European Journal of Scientific Research
ISSN 1450-216X Vol.69 No.2 (2012), pp. 188-196

http://www.europeanjournalofscientificresearch.com

Mixed Geographically Weighted Regression Model (Case Study: the Percentage of Poor Households in Mojokerto 2008)

Purhadi
Statistics of Institute of Technology Sepuluh Nopember Surabaya
E-mail: purhadi@statistika.its.ac.id

Hasbi Yasin
Statistics of Diponegoro University Semarang
E-mail: hasbiyasin@yahoo.com

Abstract

Regression analysis is a statistical analysis that aims to model the relationship between response variables with predictor variables. Geographically Weighted Regression (GWR) is statistical methods used for analyzed the spatial data in local form of regression. Where certain predictor variables influencing the response are global while others are local used the Mixed Geographically Weighted Regression (MGWR) model to solve the problem. The results showed that Weighted Least Square (WLS) can be used to estimate the parameter model and Cross Validation (CV) for the selection of the optimum bandwidth. Goodness of fits tests for a global regression model and MGWR approximated by F distribution as well as on the test of global parameters and local parameters simultaneously and for testing the partial model parameters using the t distribution. The applications of MGWR model in the percentage of poor households in Mojokerto showed that MGWR model differs significantly from the global regression model. Based on Akaike Information Criterion (AIC) values between the global regression model, GWR and MGWR model, it is known that the MGWR model with a weighting Gaussian kernel function is the best model used to analyze the percentage of poor households in Mojokerto (2008) because it has the smallest AIC value.

Keywords: Akaike Information Criterion, Cross Validation, Gaussian Kernel Function, Mixed Geographically Weighted Regression, Weighted Least Square

1. Introduction

One indicator of disadvantaged areas is the high percentage of poor people who inhabit it [1]. Poor condition is seen as the inability of the economy to meet the basic needs of food and non food as measured from the expenditure side. Therefore we need an effort to reduce poverty based on the conditions in each area. Modeling the percentage of poor households based on the characteristics of the area will be affected by the geographical location [2]. This is due to geographical differences will affect the potential possessed or used by an area.

Mixed geographically Weighted Regression (MGWR) is a combination of global linear regression model with the GWR model. So that the model will be generated MGWR estimator parameters are global and some others are localized in accordance with the location of observations ([3], [4]) onducted a modeling MGWR in the case of rural community development in the European

Union with the predictor variables was based on indicators of agricultural and socio-economics. Estimation of parameters in the model MGWR can be done with WLS method as in the GWR model [5]. It is not yet known properties of the resulting estimator. Likewise, the test statistic used to measure the significance of the MGWR parameters. Therefore we need further study on the estimator of MGWR parameters. The MGWR model will be applied to modeling percentage of poor households in Mojokerto.

2. Previous Research

GWR model represents the development of a global regression model where the basic idea is taken from the non-parametric regression [5]. This model is a locally linear regression that that produces the estimator model parameters that are local to each point or location where the data is collected. GWR model is widely used by researchers in analyzing spatial data in various fields, because the method of GWR can be used to determine the effect of predictor variables in order to response variables both globally and locally by considering the elements of geography or location as a weighting in estimating the model parameters [3]. Estimation of GWR model parameter are using the Weighted Least Squares (WLS) method that give a different weighting for each observation. Weighting that used to estimate the parameters in the GWR model are the kernel functions: Gaussian distance function, exponential functions, Bisquare functions, and Tricube functions [7]. The best model selection is done by determining the model with the smallest AIC value [10].

Poverty is seen as the inability of the economy to meet the basic needs of food and non food as measured from the expenditure side. To measure poverty, BPS uses the concept of ability to meet basic needs (basic needs' approach). With this approach, can be calculated *Headcount Index*, the percentage of poor population to total population. One method for calculating the poverty line is the poverty line Expenditure / Consumption approach [1]. Factors that influence poverty among other natural and environmental factors, infrastructure factors and socioeconomic factors [2]

3. Hypotheses

While testing the effect of each predictor variable geographic done with the following hypotheses:

$H_0 : \beta_k(u_1,v_1) = \beta_k(u_2,v_2) = \cdots = \beta_k(u_n,v_n)$ for k $(k = 0,1,2,\cdots,p)$

