IOP Conference Series: Materials Science and Engineering

PAPER • **OPEN ACCESS**

Parameter Estimation of Locally Compensated Ridge-Geographically Weighted Regression Model

To cite this article: Alfi Fadliana *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **546** 052022

Parameter Estimation of Locally Compensated Ridge-Geographically Weighted Regression Model

Alfi Fadliana[1*)], Henny Pramoedyo[2], Rahma Fitriani[3]

[1]Statistics Master Study Program, Department of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University, Malang, East Java, Indonesia
[2,3]Department of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University, Malang, East Java, Indonesia

[*)] Corresponding author: alfifadliana@gmail.com

Abstract. Geographically weighted regression (GWR) is a spatial data analysis method where spatially varying relationships are explored between explanatory variables and a response variable. One unresolved problem with spatially varying coefficient regression models is local collinearity in weighted explanatory variables. The consequence of local collinearity is: estimation of GWR coefficients is possible but their standard errors tend to be large. As a result, the population values of the coefficients cannot be estimated with great precision or accuracy. In this paper, we propose a recently developed method to remediate the collinearity effects in GWR models using the Locally Compensated Ridge Geographically Weighted Regression (LCR-GWR). Our focus in this study was on reviewing the estimation parameters of LCR-GWR model. And also discussed an appropriate statistic for testing significance of parameters in the model. The result showed that Parameter estimation of LCR-GWR model using weighted least square method is $\hat{\boldsymbol{\beta}}(u_i, v_i, \lambda_i) = [\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i, v_i)]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{y}^*$, where the ridge parameter, λ, varies across space. The LCR-GWR is not necessarily calibrates the ridge regressions everywhere; only at locations where collinearity is likely to be an issue. And the parameter significance test using t-test, $t = \frac{\hat{\beta}_k(u_i, v_i, \lambda_i)}{\hat{\sigma}\sqrt{v_{kk}}}$.

Keywords: Geographically Weighted Regression, collinearity, Locally Compensated Ridge

1. Introduction

Regression analysis is a statistical technique for constructing mathematical model that can be used to investigate the relationships among variables. In fact, regression analysis may be the most widely used statistical method. Applications of regression are numerous and can be found in almost every field. However, if this technique is applied to spatial data, it may result in a significant problem since regression examines phenomena as if these parameters were constant over the space. Brunsdon *et al.* [1] developed Geographically Weighted Regression (GWR) as an approach to incorporate the spatial non-stability in the model. The goal of GWR is to allow spatial data analysts to visualize the spatial variation in relationships of explanatory variables to a response variable by way of the estimated regression coefficients from each calibration location in the study area [2].

Some example of the application of GWR methodology can be found in Nakaya [3] that uses the GWR approach for spatial interaction modeling with local distance decay and accessibility parameters. Huang and Leung [4] apply GWR to study regional industrialization in China. Longley and Tobón [5] perform a comparative study of several global and local spatial estimation procedures including GWR to investigate the heterogeneity in patterns of intra-urban hardship. Interestingly, these applications interpret the local parameter patterns without reporting the level correlation in the estimated regression coefficients, even though there appears to be coefficient correlation in some map patterns.

The problem of collinearity amongst the explanatory variables is worsened in the GWR model. According to Brunsdon, *et al.* [6], collinearity is an issue in GWR since: (i) its effects can be more pronounced with the smaller samples that are used to calibrate each local regression; and (ii) if the data is spatially heterogeneous in terms of its correlation structure, some localities may exhibit collinearity when others do not. In both cases, collinearity may cause problems in GWR when none are found globally. In addition, the consequence of local collinearity is: estimation of GWR coefficients is possible but their standard errors tend to be large. As a result, the population values of the coefficients cannot be estimated with great precision or accuracy. Local collinearity in weighted explanatory variables can lead to GWR coefficient estimates that are correlated locally and across space, have inflated variances, and are at times counterintuitive and contradictory sign to the global regression estimates, i.e., evidence of the reversal paradox [2, 7, 8].

Ridge regression is one of methods which attempt to circumvent collinearity in global linear regression models with constant coefficients. Ridge regression was designed specifically to reduce collinearity effects by penalizing the size of regression coefficients and decreasing the influence in the model of variables with relatively small variance in the design matrix. To address the issue of collinearity in the GWR framework, Wheeler [2] implemented a ridge-regression version of GWR, called Geographically Weighted Ridge Regression (GWRR), and found it was able to constrain the regression coefficients to counter local correlation present in an existing dataset [9].

