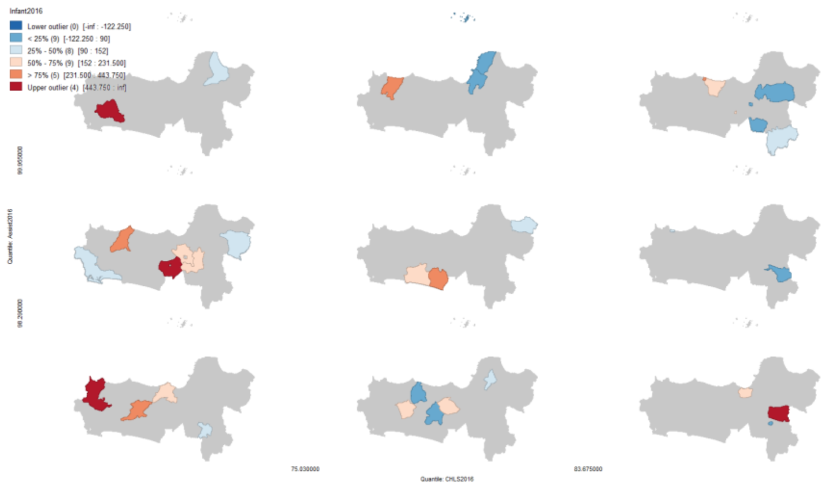

Figure 5. Spread Map between Independent Variables Significantly in GWPR model in 2016

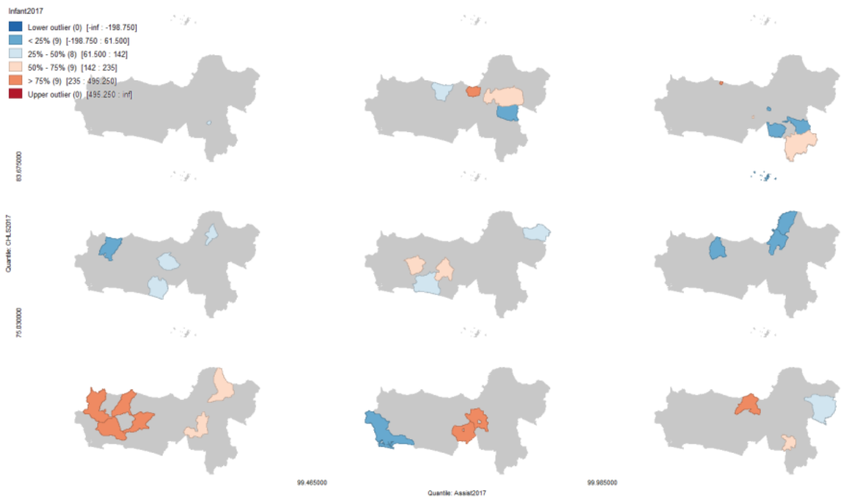

Figure 6. Spread Map between Independent Variables Significantly in GWPR model in 2016

Noted: Y-axix was %Labor Assited by Health Workes; X-Axix was %Clean and Health Lifestyle of Household