The 3rd International Seminar On EDUCATION and TECHNOLOGY – ISET
Collaborative Graduate Schools Conference

GEOGRAPHICALLY WEIGHT REGRESSION APPLICATIONS FOR SPATIAL ANALYSIS OF INEQUALITY IN CENTRAL JAVA

Lia Miftakhul Janah[1], M. Saifudin Nur[2], Tiani Wahyu Utami[3]
[1,2,3] **University of Muhammadiyah Semarang, Jl. Kedungmundu Raya No.18, Semarang 50273, Central Java Indonesia**

[*]**Corresponding author email : liamiftakhul7@gmail.com, saipungdekirby@gmail.com, tianiunimus@gmail.com**

Abstract

Inequality is a state where there is an imbalance between each other. Inequality indicates the unevenness of development that runs in an area. In Central Java, the problem of inequality among people still exists in daily life. Geographically Weight Regression method is a method that yields model parameter estimators that have localized properties at each point or location. In this study aims to modeling the inequality problem that occurred in Central Java using Geographically Weight Regression method that has the nature of localization at the point. Data taken from Central Statistics Agency (BPS) 2015. Through Geographically Weight Regression method can be concluded that with OLS method got 2 variables effect on imbalance with α 10% is variable of HDI (IPM) and PDRB. While the influential GWR method is the number of population and the amount of labor. While goodness of fit test showed there is no difference between GWR model and OLS model or in other words there is no spatial effect in the imbalance analysis in Central Java Province (0.4976 <0.1).

Keywords: inequality, GWR,Spatial

1. Introduction

The problem of inequality is one of the national problems that are still being attempted to overcome. The existence of inequality indicates that there is still no development in the region and economic growth is not yet qualified. Measurement of inequality using gini ratio data. The gini ratio is an inequality or aggregate inequality (overall) whose numbers range from zero (perfect equalization) to one (perfect inequality).

The gini ratio in Indonesia in March 2016 is 0.39 (Indonesian BPS, which means 1% of people owns 39% of wealth, while 61% is shared evenly with 99% others) It is clearly the government's job to maintain the gini ratio figure on target or even Smaller due to the many possible economic factors in 2017 that will affect the gini ratio.In Central Java the ratio of gini year 2016 amounted to 0.37.Gini ratio ratio is decreased from the year 2015 at 0.38.It means that in the middle of Java improvement Economic equity.

Berbagai penelitian yang berkaitan dengan angka ketimpangan (rasio gini) telah dilakukan, diantaranya Ketimpangan Distribusi Pendapatan dan Faktor-faktor yang Mempengaruhinya di Kabupaten Purbalingga (Badriah *et al.*, 2006), Analisis Faktor-faktor yang mempengaruhi ketimpangan Distribusi Pendapatan di Pulau Jawa (Nurlaili, 2016), Analisis Pengaruh Ketimpangan Distribusi Pendapatan Terhadap Jumlah Penduduk Miskin di Provinsi Jawa Tengah Periode 2000-2007 (Putra, 2011). Penelitian¬penelitian tersebut tidak menekankan aspek humaniora seperti karakteristik ragam budaya yang direpresentasikan karakteristik daerah (kabupaten/kota) masih terbatas untuk dikaji. Oleh karena itu dalam penelitian ini akan dikembangkan pemodelan rasio gini yang mengakomodasi adanya aspek perilaku masyarakat yang direpresentasikan dalam spasial (lokasi).

Various studies related to the rate of inequality (gini ratio) have been done, including Income Distribution Inequality and Its Affecting Factors in Purbalingga District (Badriah et al., 2006), Analysis of Factors Affecting Inequality of Income Distribution in Java (Nurlaili , 2016), Influence Analysis of Income Distribution to the Number of Poor People in Central Java Province Period 2000-2007 (Son, 2011). These studies do not emphasize aspects of humanities such as the characteristics of cultural diversity that are represented by regional characteristics (districts / cities) are still limited to be studied. Therefore, in this research will be developed gini ratio modeling that accommodate the aspect of society behavior which is represented in spatial (location).