Here is some previous research about GWRR: Wheeler [9] apply GWRR on Columbus Crime data and Sukmantoro [10] apply GWRR to predict the land value of Pondok Indah Residence in South Jakarta. The parameter estimation results of the GWRR model used by Wheeler [9] and Sukmantoro [10] used one ridge parameter for the entire observation area. Even though it is possible that not all observation areas have problems with local collinearity. Adding ridge parameters which actually does not have a problem with local collinearity between explanatory variables can actually reduce the effectiveness of the model.

In this paper, we will first introduce an extension to GWRR which allows the ridge parameter to vary across space. The term Locally Compensated Ridge-Geographically Weighted Regression (LCR-GWR) is proposed for this new method. LCR-GWR was originally proposed by Gollini, *et al.* [11]. The purpose of this paper is to (1) briefly review the parameter estimation of LCR-GWR model and (2) propose appropriate statistics for testing significance of the parameters in the model.

2. Theoretical Review

2.1 Geographically Weighted Regression (GWR)

Geographically weighted regression (GWR) is a spatial data analysis method where spatially varying relationships are explored between explanatory variables and a response variable. The specification of basic GWR model can be written as

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)\, x_{ik} + \varepsilon_i \tag{1}$$

where,

y_i : the dependent variable at location-i, for $i = 1,2,\dots,p$

(u_i, v_i) : the coordinates of i-th point in space

$\beta_0(u_i, v_i)$: the intercept parameter at location i

$\beta_k(u_i, v_i)$: the local regression coefficient for k-th explanatory variable at location i

x_{ik} : the value of the k-th explanatory variable at location i
ε_i : the error terms, which may follow an independent normal distribution with zero mean and homogeneous variance.

In matrix notation

$$\boldsymbol{y} = \boldsymbol{X\beta}(u_i, v_i) + \boldsymbol{\varepsilon} \tag{2}$$

The local parameter $\boldsymbol{\beta}(u_i, v_i)$ are estimated by weighted least square (WLS) estimator, given by

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = [\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{y} \tag{3}$$

where $\boldsymbol{X} = [\boldsymbol{X}^T(1); \boldsymbol{X}^T(2); \dots; \boldsymbol{X}^T(n)]^T$ is the design matrix of explanatory variables, which typically includes a column of 1s for the intercept, $\boldsymbol{W}(u_i, v_i) = diag\big(w_1(u_i, v_i), w_2(u_i, v_i), \dots, w_n(u_i, v_i)\big)$ is the diagonal weights matrix that varies by calibration location i, $\boldsymbol{y}$ is the $n \times 1$ vector of dependent variables, and $\widehat{\boldsymbol{\beta}}(u_i, v_i) = \left(\hat{\beta}_0(u_i, v_i), \hat{\beta}_1(u_i, v_i), \dots, \hat{\beta}_p(u_i, v_i)\right)^T$ is the vector of $(p+1)$ local regression coefficient at location i for p explanatory variables and an intercept term [9].

The weights matrix $\boldsymbol{W}(u_i, v_i)$, is calculated from a kernel function that places more emphasis on observations that are closer to the model calibration location i. There are numerous choices for the kernel function, including the Gaussian function, the bi-square nearest-neighbor function, and the exponential function [9]. The adaptive Gaussian kernel function is utilized in this paper. The weight from the adaptive Gaussian kernel function between any location j and the model calibration location i is calculated as

$$w_{ij} = exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{h_i}\right)^2\right] \tag{4}$$

where d_{ij} is the distance between the calibration location i and location j, and h_i is referred to as the bandwidth. There are a number of criteria that can be used for bandwidth selection. In this paper, we use a cross-validation (CV) approach suggested for local regression by Cleveland [12] and for kernel density estimation by Bowman [13]. CV is a technique in which the optimal bandwidth is that which minimises the following score

$$CV = \sum_{i=1}^{n}\big(y_i - \hat{y}_{\neq i}(h)\big)^2 \tag{5}$$

where $\hat{y}_{\neq i}(h)$ is the fitted value of y_i with the observations for point i omitted from the calibration process [14].

