Forum Statistika dan Komputasi : *Indonesian Journal of Statistics* available onlir
ISSN : 0853-8115 *journal.ipb.ac.id/index.php/stati*
Vol. 20 No.2, Oktober 2015, p: 96-104

RIDGE AND LASSO PERFORMANCE IN SPATIAL DATA WITH HETEROGENEITY AND MULTICOLLINEARITY

Tiyas Yulita[1], Asep Saefuddin[1], Aji Hamim Wigena[1]

[1]Department of Statistics, Bogor Agricultural University, Indonesia
E-mail : yulitatiyas@gmail.com

ABSTRACT

Spatial heterogeneity becomes a separate issue on the analysis of spatial data. GWR (Geographically Weighted Regression) is a statistical technique to explore spatial nonstationarity by form the differrent regression models at different point in observation space. Multicollinearity is a condition that the independent variables in model have linear relationship. It would be a problem for estimation parameters process, because that condition produces unstable model. This problem may be found in GWR models, which allow the linear relationship between independent variables at each location called local multicollinearity. GWRR (Geographically Weighted Ridge Regression) and GWL (Geographically Weighted Lasso) which use the concept of ridge and lasso is shrink the regression coefficient in GWR model. GWRR and GWL techniques are consider to be capable of overcoming local multicollinearity to produce more stable models with lower variance. In this study, GWRR and GWL is used to model Gross Regional Domestic Product (GRDP) in Java using kernel exponential weighted function. The results showed that GWL has better performance to predict GRDP with lower RMSE and higher R^2 value than GWRR.

Keyword : Spatial Heterogeneity, GWR, Local Multicollinearity, Ridge, Lasso

INTRODUCTION

Statistical methods are often used as a tool to determine the relationship between variables by establishing an appropriate model in describing the characteristics of the data. As in the linear regression model that is able to describe the relationship between independent variables with the response variable. The spatial statistical methods accomodate the relationships between variables in spatial data. Spatial statistical methods are used to model some specific variables at each different geographic locations so that diversity which is caused by the differences in the characteristics of the region (spatial heterogeneity) on the data can be resolved. The spatial heterogeneity can be caused by several things such as differences in geography, culture , ecnomic policies that vary in each region. If ―Ordinary Least Squares (OLS) was used to estimate the parameters of the linear regression analysis, the estimated parameters would have a high variance. One method that can be used to identify the presence of spatial heterogeneity is Geographically Weighted Regression (GWR). GWR is a method that is quite effective in parameter estimation on the data with spatial heterogeneity (Fotheringham et al. 2002).

Another problem that may arise in the modeling of more than one independent variables is multicollinearity. The multicollinearity is caused by the linear relationship was almost perfect (near dependence) on the columns of the matrix **X**, and if the linear relationship is perfect, that would lead to the $|\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}| = 0$ and this condition is called exact multicollinearity (Draper and Smith 1998). If this condition is not solved, the estimated parameters obtained would become unstable. As in modeling with multiple linear regression, multicollinearity problems can also be found on the spatial regression called local muticollinearity. There are several methods to overcome multicollinearity in multiple linear regression modeling including ridge regression and lasso. Lasso performs estimation with LARS algorithm (Least Angle Regression) that shrinks the estimated coefficients to zero. While the ridge regression adds a positive bias coefficient on parameter estimation process that shrinks the coefficient to zero, so the results are biased but has a low variance. In the spatial regression, local multicollinearity can be resolved by adopting the methods of the ridge regression and lasso into the GWR. Geographically Weighted Lasso (GWL) is a method that adopts the concept lasso GWR in estimating the parameters for addressing cases of local multicollinierity so as to obtain stable estimates for the parameters. In this study, the GWRR and GWL models is used to examine heterogeneity and multicollinearity problems for the Gross Regional Domestic Product (GRDP) in 2010 of 113 districts/cities in Java.

