Geographically and Temporally Weighted Regression Model with Gaussian Kernel Weighted Function and Bisquare Kernel Weighted Function

Harianto*, W H Nugroho and E Sumarminingsih
Department of Statistic, Malang Brawijaya University, Indonesia

*Corresponding author: harianto@student.ub.ac.id

Abstract. The Geographically and Temporally Weighted Regression (GTWR) model is a development of the Geographically Weighted Regression (GWR) model. Estimation of parameters in the GTWR model uses Weighted Least Square (WLS). The purpose of this study is to build two GTWR models using the Gaussian kernel weight function and the Bisquare kernel weight function on the percentage of poor people in East Java Province in 2010-2015. Each weighted function produces 228 different models at each location and time. The GTWR method with the Gaussian Kernel weight function produces parameter estimates at the observation location in the range between 0.7796 to 48.1435. Meanwhile, the parameter estimation of the GTWR model with the Kernel Bisquare weight function ranges from 0.7828 to 48.1444. The two weight functions can model the percentage of poor people in East Java Province in 2010-2015, with $R^2$ values of 98.994% and 98.987%, respectively.

1. Introduction
The Geographically and Temporally Weighted Regression (GTWR) method is a development of the Geographically Weighted Regression (GWR) method with the aim of being able to simultaneously handle spatial and temporal non-stationary data [1]. Furthermore, based on ref [2], the GTWR model has the advantage that the resulting model is local at each location and time, this makes the resulting model more representative in terms of location and time. Several previous studies on GTWR have been conducted in ref [3], for modeling economic growth using R and in ref [4], for modeling the proportion of poor people in districts and cities in Central Java Province. Also, in ref [5] for determining the optimum bandwidth result of the GTWR model using the cross validation method. Dwi et al. (2016) modeling of air pollutant elements No2 using GTWR [6]. One of the important elements in the GTWR method is to construct a weighted matrix with data containing spatial and temporal information that can identify spatial and temporal diversity. This weighted idea is based on Tobler's I law which says "All things are related to each other, but something closer affects something far away" [7]. A weighted function that can be used either by using kernel functions. This research uses the Gaussian kernel function and the Bisquare kernel function to build the model.

Yusuf (2013) implemented Gaussian kernel functions and Bisquare kernel functions to build a GWR model [8]. The weighting function is used because both involve the element of distance between the observed locations whose value is continuous in constructing the weighting matrix, so that each location will receive a weight according to the distance between the location and the observed location. The purpose of this research is to apply the Gaussian kernel weight function and
the Bisquare kernel function in building the GTWR model with a case study of the percentage of poor people in East Java Province in 2010-2015.

2. Literature Review

2.1. Geographically Weighted Regression (GWR)

One of the methods of regression that incorporates spatial elements in building a model is the Geographically Weighted Regression (GWR) method. GWR was developed from a global regression model based on non-parametric regression [9]. GWR is a method aimed at building a regression model that is local. The estimation of the regression parameters in the GWR method is carried out at each location by involving the surrounding information that has been weighted by geographical distance [10]. The GWR model can generally be written as follows:

\[ Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i ; k = 1, 2, ... p ; i = 1, 2, ..., n \] (1)

The parameter estimation on GWR model using Weighted Least Square (WLS) Method by providing different weighted for each location where data is observed. Here are the following formulation of WLS estimator:

\[ \hat{\beta}(u_i, v_i) = [X^TW(u_i, v_i)X]^{-1}X^TW(u_i, v_i)y \] (2)

with

\[ W(u_i, v_i) = \begin{bmatrix} w_{i1} & 0 & ... & 0 \\ 0 & w_{i2} & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & w_{in} \end{bmatrix} \]

In equation (2), estimator \( \hat{\beta}(u_i, v_i) \) is an unbiased and consistent estimator of \( \beta(u_i, v_i) \) [4].

2.2. Spatial Weight Function

Regression involving spatial effects, one that determines the results of the analysis is the choice of weighting function [11]. In 2013, Yusuf [8] applied the two weight functions to the GWR method. Meanwhile, in this study, the Gaussian Kernel weighted function and the Bisquare Kernel weight function will be used in constructing the weight matrix in the GTWR model.