2.2 Ridge Regression

The ridge regression method is known as a fairly efficient corrective action to overcome collinearity in linear regression models [15]. According to Kutner, *et al.* [16], estimation of ridge regression parameters was done by standardizing predictor variables and response variables with models

$$y_i^* = \beta_0 + \sum_{k=1}^{p}\beta_k\, x_{ik}^* + \varepsilon_i \tag{6}$$

where,

$$y_i^* = \frac{1}{\sqrt{n-1}}\left(\frac{y_i - \bar{y}}{s_y}\right), \qquad x_{ik}^* = \frac{1}{\sqrt{n-1}}\left(\frac{x_{ik} - \bar{x}_k}{s_x}\right)$$
$$s_y = \sqrt{\frac{\sum_i (y_i - \bar{y})^2}{n-1}}, \qquad s_x = \sqrt{\frac{\sum_i (x_{ik} - \bar{x}_k)^2}{n-1}} \tag{7}$$

The parameter estimation of the ridge regression model is obtained in the same way as the ordinary least square (OLS) method, namely by minimizing error sum square. The ridge regression adds an obstacle to the least square so that the coefficient shrinks near zero [17]. In the ridge regression method, the resulting parameter estimator is a biased estimator but tends to be more stable and more potential to produce better accuracy compared to the predicted results using OLS.

The estimator of ridge regression parameters is

$$\widehat{\boldsymbol{\beta}}_{ridge} = \left(\boldsymbol{X}^{*T}\boldsymbol{X}^* + \lambda \boldsymbol{I}\right)^{-1}\boldsymbol{X}^{*T}\boldsymbol{y}^* \tag{8}$$

where the constant λ is the magnitude of the bias coefficient of the parameter estimator located at the interval $0 < \lambda < 1$.

3. Result and Discussion

3.1 Estimation of Parameters in Locally Compensated Ridge-Geographically Weighted Regression Model

The Locally Compensated Ridge-Geographically Weighted Regression (LCR-GWR) model is a development of the GWRR model using one bias coefficient for a given region. That is, if there are N observation regions, there are n different ridge bias coefficients. This method produces a ridge bias coefficient locally. The parameters of the ridge are left to vary in each region adjusting to the effect of collinearity between explanatory variables in each region so that the expected parameter coefficients on the model are expected to be more accurate.

The LCR-GWR model can be expressed by

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, \lambda_i)x_{ik} + \varepsilon_i \tag{9}$$

where, $\beta_k(u_i, v_i, \lambda_i)$ is the local regression coefficient for k-th explanatory variable at location i and a specified value of the ridge bias coefficients at location i, λ_i.

The solution of parameter estimation for the LCR-GWR model is done using the WLS method on the GWR model by first centering on the y variable and centering-scaling on X variables. So that the equation of the GWR model can be written

$$y_i^* = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik}^* + \varepsilon_i \tag{10}$$

or in the matrix form

$$\boldsymbol{y}^* = \boldsymbol{X}^*\boldsymbol{\beta}(u_i, v_i) + \boldsymbol{\varepsilon}^* \tag{11}$$

By adding weighting elements $\boldsymbol{W}(u_i, v_i)$ to Equation (11), the number of error sum squares (ESS) is obtained as follows

$$\boldsymbol{\varepsilon}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{\varepsilon}^* = \boldsymbol{y}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^* - 2\boldsymbol{\beta}^T(u_i,v_i)\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^* + \boldsymbol{\beta}^T(u_i,v_i)\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{X}^*\boldsymbol{\beta}(u_i,v_i) \quad (12)$$

The parameter estimation solution for the LCR-GWR model is obtained by adding the coefficient $\lambda\boldsymbol{I}(u_i,v_i)$ which is a Locally Compensated (LC) value of λ in the region (u_i,v_i), so that Equation (12) becomes

$$\boldsymbol{\varepsilon}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{\varepsilon}^* = \boldsymbol{y}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^* - 2\boldsymbol{\beta}^T(u_i,v_i)\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^* + \boldsymbol{\beta}^T(u_i,v_i)\left(\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i,v_i)\right)\boldsymbol{\beta}(u_i,v_i) \quad (13)$$

Next, Equation (13) is derived from $\boldsymbol{\beta}^T(u_i,v_i)$ and the result is equal to zero, obtained

$$\left.\frac{\partial\boldsymbol{\varepsilon}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{\varepsilon}^*}{\partial\boldsymbol{\beta}^T(u_i,v_i)}\right|_{\widehat{\boldsymbol{\beta}}(u_i,v_i)} = 0$$

$$\Leftrightarrow -2\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^* + 2\left(\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i,v_i)\right)\boldsymbol{\beta}(u_i,v_i) = 0$$

$$\Leftrightarrow \left(\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i,v_i)\right)\boldsymbol{\beta}(u_i,v_i) = \boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^*$$

$$\Leftrightarrow \widehat{\boldsymbol{\beta}}(u_i,v_i) = \left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i,v_i)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^*$$