1.1 Geogaphically Weighted Regression (GWR)

Spatial heterogeneity problem can be addressed effectively by predicting the point by using GWR. Fotheringham et al. (2002) adopted the concept that exists in the linear regression model be weighted regression model to establish models of GWR as follows:

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

where $y_i$is a dependent variable at location (u_i, v_i), x_{ik} is independent variables at location (u_i, v_i) , β_k is local parameter estimate at each location. Local parameters of GWR model is estimate with derivative process of equation (1) by $\boldsymbol{\beta}^T(u_i, v_i)$ and produce the following function :

$$\boldsymbol{\beta}(u_i, v_i) = [\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{y} \qquad (2)$$

where $\boldsymbol{W}(u_i, v_i)$ is a weighted matrix at location (u_i, v_i).

The weighted matrix on the GWR calculated from a kernel function that puts more emphasis on observations closer to the location of the i^{th} observation. In this study, the kernel function is used to form the weighted matrix is the exponential kernel which has fuction as follows:

$$w_j(u_i, v_i) = exp\left(-\frac{d_{ij}}{h}\right) \qquad (3)$$

h is a *bandwidth* at (u_i, v_i) location, this is obtained from the Euclidean distance between the location of the observations with other locations, so that the region is still influenced by the surrounding neighbors within less than h. Estimating h is conducted using Cross Validation (CV) over locations, the CV has fuction as follows:

$$CV(h) = \sum_{i=1}^{n}[y_i - \hat{y}_{\neq i}(h)]^2 \qquad (4)$$

$\hat{y}_{\neq i}(h)$ is fitted value of $y_i$by eliminating the observation point i^{th} location on the prediction and h will be obtained with the iterative process to obtain the minimum CV.

1.2 Ridge Regression

Ridge regression is used to control the instability of least squares estimators arising due to multicollinearity (Hoerl and Kennard, 2000). The ridge regression parameter is estimate by minimizing the sum square of errors which added a constraints on squares that shrink coefficient close to zero. Specifically, the coefficient of ridge estimators obtained by minimizing the following equation:

$$\hat{\beta}_R = \sum_{i=1}^{k}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 \qquad (5)$$

with $\sum_{j=1}^{p}\beta_j^2 \leq \rho$, where ρ is controls the amount of shrinkage. The estimate parameters is obtained with derivative process of the sum square of error by $\hat{\beta}$:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{R}} = (\boldsymbol{X}^T\boldsymbol{X} + \mathrm{c}\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \qquad (6)$$

where $\boldsymbol{I}$ is an identity matrix with size $p \times p$, and c is a positive bias coefficient.

1.3 Least Absolute Shrinkage and Selection Operator (LASSO)

Least Absolute Shrinkage and Selection Operator (LASSO) was first introduced by Tibshirani in 1996. Estimates of the coefficient of lasso parameters can not be obtained in closed form as in OLS or ridge regression, but by using quadratic programming (Hastie et al., 2009). Lasso is defined as follows:

$$(\hat{\beta}) = \arg\min \sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{k=1}^{p} x_{ik}\beta_k\right)^2 (7)$$

where $\sum_{k=1}^{p}|\hat{\beta}_k| \leq t$ is treshold parameter.

It is known that *t* is a quantity that controls the amount of shrinkage in the estimation of coefficients lasso with $t \geq 0$. If β_k is a least estimate coefficient for lasso and $t_0 = \sum_{k=1}^{p}|\hat{\beta}_k^0|$, so $t<t_0$ will both lead to the solution of Ordinary Least Square (OLS) shrinks to——zero, and can make some coefficients is equal to zeto. Solution for lasso is obtained by determining $s = t / \sum_{k=1}^{p}|\hat{\beta}_k^0|$ where $t = \sum_{k=1}^{p}|\hat{\beta}_k|$and $\hat{\beta}_k^0$ is a parameter estimators for the full model or written as$|beta|/max|beta|$.

Efron et al.(2004) has a solution for lasso problem by modificating LARS algorithm (Least Angle Regression). The lasso shrinkage parameters are defined as follows:

$$s = \frac{\sum_{k=1}^{p} |\hat{\beta}_k|}{\sum_{k=1}^{p} |\hat{\beta}^0{}_k|} \quad (8)$$

where s is shrinkage parameter between 0 to 1.

RESEARCH METHODOLOGY

The analysis steps of the research is :

a. Exploration of data to determine a general overview of the condition of the data.
b. Performed Breusch-Pagan test to detect heterogeneity on the data.
c. Apply GWR model to GRDP data using kernel exponential weighted function.
d. Detected the local multicollinearity on GWR models with calculate the VIF value for independent variables at each location.
e. Modelled the GRDP data use GWRR and GWL.
f. Compared the models with criterion of RMSE and R^2 on GWRR and GWL.