(There is no significant difference in the effect of predictor variables between one location to another)

H_1 : At least there is a $\beta_k(u_i,v_i)$, for $i = 1,2,...,n$ different one

Testing parameters are partially done with the hypothesis as follows:

$H_0 : \beta_k(u_i,v_i) = 0$

$H_1 : \beta_k(u_i,v_i) \neq 0$ where $k = 1,2,\cdots,p$

Test statistic given by:

$$T = \frac{\hat{\beta}_k(u_i,v_i)}{\hat{\sigma}\sqrt{c_{kk}}} \tag{1}$$

where $\hat{\sigma}^2 = \dfrac{\mathbf{y}^T(\mathbf{I}-\mathbf{L})^T(\mathbf{I}-\mathbf{L})\mathbf{y}}{\delta_1}$ and c_{kk} is the k-th diagonal element of the matrix $\mathbf{C}\mathbf{C}^T$ where

$$\mathbf{C} = \left(\mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}(u_i,v_i). \tag{2}$$

Reject H_0 or in other words the $\beta_k(u_i, v_i)$ is a significant parameters of the model when $|T| > t_{\alpha/2, df}$, where $df = \left[\delta_1^2 / \delta_2 \right]$.

4. Research Method

The data used in this study are secondary data obtained from the Central Bureau of Statistics that is data Target Households Documenting the results of the Social Protection Program 2008 and the Village Potential data in 2008 for the Mojokerto regency. The unit of observation used are the villages in Mojokerto, East Java, consisting of 304 villages. To support the research process used the MINITAB and the computer programs and algorithms developed in MATLAB script [8].

The variable used is the percentage of poor households per village/villages in Mojokerto regency in 2008 as a response variable (Y) and variable population density (X_1), the percentage of families living on the banks / banks of the river (X_2), the percentage of families living in slums (X_3), number of people with disease outbreaks over the last year (X_4), the distance from village to district capital (X_5), the distance from the village to the city (X_6), the distance from the village to the district capital / other cities nearby (X_7), the ratio of basic educational facilities (elementary / junior high schools equal) per 100 inhabitants (X_8), the ratio of primary health facilities (PHC / Sub Health Center / the doctor or midwife / village health post) per 100 inhabitants (X_9), the ratio of health personnel per 100 inhabitants (X_{10}), the percentage of families who subscribe to the telephone wiring (X_{11}), the number of trade centers and cooperatives (without building market, mini market / restaurant / restaurant / shop / food and drink stalls / shop grocery store / cooperative) (X_{12}), the percentage of farm families (X_{13}), the percentage of people with malnutrition in the last 3 years (X_{14}) as the predictor variable, and Latitude (u_i) and Longitude (v_i) of the district capital.

To determine the factors that affect the percentage of poor households in rural and urban villages in Mojokerto regency in 2008 performed the analysis with the following steps:

1. Analyzing the global linear regression model
2. Analyzing the GWR model
3. Analyzing MGWR model
4. Compare the global regression model, GWR and MGWR

5. Results and Discussion

5.1. Parameter Estimation and Hypothesis Testing of MGWR Model

This research will be used in the WLS method of estimation parameters. Generally MGWR model can be written by:

$$\mathbf{y} = \mathbf{X}_l \boldsymbol{\beta}_l(u_i, v_i) + \mathbf{X}_g \boldsymbol{\beta}_g + \boldsymbol{\varepsilon} \tag{3}$$

where $\mathbf{X}_g$: matrix of global variables, $\mathbf{X}_l$: matrix of local variables, $\boldsymbol{\beta}_g$: vector of global parameters and $\boldsymbol{\beta}_l(u_i, v_i)$: matrix of local parameters.