Based on the above explanation will be conducted a study that examines the gini ratio in Central Java and the factors suspected to influence it by taking into account the spatial aspects. In addition, the use of GWR model is expected to produce gini ratio model in each city / district so as to provide positive information and input for the government in overcoming the problems of imbalance in Central Java.

1.1 GWR (Geographically Weight Regression)

The Geographically Weighted Regression (GWR) method is a regression model developed by Brunsdon et al. (2002) for continuous response variables that consider the location aspect. The GWR model is a local linear regression model that generates a model parameter estimator that has localized properties at each point or location. In the GWR model the values of different parameter estimator at each point of their geographical location, because each parameter value is calculated at each point of geographical location. GWR model can be written that is (Brunsdon et al., 2002)

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \;; \qquad i = 1, 2, \ldots, n \qquad (1.1)$$

keterangan :

y_i : Observation value of response variable at location i

x_{ik} : Observed value of predictor variable k at location i

$\beta_0(u_i, v_i)$: Value intercept GWR regression model

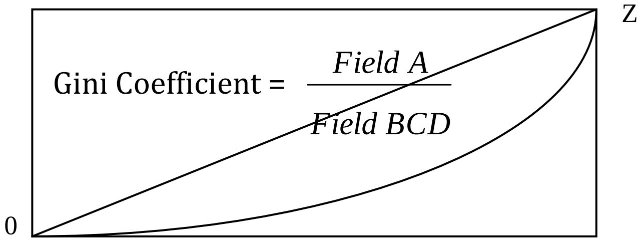
$\beta_k(u_i, v_i)$: Regression parameters for each i-location

(u_i, v_i) : Point of coordinates (latitude, longitude) at location i

ε_i : I-ass error assumed iidn (identical, independent, and normally distributed) with zero average and constant variance σ^2.

1.2 Inequality

Inequality is a state where there is an imbalance between each other. Inequality indicates the unevenness of development that runs in an area. Measurement of inequality using gini ratio data. The Gini ratio is a measure of inequality or aggregate inequality (overall) whose numbers range from zero (perfect equalization) to one (perfect inequality). The Gini ratio is an indicator for assessing an imbalance. The Gini coefficient can be obtained by calculating the ratio of the plane between the diagonal line and the Lorenz curve divided by the area of the half of the plane in which the Lorenz curve is located.



Lorenz Curve and Gini Coefficient

The further the Lorenz curve line distance from the diagonal line, the higher the degree of inequality. Instead the closer the Lorenz curve distance from the diagonal line, the higher the level of its equalization. In the picture above, the magnitude of inequality is described as a shaded area.

The level of gini ratio (inequality) by Lincolin Arsyad, 1997 is divided into 3 namely:

1. Low level of inequality 0.20 - 0, 35
2. Medium degree of inequality 0.36 - 0.49
3. High Inequality Rate 0.50 - 0.70

If the value of gini ratio close to 0 means the inequality is getting smaller. Conversely, when close to 1 means greater inequality.

2. Methods

2.1 Data Source

The main data source used in this study is secondary data sourced from the Central Statistics Agency. The data collected are 35 districts / cities in Central Java by 2015.