In this study, analysis of GWL and GWRR use a statistical software R 3.2 with GWmodel and gwrr package.

RESULTS

3.1 Data Exploration

This study used data of GRDP at 113 districts/cities in Java in 2010 as a dependent variable in billion rupiah (Y). The explanatory variables include the percentage of poverty (poor people) (X1), the number of families using electricity (X2), education (X3), the Human Development Index (X4), life expectancy (X5), the literacy rate (X6), the average length of education (X7), expenditure per capita (X8), percentage of villages using liquefied petroleum gas (X9), the number of markets (X10), the number of hotels and inns (X11) . Before modelling the data, exploration is needed to know about data condition. Descriptive variables are summarized in Table 1. The relationship between the dependent and independent variables can be seen from the value of the correlation coefficient. The correlation coefficients were obtained on Table 2.

Table 1. Descriptive statistics for each variables

Variables	Minimum	Maximum	Mean	Standard Deviation
Y	0,75	96,48	11,22	17,64
X1	1,670	25,220	12,837	5,359
X2	0,0340	1,1510	0,3275	0,1970
X3	0,183	1,333	0,5936	0,226
X4	62,940	79,520	72,523	3,404
X5	61,130	75,060	69,324	2,667
X6	76,720	99,670	92,959	5,225
X7	5,540	11,480	7,988	1,513
X8	614,83	652,80	637,39	8,03
X9	12,28	100	72,70	26,40
X10	239	55080	14938	8994
X11	4	1050	78,6	124,3

Table 2. The correlation coefficients between dependent variables and independent variables.

Variable	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
Y	-0,456	0,429	0,634	0,398	0,23	0,325	0,443	0,241	0,355	0,385	0,194
p-value	0,000	0,000	0,000	0,000	0,014	0,000	0,000	0,000	0,000	0,000	0,039

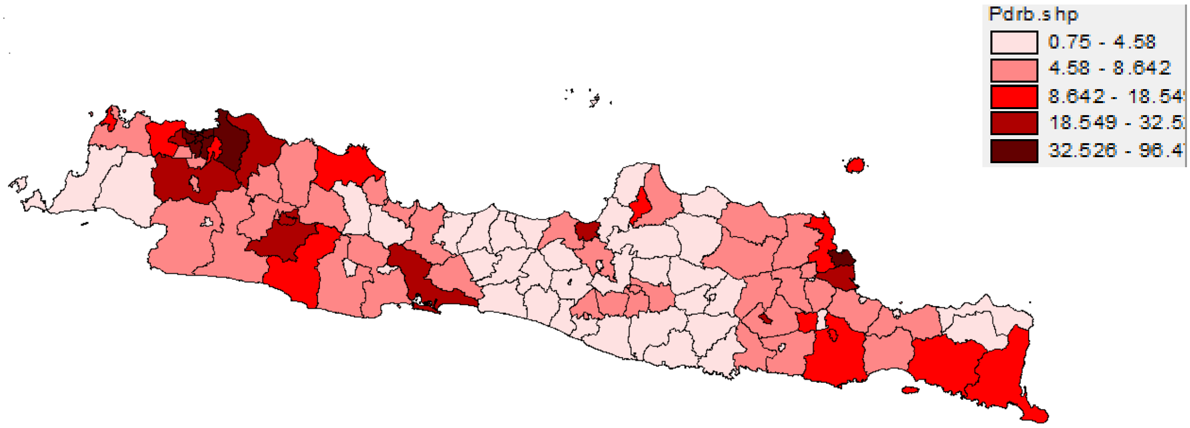

Figure 1. Map of GRDP in Java

From Table 2, we know that percentage of poverty (poor people) (X1) has negative correlation with GRDP, but other variables have positive correlation.

Figure 1 provided the information that GRDP in 113 districts/cities in Java have a heterogeneity of data distribution. Locations with high GRDP are cities around Jakarta and several other provincial capitals, thus the OLS method can not still be used to perform a data modeling with GRDP, due to the results obtained will have a large variance of parameters estimate. So we need a modeling method that captures the spatial heterogeneity to form nonstationary regression model in each location.