MGWR model was first written in the form of GWR following:

$$\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{X}_g \boldsymbol{\beta}_g = \mathbf{X}_l \boldsymbol{\beta}_l(u_i, v_i) + \boldsymbol{\varepsilon} \tag{4}$$

So the first estimator of GWR parameter GWR is:

$$\hat{\boldsymbol{\beta}}_l(u_i, v_i) = \left[\mathbf{X}_l^T \mathbf{W}(u_i, v_i) \mathbf{X}_l \right]^{-1} \mathbf{X}_l^T \mathbf{W}(u_i, v_i) \tilde{\mathbf{y}} \tag{5}$$

Let $\mathbf{x}_{li}^T = (1, x_{i1}, x_{i2}, \cdots, x_{iq})$ is the element of the i-th row of the matrix $\mathbf{X}_l$. Then the predicted value for $\tilde{y}$ at the (u_i, v_i) for the entire observation can be written by:

$$\hat{\tilde{\mathbf{y}}} = \left(\hat{\tilde{y}}_1, \hat{\tilde{y}}_2, \cdots, \hat{\tilde{y}}_n\right)^T = \mathbf{S}_l \tilde{\mathbf{y}}$$

where

$$\mathbf{S}_l = \begin{pmatrix} \mathbf{x}_{l1}^T \left(\mathbf{X}_l^T \mathbf{W}(u_1, v_1) \mathbf{X}_l\right)^{-1} \mathbf{X}_l^T \mathbf{W}(u_1, v_1) \\ \mathbf{x}_{l2}^T \left(\mathbf{X}_l^T \mathbf{W}(u_2, v_2) \mathbf{X}_l\right)^{-1} \mathbf{X}_l^T \mathbf{W}(u_2, v_2) \\ \vdots \\ \mathbf{x}_{ln}^T \left(\mathbf{X}_l^T \mathbf{W}(u_n, v_n) \mathbf{X}_l\right)^{-1} \mathbf{X}_l^T \mathbf{W}(u_n, v_n) \end{pmatrix} \tag{6}$$

Then, substitute the element of $\hat{\boldsymbol{\beta}}_l(u_i, v_i)$ into MGWR model in equation (4) to obtain:

$$\hat{\boldsymbol{\beta}}_g = \left[\mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l) \mathbf{X}_g\right]^{-1} \mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l) \mathbf{y}$$

where $\mathbf{S}_g = \mathbf{X}_g \left(\mathbf{X}_g^T \mathbf{X}_g\right)^{-1} \mathbf{X}_g^T$ (7)

By substituting $\hat{\boldsymbol{\beta}}_g$ into equation (5) then obtained an estimate for the local coefficient on the location (u_i, v_i) is:

$$\hat{\boldsymbol{\beta}}_l(u_i, v_i) = \left[\mathbf{X}_l^T \mathbf{W}(u_i, v_i) \mathbf{X}_l\right]^{-1} \mathbf{X}_l^T \mathbf{W}(u_i, v_i) \left(\mathbf{y} - \mathbf{X}_g \hat{\boldsymbol{\beta}}_g\right) \tag{8}$$

Therefore, the fitted-value of the response for n observation are:

$$\hat{\mathbf{y}} = \mathbf{S}\mathbf{y}$$

where

$$\mathbf{S} = \mathbf{S}_l + (\mathbf{I} - \mathbf{S}_l) \mathbf{X}_g \left[\mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l) \mathbf{X}_g\right]^{-1} \mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l). \tag{9}$$

Estimator $\hat{\boldsymbol{\beta}}_g$ and $\hat{\boldsymbol{\beta}}_l(u_i, v_i)$ is an unbiased estimator for $\boldsymbol{\beta}_g$ and $\boldsymbol{\beta}_l(u_i, v_i)$.

Hypothesis testing was first done is testing the suitability of global regression and MGWR model to test the significance of geographical factors. Form hypotheses are:

H_0 : $\beta_k(u_i, v_i) = \beta_k \quad k = 0, 1, 2, \cdots, q,$ and $i = 1, 2, \cdots, n$

H_1 : at least there is one $\beta_k(u_i, v_i) \neq \beta_k$

Conformance testing of global regression model and MGWR using the comparative value of the difference residual sum of squares of global regression model and MGWR model. So the test statistic is:

$$F(1) = \frac{\mathbf{y}^T \left[(\mathbf{I} - \mathbf{H}) - (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S})\right] \mathbf{y} / v_1}{\mathbf{y}^T (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S}) \mathbf{y} / u_1} \tag{10}$$

Reject H_0 if $F(1) \geq F_{\alpha, df_1, df_2}$ where $df_1 = \left(v_1^2 / v_2\right)$, $df_2 = \left(u_1^2 / u_2\right)$,

$v_i = tr\left(\left[(\mathbf{I} - \mathbf{H}) - (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S})\right]^i\right)$, $i = 1, 2$, $u_i = tr\left(\left[(\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S})\right]^i\right)$, $i = 1, 2$ and $\mathbf{H} = \mathbf{X}\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T$.