2.2 Research Variabel

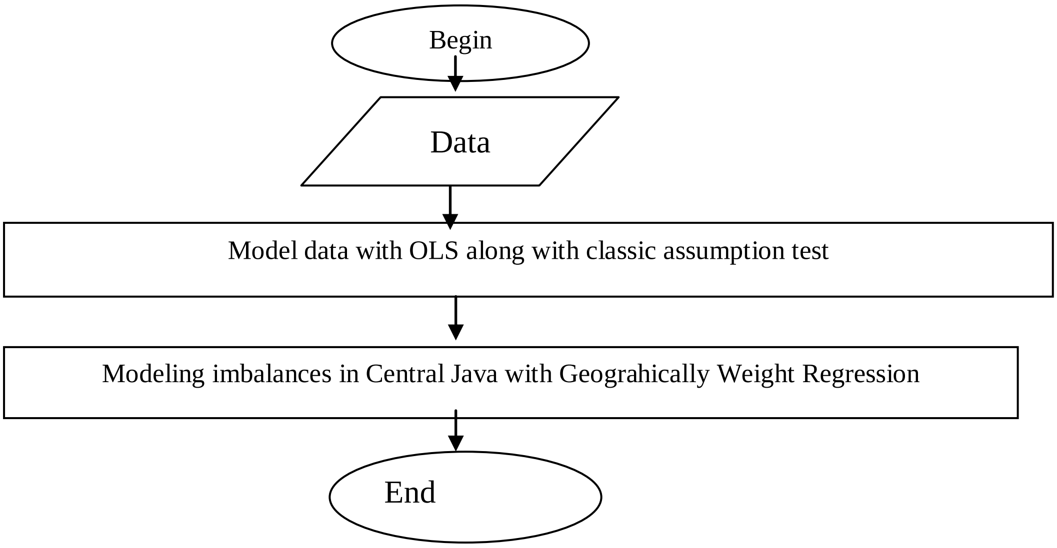
Table 2.2.1 *Dependent* Variabel and *Independent* Variabel

No	Variabel Name	Jenis variabel	Sumber
1	Gini Ratio	Dependen	BPS
2	Average per capita expenditure	Independen	BPS

3	Human Development Index	Independen	BPS
4	Gross Regional Domestic Product	Independen	BPS
5	Total population	Independen	BPS
6	Number of Poor People	Independen	BPS
7	Total employee	Independen	BPS
8	Unemployment Rate	Independen	BPS
9	Economic Growth Rate	Independen	BPS

2.3 Procedures (or research design)

The analysis steps in this study can be described in the flow diagram as shown in Figure 2.1, as follows:



3. Results

3.1 **Modeling Inequality using OLS**

The resulting OLS functionality:

$\hat{Y}$ = 0.00000081-0.7281X_1+0.8071X_2+0.2685X_3-1.476X_4+0.3682X_5+1.009X_6-0.2498X_7-0.2987X_8

Testing simultaneously OLS regression model:

$H_0 : \beta_1 = \beta_2 = ... = \beta_6 = 0$

H_1 : At least there is one where i$\beta_i \neq 0$

Obtained value of F statistics = 1.484 with p-value = 0.21 so that if α = 10% then the decision taken is accept H0. In other words, at 90% confidence level, the predictor variable used does not have a significant effect simultaneously on the response variable

Partial testing for the OLS regression model:

Hypothesis

$H_0 : \beta_i = 0$

$H_1 : \beta_i \neq 0$

Table 3.1.1 Partial Test Analysis Result OLS regression model

Predictor	**Coefficient**	**P-value**	**Decision**	**Conclusion**
Konstanta	0.00000081	1.000	Accept H0	not significant
X_1	-0.7281	0.3248	Accept H0	not significant
X_2	0.8071	0.0894	reject H0	significant
X_3	0.2685	0.0592	reject H0	significant
X_4	-1.476	0.4481	Accept H0	not significant
X_5	0.3682	0.3801	Accept H0	not significant
X_6	1.009	0.4660	Accept H0	not significant

X_7	-0.2498	0.3593	Accept H0	not significant
X_8	-0.2987	0.2066	Accept H0	not significant

In modeling using OLS there are 4 classical assumptions that must be met that is normality, not multicolinearity, not heteroscedasticity, not autocorrelation. The results of the classical assumption test OLS model is as follows:

Table 3.1.2 Classical assumption test results OLS model

Assumption	**P-value**	**Conclusion**
Normalitas	0.5732	fulfilled
Homocedasticity	0.3047	fulfilled
No autocorrelation	0.3556	fulfilled

Table 3.1.3 Results of the OLS Model multicollinearity test

Variabel	**X_1**	**X_2**	**X_3**	**X_4**	**X_5**	**X_6**	**X_7**	**X_8**
VIF	6.4494	6.3382	3.424339	81.888967	5.897719	70.476066	2.013733	3.982789
Conclusion	fulfilled	fulfilled	fulfilled	Not fulfilled	fulfilled	not ulfilled	fulfilled	fulfilled

It is seen that there are 6 independent variables that are not significant to the response variable at 10% confidence level. In the classical assumption test found that the model is not met multicollinearity test. OLS regression model has a poor performance because there is multicollinearity. Therefore it is necessary to do spatial modeling using geographically weighted regression (GWR).