3.2 Geographically Weighted Regression (GWR)

Before modeling GWR, there is a way to determine the spatial heterogeneity between the location of observation use Breusch-Pagan test (Anselin 1988). The results were significant with p-value of 0,014 it means that there is a spatial heterogeneity between the location of observations. The results of parameters estimate in GWR model are summarized in the following table.

Table 3 informs that the mean value of the coefficients for X1 is negative corresponding to the correlation coefficient between Y and X1. But for mean of coefficient estimation at the X6, X7, X8, X9 and X11 have negative sign which opposite with correlation coefficient. The indicates for multicollinearity in independent variables can be known from VIF (Variance Inflation Factor) value of local observations.

Table 3. Minimum, mean and maximum value of coefficient estimate of GWR

Coefficient	Minimum	Mean	Maximum
$\hat{\beta}_0$	-269,682	-63,9524	36,7215
$\hat{\beta}_1$	-2,1748	-0,5999	0,4644
$\hat{\beta}_2$	-25,498	25,0572	102,9368
$\hat{\beta}_3$	11,6985	27,0232	43,0811
$\hat{\beta}_4$	0,9276	1,4209	2,0446
$\hat{\beta}_5$	-1,8559	0,4102	4,4610
$\hat{\beta}_6$	-0,7356	-0,2389	-0,0876
$\hat{\beta}_7$	-4,9464	-0,0355	3,1272
$\hat{\beta}_8$	-0,1690	-0,0812	0,0957
$\hat{\beta}_9$	-0,0732	-0,0207	0,0588
$\hat{\beta}_{10}$	-0,0010	0,0003	0,0013
$\hat{\beta}_{11}$	-0,0230	-0,0157	0,0090

Table 4. VIF value summary for each variables of over location observation

Variables	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
Minimum	1,98	7,01	1,92	20,81	6,97	6,86	6,68	2,68	2,02	7,44	1,22
Mean	2,43	8,86	2,94	289,03	61,11	23,25	42,46	11,21	2,94	8,60	1,54
Maximum	3,24	10,32	5,16	1410,59	321,13	59,90	174,91	42,95	4,66	9,84	1,94
VIF>10	0	9	0	113	81	71	76	35	0	0	0

VIF value in Table 4 shows that there are local multicollinearity between independent variables seen from the location which has VIF value greater than 10. The presence of multicollinearity could cause the results of coefficient parameter has a large variance. In addition, multicoliniearity will produce an unstable model.

3.3 Geographically Weighted Ridge Regression (GWRR)

GWRR is a developing methods of ridge regression for spatial regression. The difference between the ridge and GWRR is additional information of weighted. GWRR predict on the local parameters of the observation (u_i, v_i) by adding a weighted $w_j(u_i, v_i)$which can be written as follows:

$$w_j(u_i, v_i)Y = w_j(u_i, v_i)(\beta X + \varepsilon) \quad (9)$$

Estimation of the parameters on GWRR obtained with derivative of the sum of square error by $\hat{\boldsymbol{\beta}}_{\boldsymbol{R}}(u_i, v_i)$, thus obtained:

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{R}}(u_i, v_i) = (\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X} + c\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} \quad (10)$$

where $\boldsymbol{I}$ is a $p \times p$ identity matrix, c is positive coefficient bias, and $\boldsymbol{W}(u_i, v_i)$ is $n \times n$ weighted matrix for observation (u_i, v_i). $\boldsymbol{Y}$ and $\boldsymbol{X}$ is matrix of centering and scalling dependent and independent variables (Wheeler 2007).

GWRR modelling steps are i) Predict the bandwidth values to form the weighted matrix $\boldsymbol{W}(u_i, v_i)$, ii) Suppose that minimize bias coefficient CV, iii) Use predicted bandwidth and bias coefficient obtained in i and ii using equation (10). In GWRR, the value of c is obtained simultaneously with the iteration method for each h is selected by minimize value of CV, then use these results to estimate the coefficients on GWRR. Value of c which obtained from GRDP data for GWRR is 1,484 with bandwidth is 0,1390. The examples of results on five location in Jakarta is shown in Table 5.