Testing parameters simultaneously on a global predictor variables performed with the hypothesis :

H_0 : $\beta_{q+1} = \beta_{q+2} = \cdots = \beta_p = 0$

H_1 : at least one of $\beta_k \neq 0$

The test statistic is:

$$F\left(2\right)=\frac{\mathbf{y}^T\left[\left(\mathbf{I}-\mathbf{S}_l\right)^T\left(\mathbf{I}-\mathbf{S}_l\right)-\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\right]\mathbf{y}\Big/r_1}{\mathbf{y}^T\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\mathbf{y}\Big/u_1} \tag{11}$$

Reject H$_0$ if $F\left(2\right)\geq F_{\alpha,df_1,df_2}$ where $df_1=\left(r_1^2\big/r_2\right)$, $df_2=\left(u_1^2\big/u_2\right)$

and $r_i=tr\left(\left[\left(\mathbf{I}-\mathbf{S}_l\right)^T\left(\mathbf{I}-\mathbf{S}_l\right)-\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\right]^i\right)$, $i=1,2$.

Testing parameters simultaneously on a local predictor variables $x_k\left(1\leq k\leq q\right)$ performed with the hypothesis:

H$_0$: $\beta_1\left(u_i,v_i\right)=\beta_2\left(u_i,v_i\right)=\cdots=\beta_q\left(u_i,v_i\right)=0$

H$_1$: at least one of $\beta_k\left(u_i,v_i\right)\neq 0$

The test statistic is:

$$F\left(3\right)=\frac{\mathbf{y}^T\left[\left(\mathbf{I}-\mathbf{S}_g\right)^T\left(\mathbf{I}-\mathbf{S}_g\right)-\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\right]\mathbf{y}\Big/t_1}{\mathbf{y}^T\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\mathbf{y}\Big/u_1} \tag{12}$$

Reject H$_0$ if $F\left(3\right)\geq F_{\alpha,df_1,df_2}$ where $df_1=\left(t_1^2\big/t_2\right)$, $df_2=\left(u_1^2\big/u_2\right)$

and $t_i=tr\left(\left[\left(\mathbf{I}-\mathbf{S}_g\right)^T\left(\mathbf{I}-\mathbf{S}_g\right)-\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\right]^i\right)$, $i=1,2$.

Tests carried out by testing the parameters of the model parameters partially. This test is to determine which parameters significantly affect the response variable. To test the significance of the global variable $x_k\left(q+1\leq k\leq p\right)$ is used the hypothesis:

H$_0$: $\beta_k=0$ (global variables x_k are not significant)

H$_1$: $\beta_k\neq 0$ (global variables x_k are significant)

The test statistic is:

$$T_g=\frac{\hat{\beta}_k}{\hat{\sigma}\sqrt{g_{kk}}} \tag{13}$$

where g_{kk} is the k-th diagonal element of matrix $\mathbf{GG}^T$

and $\mathbf{G}=\left[\mathbf{X}_g^T\left(\mathbf{I}-\mathbf{S}_l\right)^T\left(\mathbf{I}-\mathbf{S}_l\right)\mathbf{X}_g\right]^{-1}\mathbf{X}_g^T\left(\mathbf{I}-\mathbf{S}_l\right)^T\left(\mathbf{I}-\mathbf{S}_l\right)$, $\hat{\sigma}^2=\dfrac{\mathbf{y}^T\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\mathbf{y}}{tr\left(\left(\mathbf{I}-\mathbf{S}\right)^T\left(\mathbf{I}-\mathbf{S}\right)\right)}$.

Reject H$_0$ if $\left|T_{g\_hit}\right|>t_{\alpha/2,df}$, where $df=\left[u_1^2\big/u_2\right]$.