So that the estimator of the LCR-GWR model, $\widehat{\boldsymbol{\beta}}(u_i,v_i)$, is obtained at the specified value λ for each location as follows:

$$\widehat{\boldsymbol{\beta}}(u_i,v_i,\lambda_i) = \left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i,v_i)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_i,v_i)\boldsymbol{y}^* \quad (14)$$

where, $\widehat{\boldsymbol{\beta}}(u_i,v_i,\lambda_i) = \begin{bmatrix}\hat{\beta}_0(u_1,v_1,\lambda_0)\\ \hat{\beta}_1(u_1,v_1,\lambda_1)\\ \hat{\beta}_2(u_2,v_2,\lambda_2)\\ \vdots\\ \hat{\beta}_p(u_i,v_i,\lambda_i)\end{bmatrix}$.

The value of the ridge regression parameter is obtained by connecting the eigenvalue and conditional number (c) of matrix multiplication $(\boldsymbol{X}^T\boldsymbol{X})$. If the eigenvalues of matrix $(\boldsymbol{X}^T\boldsymbol{X})$ are $\epsilon_1,\epsilon_2,\dots,\epsilon_p$ then the eigenvalues of matrix $(\boldsymbol{X}^T\boldsymbol{X}+\lambda\boldsymbol{I})$ are obtained $\epsilon_1+\lambda,\epsilon_2+\lambda,\dots,\epsilon_p+\lambda$. The conditional number (c) of a square matrix is defined as ϵ_1/ϵ_p, where ϵ_1 is the largest eigenvalue and ϵ_p is the smallest eigenvalue, so the conditional number (c) for the ridge-adjusted matrix is defined as $\epsilon_1+\lambda/\epsilon_p+\lambda$. By re-arranging the terms, the ridge bias coefficient obtained from a particular conditional number (c) is $\lambda = \left((\epsilon_1-\epsilon_p)/(c-1)\right)-\epsilon_p$.

The LCR-GWR model fits local ridge regressions with their own ridge parameters (i.e., the ridge parameter varies across space), and only fits such ridge regressions at locations where the local condition number is above a user-specified threshold. Thus a biased local estimation is not necessarily used everywhere, only at locations where collinearity is likely to be an issue. At all other locations, the usual un-biased estimator is used, hoping to produce a more accurate model with local collinearity problems and spatial heterogeneity [11].

3.2 Parameters Significance Test of Locally Compensated Ridge-Geographically Weighted Regression Model

Testing the parameters significance of the GWRR model using the following hypothesis

$H_0: \beta_k(u_i, v_i, \lambda_i) = 0, \;\; i = 0, 1, 2, \dots, n$

versus

$H_1: \exists \beta_k(u_i, v_i, \lambda_i) \neq 0.$

The estimator of the parameter $\boldsymbol{\beta}(u_i, v_i, \lambda_i)$ is exactly distributed as a normal distribution with $\boldsymbol{\beta}(u_i, v_i, \lambda_i)$ mean and $\boldsymbol{V}\boldsymbol{V}^T\sigma^2$ variance-covariance matrices, where $\boldsymbol{V} = \left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i, v_i)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)$. Mathematically the test statistics used can be written

$$t = \frac{\hat{\beta}_k(u_i, v_i, \lambda_i)}{\hat{\sigma}\sqrt{v_{kk}}} \tag{15}$$

where v_{kk} is the k-th diagonal element of the matrix $\boldsymbol{V}\boldsymbol{V}^T$and $\hat{\sigma}$ is the root of $\hat{\sigma}^2 = \frac{ESS_{LCR-GWR}}{p_1}$, with

$$ESS_{LCR-GWR} = \hat{\boldsymbol{\varepsilon}}^T\boldsymbol{W}(u_i, v_i)\hat{\boldsymbol{\varepsilon}} = (\boldsymbol{y}^* - \hat{\boldsymbol{y}}^*)^T(\boldsymbol{y}^* - \hat{\boldsymbol{y}}^*) = (\boldsymbol{y}^*)^T(\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})\boldsymbol{y}^* \tag{16}$$

and

$$p_1 = tr[(\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})] \tag{17}$$

then

$$\boldsymbol{S} = \begin{bmatrix} x_1^{*T}\left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_1, v_1)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_1, v_1)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_1, v_1) \\ x_2^{*T}\left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_2, v_2)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_2, v_2)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_2, v_2) \\ \vdots \\ x_n^{*T}\left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_n, v_n)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_n, v_n)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_n, v_n) \end{bmatrix} \tag{18}$$

The H_0 rejection criterion is if $|t| > t_{\alpha/2\left(\frac{p_1^2}{q_2}\right)}$. With, $q_2 = tr(\boldsymbol{B}^2)$, where $\boldsymbol{B} = (\boldsymbol{I} - \boldsymbol{K}) - (I - S)^{\boldsymbol{T}}(I - S)$ and $\boldsymbol{K} = \boldsymbol{X}^*\left[X^{*\boldsymbol{T}}W(u_i, v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i, v_i)\right]^{-1}X^{*\boldsymbol{T}}W(u_i, v_i)$.