Local model for South Jakarta is :

$$\begin{aligned}\hat{Y}^* = 0{,}1847 - 0{,}0856X_1^* + 0{,}0557X_2^* \\ + 0{,}0595X_3^* + 0{,}1882X_4^* \\ + 0{,}3618X_5^* + 0{,}2054X_6^* \\ + 0{,}1258X_7^* + 0{,}1677X_8^* \\ + 0{,}1712X_9^* + 0{,}1692X_{10}^* \\ - 0{,}0217X_{11}^*\end{aligned}$$

The model showed that the percentage of poverty in South Jakarta has a negative influence on GRDP in GWRR models, as well as on the number of hotels and inns. As for the other variables have a positive influence. Summary results of parameter estimation with GWRR shown in Table 6.

Table 5 show the results of the coefficient of GWRR model different from GWR. The coefficient of GWRR has smaller range than GWR, this indicates that the coefficient is shrinking at GWRR models. Of the average value of the estimate parameter for the variables percentage of poverty is negative in accordance with the correlation coefficient, as well as on other variables.

Table 5. Coefficient estimate of GWRR in Jakarta.

Coefficient	South Jakarta	East Jakarta	Central Jakarta	West Jakarta	North Jakarta
$\hat{\boldsymbol{\beta}}_0$	0,1847	0,2087	0,2726	0,2159	0,2728
$\hat{\boldsymbol{\beta}}_1$	-0,0856	-0,1294	-0,1543	-0,1164	-0,1572
$\hat{\boldsymbol{\beta}}_2$	0,0557	0,0259	-0,1023	-0,0225	-0,0877
$\hat{\boldsymbol{\beta}}_3$	0,0595	0,0664	0,0774	0,0673	0,0823
$\hat{\boldsymbol{\beta}}_4$	0,1882	0,1665	0,2074	0,1948	0,2048
Coefficient	**South Jakarta**	**East Jakarta**	**Central Jakarta**	**West Jakarta**	**North Jakarta**
$\hat{\boldsymbol{\beta}}_5$	0,3618	0,3309	0,3492	0,3219	0,3349
$\hat{\boldsymbol{\beta}}_6$	0,2054	0,1981	0,2649	0,2287	0,2611
$\hat{\boldsymbol{\beta}}_7$	0,1258	0,1078	0,1554	0,1402	0,1547
$\hat{\boldsymbol{\beta}}_8$	0,1677	0,1260	0,2159	0,2026	0,2117
$\hat{\boldsymbol{\beta}}_9$	0,1712	0,1943	0,2596	0,2075	0,2612
$\hat{\boldsymbol{\beta}}_{10}$	0,1692	0,1263	0,0829	0,1392	0,1145
$\hat{\boldsymbol{\beta}}_{11}$	-0,0217	-0,0004	0,1262	0,0967	0,1148

Table 6 Minimum value, mean and maximum value of coefficient estimate GWRR model

Coefficient	Minimum	Mean	Maximum
$\hat{\boldsymbol{\beta}}_0$	-0,1112	-0,0151	0,2911
$\hat{\boldsymbol{\beta}}_1$	-0,3090	-0,0419	0,0488
$\hat{\boldsymbol{\beta}}_2$	-0,1023	0,0681	0,3802
$\hat{\boldsymbol{\beta}}_3$	-0,1133	0,0842	0,3556
$\hat{\boldsymbol{\beta}}_4$	-0,0478	0,0366	0,2170
$\hat{\boldsymbol{\beta}}_5$	-0,1512	0,0382	0,3987
$\hat{\boldsymbol{\beta}}_6$	-0,1074	0,0309	0,2649
$\hat{\boldsymbol{\beta}}_7$	-0,0664	0,0397	0,1748
$\hat{\boldsymbol{\beta}}_8$	-0,0951	0,0306	0,2159
$\hat{\boldsymbol{\beta}}_9$	-0,1763	0,0302	0,2884
$\hat{\boldsymbol{\beta}}_{10}$	-0,0517	0,0780	0,3056
$\hat{\boldsymbol{\beta}}_{11}$	-0,1143	0,0582	0,4881

3.4. Geographically Weighted Lasso (GWL)

The concept of Least Absolute Shrinkage and Selection Operator (LASSO) that is applied in a GWR model who later became known as Geographically Weighted Lasso (GWL) is a method used to overcome spatial heterogeneity and local multicollinearity. GWL produce not biased coefficient estimate and efficiently so the prediction results obtained more accurate (Wheeler, 2009). Estimation parameters of the lasso conducted simultaneously so that the final solution depends on bandwidth kernel that has been previously before. In step to estimate the parameters in the model GWL, shrinkage (s) must be estimated before the final solution lasso. Estimation parameters in the model shrinkage lasso GWL was conducted using cross validation (CV), so there is a shrinkage parameter at each location. So every point of geographical location has a value of different regression.