To test the significance of a local variable $x_k\left(1\leq k\leq q\right)$ is used hypothesis as follows:

H$_0$: $\beta_k\left(u_i,v_i\right)=0$ (local variables x_k at the i-th location is not significant)

H$_1$: $\beta_k\left(u_i,v_i\right)\neq 0$ (local variables x_k at the i-th location is significant)

The test statistic is:

$$T_l=\frac{\hat{\beta}_k\left(u_i,v_i\right)}{\hat{\sigma}\sqrt{m_{kk}}} \tag{14}$$

where m_{kk} is the the k-th diagonal element of matrix $\mathbf{MM}^T$

and $\mathbf{M} = \left[\mathbf{X}_l^T \mathbf{W}(u_i, v_i) \mathbf{X}_l \right]^{-1} \mathbf{X}_l^T \mathbf{W}(u_i, v_i)\left(\mathbf{I} - \mathbf{X}_g \mathbf{G}\right)$

Reject H_0 if $\left|T_{l\_hit}\right| > t_{\alpha/2, df}$, where $df = \left[u_1^2 / u_2 \right]$.

5.1. Percentage of Poor Households in Mojokerto

As an initial step for analysis MGWR model, it is necessary to set up a global regression and GWR. With the global regression analysis was obtained the ANOVA table in Table 1. Since the p-value less than 0.05 then it can be concluded that the predictor variables simultaneously affect the response variables.

Table 1: Analysis of Variance of Global Regression Model

Source of Variation	Sum of Squares	Degree of Freedom	Mean of Squares	F	p-value
Regression	12383,18	14	884,51	15,715	0,000
Error	16265,90	289	56,28		
Total	**28649,09**	**303**			

Since there are differences in the data unit, the first standardized predictor variable (variable Z) it minus the mean and divided by the standard deviation. Using the Stepwise Regression at α 5% was obtained eight significant parameters namely β_0, β_4, β_6, β_8, β_9, β_{10}, β_{11} and β_{13} so the regression model to percentage of poor households in Mojokerto regency in 2008 is:

$$\hat{y}_i = 24{,}1588 - 0{,}9641\mathrm{Z}_{4i} - 1{,}3271\mathrm{Z}_{6i} + 1{,}5745\mathrm{Z}_{8i} + 4{,}3165\mathrm{Z}_{9i} - 2{,}0483\mathrm{Z}_{10i} - 2{,}6712\mathrm{Z}_{11i} - 1{,}0832\mathrm{Z}_{13i}$$

Test the assumption of homogeneity variance using the Breusch-Pagan Test and was obtained BP values 59.0302 with p-value is 0.000 indicates that spatial heterogeneity case. Therefore it would be better if it used a model that accommodates the observation that the location factor with GWR model.

Then performed modeling using the GWR model. The first step to building a GWR model is to determine the geographic location of each village/villages in Mojokerto, then choose the optimum bandwidth to form the weighting matrix at each observation. With four weighting function is obtained that the GWR model with a gaussian kernel function weighted GWR model is best because it has the smallest AIC.

Testing the suitability model to analyze the differences GWR done with the residual sum of squares GWR model and global regression models. Based on Table 2 obtained F value of 4.6842 with a p-value of 0.000. By using a significance level (α) of 5% we can conclude that GWR models differ significantly from the global regression model.

Table 2: Goodness of fits of GWR Model with *Gaussian* Weighting

Source of Error	Sum of Squares	Degree of Freedom	Mean of Squares	F	p-value
Improvement	8913,5359	59,4186	150,0127	4,6842	0,0000
GWR	7352,3685	229,5814	32,0251		
Regression	16265,9044	289,0000			

To see any predictor variables that affect different at each observation site, we can use a partial test of the influence of geographical factors for each predictor variable. Table 3 shows that by using a significance level $(\alpha)$of 5% it can be concluded that the variables that affect locally is the distance from the village to the capital district (X_5), distance from the village to the district capital / other cities nearby (km) (X_7), ratio of health per 100 inhabitants (X_{10}), percentage of families who subscribe to the telephone wiring (X_{11}), percentage of farm families (X_{13}), and the percentage of severely malnourished in the last 3 years (X_{14}). Meanwhile, eight other predictor variables remain influential globally at all locations of observation.