2.2.1. Gaussian Kernel function

\[ w_{ij} = \exp\left(-\frac{1}{2} \left( \frac{d_{ij}}{h} \right)^2 \right) \] (3)

where \( d_{ij} \) is the euclid distance from location-\( i \) to location-\( j \), and \( b \) are the optimum window width.

In the gaussian kernel weight function, the weight \( (w_{ij}(i)) \) will approach one along the closer the distance between the location-\( i \) with location-\( j \) and weighted value \( (w_{ij}(i)) \) decreases as the distance grows between location-\( i \) and location-\( j \).

2.2.2. Bisquare Kernel function

\[ w_{ij}(u_i, v_i) = \begin{cases} 1 - \left( \frac{d_{ij}}{h} \right)^2 & ; d_{ij} < h \\ 0 & ; another \end{cases} \] (4)

where \( d_{ij} \) is the euclid distance from location-\( i \) to location-\( j \), and \( b \) is the optimum window width, namely the optimum distance to a location still give influence to the location is being observed. In the bisquare kernel weighting function, if the distance between location-\( i \) and location-\( j \) greater than or equal to the window width, then the location will be given a zero weight. While the distance between locations is less than the width of the window will be weighted close to one as the distance between location-\( i \) and location-\( j \) gets closer.
While, to determine the optimum bandwidth using the Cross Validation (CV) method, mathematically it can be written in equation 5 [11].

\[ CV(h) = \sum_{i=1}^{n}(y_i - \hat{\gamma}_x(h))^2 \]  

(5)

2.3. Geographically and temporally Weighted Regression (GTWR)

The GWR model only considers spatial effects, while GTWR combines spatial and temporal data information in a weighted matrix to identify spatial and temporal heterogeneity [4]. The GTWR model for the independent variable \( p \) with the dependent variable at the location \( \{u_i, v_i, t_i\} \) for each observation is written in equation (6):

\[ Y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i) x_{ik} + \epsilon_i \]  

(6)

Estimation of the regression coefficient parameter \( \hat{\beta}(u_i, v_i, t_i) \) at point \( i \) can be obtained using the Weighted Least Square (WLS) method as follows:

\[ \hat{\beta}(u_i, v_i, t_i) = (X^TW(u_i, v_i, t_i)X)^{-1}X^TW(u_i, v_i, t_i)Y \]  

(7)

Where \( W(u_i, v_i, t_i) = diag \{w_{i1}, w_{i2}, ..., w_{in}\} \) is a matrix on the observation location \( (u_i, v_i) \) and time \( t_i \). Diagonal element \( w_{ij} (1 < j < n) \) is a spatial-temporal distance function at the observation point \( (u_i, v_i, t_i) \). The built in spatial-temporal distance is a combination of spatial distance function and temporal distance function [12]. In 2010 Huang et al [2] given Illustration of spatial-temporal distance uses an ellipsoidal coordinate system to measure the closeness between the observation point and the observation point that surrounds it. The following illustration of ellipsoidal coordinates is shown in Figure 1.

![Figure 1. illustration of spatial-temporal distance at ellipsoidal coordinates](image)

Spatial-temporal distance function \((d_{ij}^{ST})\) is a combination of spatial distance function \((d_{ij}^S)\) and temporal distance function \((d_{ij}^T)\), which are written as follows [2]:

\[(d_{ij}^{ST})^2 = \lambda (d_{ij}^S)^2 + \mu (d_{ij}^T)^2\]  

(8)

where

\[
\begin{align*}
(d_{ij}^S)^2 &= (u_i - u_j)^2 + (v_i - v_j)^2 \\
(d_{ij}^T)^2 &= (t_i - t_j)^2 \\
(d_{ij}^{ST})^2 &= \lambda [(u_i - u_j)^2 + (v_i - v_j)^2] + \mu [(t_i - t_j)^2]
\end{align*}
\]  