4. Concluding Remarks

Based on results and discussion, the following conclusions is obtained. Parameter estimation of LCR-GWR model using weighted least square method is $\widehat{\boldsymbol{\beta}}(u_i, v_i, \lambda_i) = \left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{X}^* + \lambda\boldsymbol{I}(u_i, v_i)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{y}^*$, where the ridge parameter, λ, varies across space. The LCR-GWR is not necessarily calibrates the ridge regressions everywhere; only at locations where collinearity is likely to be an issue. And the parameter significance test using t-test, $t = \frac{\widehat{\beta}_k(u_i, v_i, \lambda_i)}{\hat{\sigma}\sqrt{v_{kk}}}$.

References

[1] Brunsdon, C., Fotheringham, A.S., Charlton, M. 1996. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, **28(4)**: 281-298.

[2] Wheeler, D.C. 2007. Diagnostic Tools and A Remedial Method for Collinearity in Geographically Weighted Regression. *Environment and Planning A*, **39**: 2464-2481.

[3] Nakaya, T. 2001. Local Spatial Interaction Modelling Based on the Geographically Weighted Regression Approach. *GeoJournal*, **53**: 347–358.

[4] Huang, Y. and Leung, Y. 2002. Analysing Regional Industrialisation in Jiangsu Province using Geographically Weighted Regression. *J Geogr Syst*, **4**: 233–249.

[5] Longley, P.A. and Tobón, C. 2004. Spatial Dependence and Heterogeneity in Patterns of Hardship: An Intra-Urban Analysis. *Ann Assoc Am Geogr*, **94**:503–519.

[6] Brunsdon, C., Charlton, M., and Harris, P. 2012. Living with Collinearity in Local Regression Models. Proceeding in International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 10th – 13th July, 2012, Florianópolis, SC, Brazil.

[7] Tu, Y-K., Gunnel, D., and Gilthorpe MS. 2008. Simpson's Paradox, Lord's Paradox, and Suppression Effects are the Same Phenomenon – the Reversal Paradox. *Emerg Themes Epidemiol* **5**: 2.

[8] Czarnota, J., Wheeler, D.C., and Gennings, C. 2015. Evaluating Geographically Weighted Regression Models for Environmental Chemical Risk Analysis. *Cancer Informatics*, **14(S2)**: 117-127.

[9] Wheeler, D.C. 2009. Simultaneous Coefficient Penalization and Model Selection in Geographically Weighted Regression: The Geographically Weighted LASSO. *Environment and Planning A*, **41**: 722-742.

[10] Sukmantoro, D. 2014. *Geographichally Weighted Ridge Regression dalam Pemodelan Nilai Tanah* (*Studi Kasus Perumahan Pondok Indah Jakarta Selatan Kondisi* 1 *Januari* 2011). Master Thesis (unpublished). Surabaya: Institut Teknologi Sepuluh Nopember.

[11] Gollini, I., Lu, B., Charlton, M., Brunsdon, C., and Harris, P. 2015. GW Model: An R Package for Exploring Spatial Heterogeneity using Geographically Weighted Models. *Journal of Statistical Software*, **63(17)**: 1-50.

[12] Cleveland W. S. 1979. Robust Locally-Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, **74**: 829-836.

[13] Bowman, A.W. 1984. An Alternative Method of Cross-Validation for the Smoothing of Density Estimates. *Biometrika* **71**: 353-360.

[14] Fotheringham, S.A., Brunsdon, C., and Charlton, M. 2002. *Geographically Weighted Regression*: *The Analysis of Spatially Varying Relationships*. New York: John Willey & sons.

[15] Kibria, B.M.G., Mansson, K., and Shukur, G. 2012. Performance of Some Logistic Ridge Regression Estimators. *Jurnal Comput Econ*, **40**: 401-414.

[16] Kutner, M.H., Nachtsheim, C.J., and Neter, J. 2004. *Applied Linear Regression Models*, 4th *Edition*. New York: McGraw-Hill Companies, Inc.

[17] Hastie, T., Tibshirani, R., and Friedman, J. 2016. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd *Ed*. New York: Springer.