Steps of GWL-local parameter estimation in this study as follows :

1. Estimate the optimum bandwidth kernel use Cross Validation (CV)
2. Form the $n \times n$ weighted matrixW.
3. For each locationi $= 1,2, \dots, n$.

- $W^{\frac{1}{2}}(i) = sqrt\left(diag\left(W(i)\right)\right)$
- $\boldsymbol{X_W} = \boldsymbol{W^{\frac{1}{2}}(i)X}$ and $\boldsymbol{y}_w = \boldsymbol{W}^{1/2}(i)\boldsymbol{y}$ for each location i

4. Call lars $(\boldsymbol{X_W}, \boldsymbol{y}_w)$ on R software, and save the serries of lasso solution
5. Save the solution of lasso that minimize CV value based on fraction from *shrinkage*(s_i) value and indicator **b**.

All of these steps was covered by gwrr package on R 3.2 with function gwl.est. As in modeling using the lasso, the regression coefficient at GWL will also be depreciated to zero over a given shrinkage. Thus, the coefficient of zero is certainly no effect on the model. Bandwidth values obtained from the iteration process using the CV in GWL is 0,09. The bandwidth and skrinkage value is then used to estimate the parameters of GWL. Table 7 gives the results of the model parameters GWL estimations at some locations in West Java.

Local model for Bogor is :

$$\hat{Y}^* = 0{,}0222 + 0{,}0683X_3^* + 0{,}9631X_5^* - 0{,}1729X_8^* + 0{,}1406X_9^* + 0{,}2400X_{10}^*$$

The model explains that the variables that affect GRDP in Bogor is education, life expectancy, expenditure per capita, percentage of village using using liquefied petroleum gas and number of market. So that the model is informed that the increasing educational facilities in Bogor, it will increase the value of GRDP. Similarly, if an increase in life expectancy, the percentage of village using liquefied petroleum gas and the number of market, it can increase the GRDP in Bogor. But if expenditure per capita of the population increases, it is not a positive influence on the increase of GRDP.

In addition to seeing the value of VIF, multicollinearity can seen by the differences of signs of regression coefficients from correlation coefficient. Different sign on GWR regression coefficients, GWRR, and GWL with the correlation coefficient is shown in Table 8.

Table 7 Coefficient estimate of GWL in West Java

Coefficient	**Bogor**	**Sukabumi**	**Cianjur**	**Bandung**	**Garut**
$\hat{\beta}_0$	0,0222	0	0	-0,0113	-0,1014
$\hat{\beta}_1$	0	0	0	-0,3381	-0,1745
$\hat{\beta}_2$	0	0	0	-0,2394	0
$\hat{\beta}_3$	0,0683	0	0,1106	0	0
$\hat{\beta}_4$	0	0	0	0	0,0543
$\hat{\beta}_5$	0,9631	0	0	0,3577	0,1758
$\hat{\beta}_6$	0	-0,1629	-0,1809	-0,0251	0
$\hat{\beta}_7$	0	0	0	0	0
$\hat{\beta}_8$	-0,1729	0	0	0,0800	0,0842
$\hat{\beta}_9$	0,1406	0	0	-0,0791	0,1004
$\hat{\beta}_{10}$	0,2400	0	0	0,5105	0,2137
$\hat{\beta}_{11}$	0	0	0,09704	0,1032	0,2552

Table 8. Number of locations observation with a sign (-) and (+) on each model coefficients

Model	Sign	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$	$\hat{\beta}_{11}$
GWR	(+)	**27**	73	113	113	50	0	69	10	37	76	8
	(-)	86	**40**	**0**	**0**	**63**	**113**	**44**	**103**	**76**	**37**	**105**
GWRR	(+)	**17**	103	108	81	53	85	90	76	85	108	93
	(-)	96	**10**	**5**	**32**	**60**	**28**	**23**	**37**	**28**	**5**	**20**
GWL	(+)	**23**	50	62	2	27	9	45	12	39	66	55
	(-)	41	**22**	**9**	**14**	**15**	**14**	**11**	**11**	**21**	**19**	**29**

Table 9 Comparison between models of GWR, GWRR, and GWR.