(9)

where \( \lambda \) and \( \mu \) are used as a balancing parameter between the influence of location and time on the measurement of spatial-temporal distances. Based on the equation above obtained:

\[ w_{ij} = \exp \left\{ \frac{-\left(\frac{1}{h_{ST}^2}[(u_i - u_j)^2 + (v_i - v_j)^2] + \frac{1}{h_{T}^2}[(t_i - t_j)^2]\right)}{h_{ST}^2 + h_{T}^2} \right\} \]  

(10)

For example \( h_{ST}^2 = \frac{h_{ST}^2}{\lambda} \) and \( h_{T}^2 = \frac{h_{T}^2}{\mu} \), it is obtained:

\[ w_{ij} = \exp \left\{ \frac{-\left(\frac{(d_{ij}^S)^2}{h_{S}^2} + \frac{(d_{ij}^T)^2}{h_{T}^2}\right)}{h_{S}^2 + h_{T}^2} \right\} \]
\[ e^{-\left(\frac{(d_{ij}^s)^2}{h_s^2}\right)} \times e^{-\left(\frac{(d_{ij}^t)^2}{h_t^2}\right)} \]

\[ w_{ij}^s \times w_{ij}^t \]

Where \( w_{ij}^s = e^{-\left(\frac{(d_{ij}^s)^2}{h_s^2}\right)} \) and \( w_{ij}^t = e^{-\left(\frac{(d_{ij}^t)^2}{h_t^2}\right)} \)

As for the parameter values \( \tau \), for example \( \tau \) is the ratio parameter of \( \frac{\mu}{\lambda} \) with \( \lambda \neq 0 \), the equation is obtained [12]:

\[ \frac{(d_{ij}^T)^2}{\lambda} = \left[ (v_i - v_j)^2 + (v_i - v_j)^2 \right] + \tau\left[ (t_i - t_j)^2 \right] \]  \( (11) \)

For example \( \lambda = 1 \), which aims to reduce unknown parameters. There is one unknown parameter, which is \( \tau \). The parameter \( \tau \) is useful to balance the effect of temporal distance on spatial distance. This parameter is obtained from the minimum Cross Validation (CV) method through the initial value \( \tau \) written in equation (12):

\[ CV(\tau) = \sum_{i=1}^{n} (y_i - \hat{y}_i(\tau))^2 \]  \( (12) \)

In this study, the model accuracy test (Goodness of fit) uses the coefficient of determination \( (R^2) \). The coefficient of determination can explain the large variance of response variables that can be explained by the predictor variable. The value of \( R^2 \) is obtained through a mathematical equation as follows [11]:

\[ R^2(u,v,t) = \frac{\sum_{i=1}^{P} \sum_{j=1}^{P} w_{ij}(y_j - \hat{y}_j)^2}{\sum_{i=1}^{P} \sum_{j=1}^{P} w_{ij}(y_j - \bar{y})^2} \]  \( (13) \)

3. Result and Discussion

This study uses secondary data sourced from the Central Statistics Agency (BPS) of East Java Province during 2010-2015. Spatial units used in this study were 38 regencies / cities in East Java Province consisting of 29 regencies and 9 cities. The response variable in this study was the Percentage of poor population \( (Y) \). While the Predictor variables used were the School Participation Rate aged 16-18 years \( (x_1) \), Morbidity \( (x_2) \) and Unemployment Rate \( (x_3) \). The analysis steps are as follows: (1) Describe the research variables. (2) Conduct a test of spatial and temporal heterogeneity. (3) Calculating the Euclidean Spatial-temporal distance at coordinates \( (u_i, v_i, t_i) \) Include: determine the parameters of the spatial-temporal ratio \( (\tau) \), determine spatial parameters \( (\mu) \) and temporal parameters \( (\lambda) \) and Calculate spatial-temporal bandwidth \( (h_{ST}) \). (4) Determine the weighted value using the Gaussian kernel function and the Bisquare kernel function. (5) Estimated parameters \( \hat{\beta}(u_i, v_i, t_i) \). (6) Calculating the value of \( R^2 \) in each model with equation (13). The statistical description of the research data is presented in Table 1.

