IJCSN - International Journal of Computer Science and Network, Volume 6, Issue 6, December 2017
ISSN (Online) : 2277-5420
www.IJCSN.org
Impact Factor: 1.5

Geographically and Temporally Weighted Regression (GTWR) for Modeling Economic Growth using R

[1] Miftahus Sholihin; [2]Agus Mohamad Soleh; [3]Anik Djuraidah

[1] Department of Statistics, Bogor Agricultural University, Indonesia
Bogor, Indonesia

[2] Department of Statistics, Bogor Agricultural University, Indonesia
Bogor, Indonesia

[3] Department of Statistics, Bogor Agricultural University, Indonesia
Bogor, Indonesia

Abstract - Economic growth is a main condition for the sustainability of regional economic development. Spatially, the highest economic growth in Indonesia is dominated by provinces in Java. However, the economic growth rate of Central Java Province is the lowest economic growth compared to other provinces. The Geographically and Temporally Weighted Regression (GTWR) method performed to model the economic growth of the Central Java Provincial districts by accommodating the influence of spatial-temporal heterogeneity. This modeling involves four explanatory variables e.g, number of labor force, local revenue, district minimum wage, and human development index with response variable gross regional domestic product. The results of the analysis showed that GTWR method has better coefficient determination (99.8%) with root mean squared error and Akaike's Information Criterion values of 0.84 and 1051.98. In general, HDI gives more influence to economic growth at each regency / city in Central Java during 2011-2015.

Keywords – *Coefficient determination, Economic growth, GTWR, Spatial-temporal heterogeneity*

1. Introduction

Slow and uneven growth of the world economy causes the dynamics of global economy. Global economic growth sustained by the growth of the developed countries e.g. US especially when European and Japanese economic growth weakened [2]. High economic growth is a main condition for the sustainability of regional economic development. However, on the data Economic growth in a certain period there is dependence of space and time [5].

National income of country shows economic rate activity overall. The concept of national income is the most used indicator of economic growth. However, is not the only one indicator of economic growth [12]. Indonesia's economy in the first quarter of 2015 grew by 4.71% compared to the previous year. Spatially, Economic structure in Indonesia dominated by provincial groups such as Java and Sumatra. Java contributes the most to economic growth through Gross Domestic Product of 58.30% [3].

In the Gross Regional Domestic Product per capita by province in Java Island, Central Java Province has the lower Gross Regional Domestic Product compared to other provinces [4]. The low Gross Regional Domestic

Product per capita of Central Java Province compared to other provinces indicates that the level of welfare of the people of Central Java is lower with other provinces in Java [11]. Regardless of its point of view, Central Java is located between West Java and East Java so that there is high population mobility. If properly utilized can certainly help the economic growth of Central Java. This problem indicates that prosperity level of Central Java needs to be improved so as not to lose with other provinces in Java.

Research on Modeling of Particle Concentration Dust (PM_{10}) on Air Pollution in Surabaya City showed that GTWR method has more accurate results in predicting concentration of dust particles in Surabaya [1]. Another research on the GTWR approach with mileage to estimate home prices showed that the GTWR model that takes into account mileage is a more efficient approach to temporary non-stationary (spatial and temporal) testing [9].

In this research, GTWR method used to model economic growth of regency / city of Central Java Province. Analysis used R programs that have not previously been developed.

2. Literature Review

2.1 Geographically Weighted Regression

The Geographically Weighted Regression (GWR) model is an expansion of the global regression model based on nonparametric regression [10]. This model performed to analyze the effect of explanatory variables on the response to spatial heterogeneity. Spatial heterogeneity occurs when an equal single explanatory variable gives an unequal response at different locations or regions [13]. The GWR model can be expressed as follows [8]:

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

where y_i is the observed value of the response variable for the *i*-location, (u_i, v_i) is the geographical location coordinate of the *i*-th observation location, $\beta_0(u_i, v_i)$ is the intercept value, $\beta_k(u_i, v_i)$ is the regression coefficient of the *k*-th explanatory variable at the *i*-th observed location, x_{ik} is the observed value of the *k*-th explanatory variable at the *i*-th observation location, ε_i is the observed error *i*-th that is assumed to be identical, independent and normal distribution.

The parameter estimation on GWR model using Weighted Least Square (WLS) method by giving different weight for each location of observation with the following formula:

$$\hat{\beta}(u_i, v_i) = \left[X^T W(u_i, v_i) X \right]^{-1} X^T W(u_i, v_i) y \quad (2)$$

$\boldsymbol{W}(u_i, v_i)$ is matrix sized *n* x *n* which is diagonal elements indicate geographical weighting at the *i*-th observation location [7].

2.2 Geographically and Temporally Weighted Regression

The Geographically and Temporally Weighted Regression Model (GTWR) is an effective approach for solving spatial and temporal non-stationary problems [8]. GTWR model is an expansion of the GWR model by adding a time element (temporal) which is formulated as followed:

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i) x_{ik} + \varepsilon_i \quad (3)$$

y_i is the observed value of the response variable at location (u_i, v_i) and time t_i.