Table 3: Geographical Test of Predictor Variables GWR model with *Gaussian* Weighting

Parameter	β_0	$\beta 1$	β_2	β_3	β_4	β_5	β_6	β_7
F	24,8967	1,0470	0,8103	1,3909	1,1579	2,9969	1,9418	3,3690
p-value	0,0000*	0,3856	0,5180	0,2489	0,3302	0,0119*	0,0870	0,0063*
Parameter	β_8	β_9	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	
F	0,7190	15,275	29,921	36,611	0,6791	246,993	31,405	
p-value	0,6089	0,1876	0,0153*	0,0033*	0,6171	0,0000*	0,0118*	

Note: * significan at $\alpha = 5\%$

Modeling MGWR with four weighting functions obtained that MGWR model with a gaussian kernel function weighted is best model because it has the smallest AIC. Testing suitability MGWR model by analyzing differences in the amount of residual squares of global regression model and the MGWR model. Based on Table 4 obtained F value of 5.0403 with a p-value of 0.000. By using a significance level (α) of 5% it can be concluded that the models differ significantly MGWR with global regression model.

Table 4: Goodness of fits of MGWR Model with *Gaussian* Weighting

Source of Error	Sum of Squares	Degree of Freedom	Mean of Squares	F	p-value
Improvement	10030,5645	69,9210	143,4557	5,0403	0,0000
MGWR	6235,3399	219,0790	28,4616		
Regression	16265,9044	289,0000			

Testing the influence of global variables simultaneously performed using the residual sum of squares difference between MGWR model and MGWR model without using global variables. Table 5 shows that the F test statistic value of 2.2605 with a p-value of 0.0244. By using a significance level (α) of 5% it can be concluded that the global predictor variables simultaneously affect the modeling of the percentage of poor households in Mojokerto. While the significance of global variables can be partially seen in Table 6 with the conclusion that global variables which have a significant predictor is the ratio of basic health facilities (PHC/Sub Health Center/the doctor or midwife / village health post / polindes) per 100 inhabitants (X_9).

Tabel 5: Test the global parameter in the MGWR model with *Gaussian* Weighting simultaneously

Source of Error	Sum of Squares	Degree of Freedom	Mean of Squares	F	p-value
Improvement	435,5111	6,7693	64,3363	2,2605	0,0244
MGWR	6235,3399	219,0790	28,4616		
Reduced	6670,8510	225,8483			

Tabel 6: Test the local parameter in the MGWR model with *Gaussian* Weighting partially

Global Parameter			
Variable	**Beta**	**t**	**p-value**
Z_1	-0,4848	-0,6259	0,2660
Z_2	-0,0290	-0,0632	0,4748
Z_3	0,0114	0,0282	0,4888
Z_4	0,0005	0,0012	0,4995
Z_6	0,0465	0,0556	0,4779
Z_8	0,3640	0,7764	0,2191
Z_9	1,8311	2,9351	0,0018*
Z_{12}	-0,4099	-1,0887	0,1387

Note: * significan at $\alpha = 5\%$

Testing the influence of local variables simultaneously performed using the residual sum of squares difference between MGWR model and MGWR model without using local variables. Table 7 shows that the F test statistic value of 86.7455 with a p-value of 0.0000 so that it can be concluded that the local predictor variables simultaneously affect the modeling of the percentage of poor households in Mojokerto. While the summary statistics of the locally generated parameters shown in Table 8.

Tabel 7: Test the local parameter in the MGWR model with *Gaussian* Weighting simultaneously

Source of Error	Sum of Squares	Degree of Freedom	Mean of Squares	F	p-value
Improvement	189911,4236	76,9210	2468,9152	86,7455	0,0000
MGWR	6235,3399	219,0790	28,4616		
Reduced	196146,7635	296,0000			

Tabel 8: Statistic of Local Parameter MGWR using *Gaussian* Weighted

Parameter	Min	Max	Mean	Range	Standard deviation
β_0	18,0175	32,9390	24,1410	14,9214	3,6897
β_5	-3,9354	3,5304	-0,3740	7,4658	1,4367
β_7	-1,5012	5,5837	1,9873	7,0849	1,3851
β_{10}	-3,0990	0,1861	-1,1655	3,2850	0,7393
β_{11}	-7,2204	3,8269	-1,6150	11,0472	2,2710
β_{13}	-18,7926	2,3739	-2,9945	21,1665	5,6844
β_{14}	-3,7460	1,3521	-0,4373	5,0981	0,9991