| Table 1. Data Description |
|---------------------------|
| **Variable** | **Min** | **Max** | **Mean** |
| Percentage of poor people \( (Y) \) | 4.45 | 32.47 | 23.12 |
| School participation rate Age 16-18 years \( (x_1) \) | 38.61 | 93.75 | 66.38 |
| Morbidity \( (x_2) \) | 7.02 | 27.72 | 15.34 |
| unemployment rate \( (x_3) \) | 0.87 | 10.62 | 4.52 |

Table 1 shows that in the period of 6 years the average percentage of poor people in East Java province amounted to 23.12% with the highest percentage of poverty at 32% and the lowest 4.45%. As for predictor variables, the 16-18 years old school participation rate has the highest number of 93.75%, the morbidity variable has the lowest value of 7.02%. While the unemployment rate variable for a period of 6 years has an average of 4.52% with the lowest number of 0.87%. The relationship...
variable responses with each of the predictor variables in a linear can be described by using a Pearson correlation. The relationship between response variable and predictor variables can be seen in Table 2.

Table 2. Correlation coefficient of the response variable and the predictor variables

| Variable | Pearson Correlation | P-value     |
|----------|---------------------|-------------|
| Y with X1 | -0.5849776          | 2.2 x 10^{-16} ** |
| Y with X2 | 0.121656            | 0.0667 *    |
| Y with X3 | -0.472064           | 4.692 x 10^{-14} ** |

( ** ) significant at a level of 5%, ( * ) significant at a level of 10%

Table 2 shows the relationship between variables Y with variables X1 and X3 which are negative or have an inverse relationship with correlation coefficient values of 58% and 47% at the 95% confidence level. As for the variable Y with the variable X2 has a positive sign or has a direct relationship with a coefficient of 12% at a 90% confidence level. Furthermore the relationship between predictor variables needs to be seen to meet the assumption of non-multicollinearity. linear correlation between predictor variables can be tested with multicollinearity assumption test.

Table 3. VIF value on each predictor variable

| Years          | X1            | X2            | X3            |
|----------------|---------------|---------------|---------------|
| Full (2010-2015) | 1.661568      | 1.031367      | 1.539207      |
| 2010           | 2.829173      | 1.113492      | 3.278216      |
| 2011           | 2.268411      | 1.163203      | 2.501853      |
| 2012           | 1.667146      | 1.075252      | 1.526519      |
| 2013           | 1.565432      | 1.025933      | 1.354471      |
| 2014           | 1.436456      | 1.107825      | 1.545898      |
| 2015           | 1.730509      | 1.076934      | 1.514770      |

Table 3 shows the value of the Variance Inflation Factor (VIF) of the multicollinearity assumption test obtained from each predictor variable entirely under 10, meaning that the risk of multicollinearity is relatively low. In the spatial heterogeneity test showed the p-value during 2010-2015 was 3.027 x 10^{-11}. These results can be concluded that there is spatial heterogeneity. The following Breusch-pagan test results can be seen in Table 4.

Table 4. Breusch-Pagan Test

| Years          | Breusch-pagan value | p-value     |
|----------------|---------------------|-------------|
| Full (2010-2015)| 55.147              | 3.027 x 10^{-11} ** |
| 2010           | 9.2353              | 0.05548 *   |
| 2011           | 9.304               | 0.05393 *   |
| 2012           | 14.022              | 0.007225 ** |
| 2013           | 9.5494              | 0.04874 **  |
| 2014           | 10.184              | 0.03744 **  |
| 2015           | 10.949              | 0.02715 **  |

( ** ) significant at a level of 5%, ( * ) significant at a level of 10%

While from the temporal dimension, the percentage of poor people in the province of East Java tends to decrease every year. The differences in each year indicate temporal heterogeneity. The following temporal heterogeneity with boxplot can be seen in Figure 2.