Model	H	RMSE	R^2
GWR	1,3430	8,8984	0,7431
GWRR	0,1390	0,3725	0,8599
GWL	0,09	0,2071	0,9567

Table 8 give an information that is not all the sign of coefficients same with coefficient correlations in locations observation. In estimate parameters of GWRR few parameters have opposite sign of the correlation coefficient, so the model GWRR considered more stable and closer to the real situation. Model GWL make some coefficients to zero, so it is not obtain the full model in the entire region. The comparison of GWR, GWRR, and GWL models can be seen from table below.

Table 9 showed the widest bandwidth coverage owned by GWR models with bandwidth value is 1,3430. That is another location that is still within the radius are considered still give effect to estimate the parameters at the point of observation location. While the bandwidth generated by GWRR and GWL only as far as 0,1390 and 0,09. The stability of the model in predicting the parameters can be seen from the resulting RMSE values. GWR models produce the highest RMSE value, this is because the GWR models still has local multicollinearity between its independent variables so the estimated parameters have large variance.The ability of variables in the model to explain the variance of GRDP can be seen from R^2 value. GWL has highest R^2 value it means that variables on GWL model can explain well the variance of GRDP.

CONCLUSSION

GWRR and GWL considered to be able to overcome problems due to multicollinearity in GWR model to form more stable regression model views from RMSE value. Based on a smaller RMSE and higher R^2 value, GWL model has better performance to predict GRDP of 113 Regency / City of Java in 2010. The development of GWRR and GWL is still needed to improve the capability of modeling GWRR and GWL which robust against outliers.

REFFERENCES

Anselin L. 1988. *Spatial Econometrics : Methods and Model*. Dordrecht: Kluwer Academic Press.

[BPS]. Badan Pusat Statistik. 2014. *Produk Domestik Regional Bruto Kabupaten/Kota di Indonesia 2009-2010.*Jakarta: BPS.

[BPS]. Badan Pusat Statistik. 2014. www.bps.go.id.

Draper NR, Smith H. 1998. *Applied Regression Analysis.* Edisi ke-3. New York (US) : John Wiley & Sons..

Efron B, Hastie T, Johnstone I, Tibshirani R. 2004. Least Angle Regression. *The Annals of Statistics* 32(2), 407-451.

Fatulloh 2013. Penerapkan metode Regresi Terboboti Geografis (RTG) untuk data Produk Domestik Regional Bruto (PDRB) di Pulau Jawa tahun 2010. [skripsi]. Institut Pertanian Bogor. Bogor

Fotheringham AS, Charlton M. 2002. *Geographically Weighted Regression the Analysis of Spatially Varying Relationship*. New York: John Wiley and Sons, Inc.

Hastie T, Tibshirani R, Friendman J. 2009. *The Elements od Statistical Learning DataMining, Inference, and Prediction*. New York: Springer

Hoerl AE, Kennard RW. 2000. Ridge Regression : biased Estimation for NonortogonalProblems. *Technometrics* 12 , 80-86.

Sukmantoro D. 2014. *Geographically Weighted Ridge Regression* dalam Pemodelan Nilai Tanah. *Journal-Institut Teknologi Sepuluh Nopember*.

Tibshirani R. 1996. Regression Shrinkage and Selection Via The Lasso. *Journal of the Royal Statistical Society B* 58(1), 267-288.

Wheeler D, Tiefelsdorf M. 2005. Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression. *J Geograph Syst* (2005) 7, 161-187.

Wheeler DC. 2007. Diagnostic Tools and a Remedial Method for Collinearity in Geographically Weighted Regression. *Environment and Planning A* 39 : 2464-2481.

Wheeler DC. 2009. Simultaneous Coefficient Penalization and Model Selection in Geographically Weighted Regression : The Geographically Weighted Lasso. *Journal of Environment and Planning A* 41 (3) : 722-742.