The selection of the best model is done by using the AIC criterion. Table 9 shows the comparison of the global regression model with the GWR model and MGWR model either by using a weighted gaussian, exponential, bisquare and tricube. Based on Table 9 is obtained that MGWR model by using weighted gaussian kernel function is better used for modeling the percentage of poor households in Mojokerto because it has the largest R^2 value with the smallest MSE and AIC.

Tabel 9: Model Comparison

	Model	MSE	R^2	AIC
Global Regression		56,2834	0,4322	2106,4689
	Gaussian	32,0251	0,7434	1976,1832
GWR	Exponential	35,5844	0,6922	1991,2417
	Tricube	34,6502	0,6936	1988,0116
	Bisquare	33,6916	0,7056	1984,4697
MGWR	Gaussian*	28,4616	0,7824	1956,0853
	Exponential	35,5443	0,6695	1978,1935
	Tricube	33,6822	0,6866	1967,4229
	Bisquare	33,2043	0,6966	1964,7980

Note : * Best Model

5. Summary and Concluding Remarks

From the results of data analysis and discussion above the conclusions can be obtained as follows MGWR model is a modeling method that combines the global regression model with a local regression model (GWR). The assumption is normally distributed errors with mean zero and constant variance. Estimation of the MGWR model parameters using the Weighted Least Square (WLS) approach. Testing equality of regression models of global and MGWR approximated by the distribution F. Testing global parameters and local parameters are also approximated by a distribution F, while a partial test of the model parameters using the t test. The selection of the best models on MGWR model using AIC. MGWR model by using gaussian weighted function is better used to analyze the percentage of poor households in the District of Mojokerto in 2008 because it had the largest R2 value with the smallest MSE and AIC.

MGWR model can be used to analyze spatial data, therefore this model should also be applied in some other cases are more in line with the characteristics of the model MGWR eg agriculture, transportation, urban planning and others.

References

[1] BPS, 2005, *Identification and Determination Villages 2002*, Central Bureau of Statistics, Jakarta, Indonesia.

[2] Dimulyo, S., 2009, The use of geographically-Weighted Regression Kriging for Classification backward villages, *Proceedings of National Seminar on Application of Information Technology 2009*, Yogyakarta, pp. D 71-77.

[3] Fotheringham, A.S., Brunsdon, C., & Charlton, M. 2002, *Geographically Weighted Regression*, Jhon Wiley & Sons, Chichester, UK.

[4] Mei, C.L., Wang, N., & Zhang, W.X., 2006, Testing the importance of the explanatory variables in a mixed geographically weighted regression model, *Environment and Planning A*, vol. 38, pp. 587-598.

[5] Pecci, F. & Sassi, M., 2008, A Mixed Geographically Weighted Approach to Decoupling and Rural Development in the EU-15, *Seminar EAAE to 107 "Modelling of Agricultural and Rural Development Policies"*, Sevilla, Spanyol.

[6] Rencher, A.C., 2000, *Linear Model in Statistics*, John Wiley & Sons Inc, Singapore.

[7] Chasco, C., Garcia, I., & Vicens, J. 2007, *Modeling Spastial Variations in Household Disposible Income with Geographically Weighted Regression,* Munich Personal RePEc Arkhive (MPRA) Working Papper No. 1682.

[8] LeSage, J.P., 2001, A Family of Geographically Weighted Regression, Departement of Economics University of Toledo.

[9] Leung, Y., Mei, C.L., & Zhang, W.X., 2000a, Statistic Tests for Spatial Non-Stationarity Based on the Geographically Weighted Regression Model, *Environment and Planning* A, 32 pp. 9-32.

[10] Nakaya, T., Fotheringham, A.S., Brunsdon, C., & Charlton, M., 2005, Geographically Weighted Poisson Regression for Disease Association Mappin, *Statistics in Medicine*, Volume 24 Issue 17, pp. 2695-2717.