Boxplot describe temporal diversity of data. Figure 2 above shows that each year the size of the boxplot tends to narrow. Outliers in boxplot also appear to be decreasing every year. This can be explained that there is one regenci / city that has the highest percentage of poverty compared to other regencies / cities in a period of 6 years.
**Table 5. Description of Parameter Estimation Models**

| Weight       | Parameter | Min     | Max     | Mean     | Std. Deviation |
|--------------|-----------|---------|---------|----------|----------------|
| Kernel       | $\hat{\beta}_0$ | 0.7796  | 48.1435 | 19.2562  | 11.8940        |
| Gaussian     | $\hat{\beta}_1$ | -0.6466 | 0.06396 | -0.1241  | 0.1549         |
|              | $\hat{\beta}_2$ | -0.2708 | 0.8837  | 0.0604   | 0.2294         |
|              | $\hat{\beta}_3$ | -1.5191 | 1.9152  | 0.0473   | 0.7399         |
| Kernel       | $\hat{\beta}_0$ | 0.7828  | 48.1444 | 19.256   | 11.8947        |
| Bisquare     | $\hat{\beta}_1$ | -0.647  | 0.0639  | -0.1242  | 0.1549         |
|              | $\hat{\beta}_2$ | -0.2708 | 0.8836  | 0.0604   | 0.2295         |
|              | $\hat{\beta}_3$ | -1.5190 | 1.915   | 0.0474   | 0.7401         |
4. Conclusion

GTWR modeling in poverty cases in 38 regencies / cities in East Java Province during 2010-2015 using the Gaussian Kernel weight function and the Bisquare Kernel Weight function each produce 228 models in each location and time. GTWR modeling uses the Gaussian Kernel weight function and the Bisquare Kernel weight function in poverty cases in East Java Province, resulting in almost the same estimated parameters. The two weight functions can model the percentage of poor people in East Java Province in 2010-2015, with $R^2$ values of 98.994% and 98.987%, respectively. This study is limited to the Gaussian kernel function and Bisquare kernel function in constructing a weight matrix.

References

[1] Wang P 2006 Exploring spatial effects on housing price: the case study of the city of Calgary, Master dissertation, University of Calgary, Canada
[2] Huang B, Wu B and Barry M 2010 Geographically and Temporally Weighted Regression for Modeling Spatio-Temporal Variation in House Prices International Journal of Geographical Information Science 24 (3) 383-401
[3] Sholihin M, Soleh AM and Djuaridah A 2017 Geographically and Temporally Weighted Regression (GTWR) for Modeling Economic Growth using R International Journal of Computer Science and Network 800-805
[4] Widiyanti KY, Yasin H and Sugito 2014 Pemodelan Proporsi Penduduk Miskin Kabupaten dan Kota di Provinsi Jawa Tengah Menggunakan Geographically and Temporally Weighted Regression Jurnal Gaussian 3 (4) 691-700
[5] Farihah I 2017 Penentuan Hasil Bandwidth Optimum Model Geographically and Temporally Weighted dengan metode Cross Validation Master dissertation UIN Malang, Indonesia.
[6] Dewi I, Yasin H and Hoyyi A 2016 Pemodelan Unsur Pencemar Udara No2 Menggunakan Geographically and Temporally Weighted Regression (GTWR) Prosiding Seminar Nasional Variansi Universitas negeri Makasar Forum Pendidikan Tinggi Statistika Indonesia 1-10
[7] Miller H J 2004 Tobler’s First Law and Spatial Analysis Annals of the Association of America Geographers 94 (2) 284-289
[8] Yusuf L M 2013 Regresi Terboboti Geografis dengan Fungsi Pembobot Kernel Gaussian dan Kernel Bisquare pada angka Harapan Hidup Studi kasus: Angka Harapan Hidup Kabupaten/Kota Provinsi Jawa Timur Master dissertation Institut Pertanian Bogor
[9] Mei C I 2005 Geographically Weighted Regression Technique for Spatial Data Analysis Master dissertation Xi’an Jiaotong University
[10] Brunsdon C, Fotheringham AS and Charlton M 1999 Some Notes on Parametric Significance Test for Geographically Weighted Regression Journal of Regional Science 39 (3) 497-524
[11] Fotheringham A S, Brunsdon C and Charlton M 2002 Geographically Weighted Regression, The Analysis of Spatially Varying Relationships England: John Wiley & Sons
[12] Liu J, Zhao Y, Yang Y, Xu S, Zhang F, Zhang X, Shi L and Qiu A 2017 A Mixed Geographically and Temporally Weighted Regression: Exploring Spatial-Temporal Variations from Global and Local Perspectives Entropy 53 (19) 1-20