3.2 Modeling Inequality using GWR

Estimation of GWR model parameters was obtained by inserting spatial weights in the calculations using the weighted least squares (WLS) method. The estimation results of GWR model parameters are presented in Table 3.2:

Table 3.2.1 Estimation Result of GWR Model Parameter

Parameter	p-value	Conclusion
Intercept	0.14262	not significant
X_1	0.86256	not significant
X_2	0.71056	not significant
X_3	0.99625	not significant
X_4	0.04078	significant
X_5	0.92123	not significant
X_6	0.09503	significant
X_7	1.000	not significant
X_8	0.71056	not significant

After obtaining parameter estimation for regression model of OLS and GWR then tested suitability of model (goodness of fit) to see whether geographical factor influence to imbalance in Central Java. The form of the hypothesis is as follows.

H_0 : There is no significant difference between the GWR model and the OLS model

H_1 : There is a significant difference between the GWR model and the OLS model.

This test is done by using Leung test and F test statistic. The result of parameter test by using software R is as follows:

Table 3.2.2 Test of Conformity Model with Software R

Model	**SSR**	**F-statistic**	**P-value**
Ols	23.34320	-	-
GWR	22.64124	0.9975	0.4976

4. Discussion

From the result of modeling with GWR method that there are two predictor variables that have significant effect on response Y (inequality) that is X4 (Total Population) and X6 (Total Labor). It can be seen from the model obtained p-value of test f of 0.4976, which is greater than 0.1 so that the decision taken is the failure to reject H0 and the conclusion there is no significant difference between the GWR model and the OLS model or in other words the predictor variables generally are not Influence spatial response variables (Inequality). However, the residual GWR model has a smaller value than the residual value of OLS (22.64124 <23.34320) so that the GWR model is better than the OLS model.

5. Conclusions

In the case of inequality in Central Java with OLS regression predictor variables that have significant effect are HDI and PDRB. The result of model conformity test stated that there is no significant difference between GWR model and OLS model or in other words generally predictor variables do not influence the response variable (inequality in Central Java) spatially. However, the residual value of GWR model is better than OLS model.

6. References

Anselin, L. & Bera, A. 1998. *Spatial Dependence In Linear Regression Models With An Introduction To Spatial Econometrics. Dalam Ullah, A. & Giles, D. Handbook of Applied Economics Statistic*. Selected Reading, hlm. 237-289. New York: Marcel Dekker.

Badriah Lilis S, S Herman, S Irma. 2006.*Ketimpangan Distribusi Pendapatan dan Faktor-faktor yang mempengaruhinya di Kabupaten Purbalingga. Eko-Regional, Vol 1, No 2, September 2006*

Brunsdon C., Fotheringham, A. S., dan Charlton, M. E. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship.* England: John Wiley and Sons Ltd

Intan. 2010. *Pemodelan Jumlah Penderita Tuberculosis (TB) di Propinsi Jawa Timur Tahun 2010 dengan Mengggunakan Metode Geographically Weighted Regression.* Final Project Statistics of FMIPA ITS, Surabaya.

Nurlaili Ani. 2016. *Analisis Faktor-faktor yang Mempengaruhi Ketimpangan Distribusi Pendapatan di Pulau Jawa tahun 2007-2013.* Final ProjectEconomics Education of Economy Faculty UNY,Yogyakarta.

Putra Linggar D. 2011. *Analisis Pengaruh Ketimpangan Distribusi Pendapatan Tterhadap Jumlah Penduduk Miskin di Provinsi Jawa Tengah Periode 2000-2007.* Final Project IESP of Economy Faculty UNDIP, Semarang.

Saefuddin A, Nur Andi S, Noer Azam A. 2011. On Comparisson between Ordinary Linear Regression and Geographically Weighted Regression: With Application to Indonesian Poverty Data. *European Journal of Scientific Research. Euro Journals Publishing, Inc.*