The regression coefficient $\hat{\beta}_i(u_i, v_i, t_i)$ at point *i* obtained by using Weighted Least Square (WLS) with the following criteria:

$$\hat{\beta}(u_i, v_i, t_i) = \left[X^T W(u_i, v_i, t_i) X \right]^{-1} X^T W(u_i, v_i, t_i) y \quad (4)$$

$W(u_i, v_i, t_i) = diag(w_{i1}, w_{i2}, ..., w_{in})$ is a weighted matrix at the observed location (u_i , v_i) and time t_i. The spatial - temporal spacing function of the GTWR method is written as follows [8]:

$$\left(d_{ij}^{ST}\right)^2 = \lambda\left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \mu\left[\left(t_i - t_j\right)^2\right] \quad (5)$$

Let τ be the ratio parameter of $\tau = \mu / \lambda$ with $\lambda \neq 0$ then the equation obtained as follows:

$$\frac{\left(d_{ij}^{ST}\right)^2}{\lambda} = \left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \tau\left[\left(t_i - t_j\right)^2\right] \quad (6)$$

The parameter τ obtained from the minimum cross-validation criteria by initializing the initial value τ as written below:

$$CV(\tau) = \sum_{i=1}^{n} \left[y_i - \hat{y}_{\neq i}(\tau) \right]^2 \quad (7)$$

Gaussian Kernel Function is the most used weighting function in the GTWR model written as follows:

$$w_{ij} = \exp\left[\left(\frac{d_{ij}^{ST}}{h_{ST}}\right)^2\right] \quad (8)$$

3. Methodology

This study used secondary data on economic growth through gross regional domestic product in Central Java Province obtained from the Central Agency on Statistics of Central Java Province. There are 29 districts and 6 cities that serve as the location of observed.

Table 1: Variables in research

Variables	Information
Y	Gross Regional Domestic Product
X_1	Workforce
X_2	Locally-generated revenue
X_3	District Minimum Wage
X_4	Human Development Index

This research used software R version 3.2.2. The GTWR modeling steps as shown in the flowchart in Figure 1.

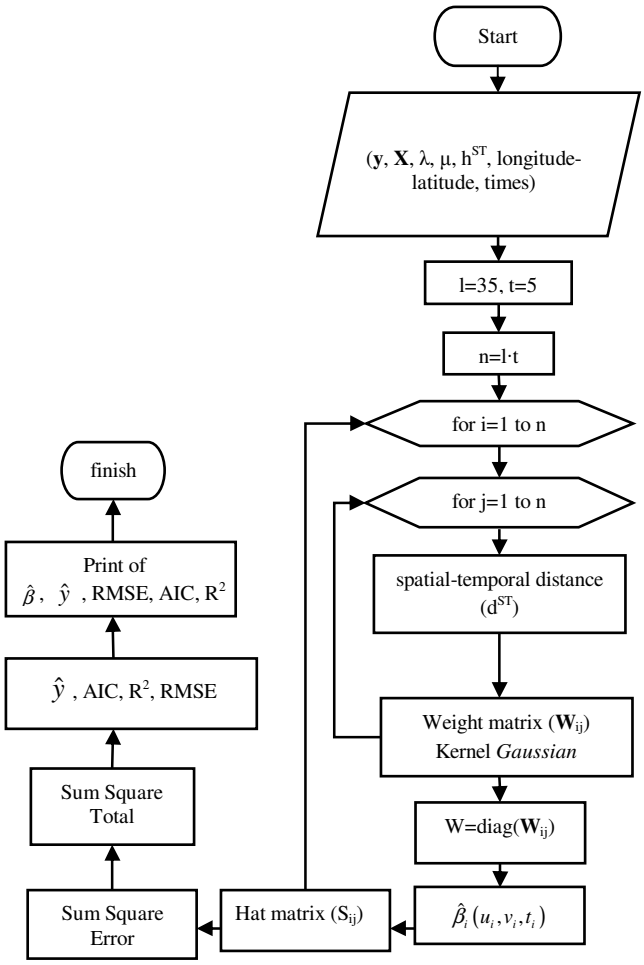

Fig. 1 Flowchart of GTWR model

4. Results and Discussion

4.1 Descriptions of Data

Table2: Descriptions of research data

Variables	Minimum	Maximum	Mean	Standard Deviation
Y	4.256	109.100	20.820	19.63652
X_1	57133	836837	462851	184499.7
X_2	60155	187212	1138364	144879.1
X_3	717000	1685000	966729	183892.7
X_4	59.66	80.96	68.71	4.69805

Table 3 shows that the correlation between Y with X_3 and X_4 has a fairly low correlation. However, these variables have an important part in developing economic growth in an area. Therefore, these variables remain involved in modeling.

Table 3: Correlation between response and explanatory variables

Variables	*Pearson* Correlation
Y with X_1	0.5193
Y with X_2	0.7306
Y with X_3	0.2688
Y with X_4	0.2690

4.2 Multicollinearity Test

The statistical test to see indication of linear correlation which is risky or not between explanatory variables is done multicollinearity assumption test by attention at Variance Inflation Factor (VIF) value. Table 4 shows that the VIF value for each variable is less than 5 either from the combined data in each year.

Table 4: VIF value on each explanatory of variables

Years	X_1	X_2	X_3	X_4
Full	2.2232	3.1103	1.7209	2.1607
2011	2.4717	2.8497	1.9311	2.5984
2012	2.1919	2.4139	1.6165	2.2868
2013	2.2393	3.3992	1.9889	2.1931
2014	2.4930	2.7423	1.5713	2.2104
2015	2.5462	2.9674	1.5753	2.3172

4.3 Heterogeneity Spatial Test

Table 5 shows that p-value obtained in the 5-year period is $4.367x10^{-10}$. At each time period also yields significant value. These result shows that there is spatial heterogeneity in economic growth data of Central Java. Syntax R program as follows:

```
dataGTWR.dat.f<-data.frame(dataGTWR)
require(lmtest)
bptest(  V1~X1  +  X2  +  X3  +  X4,  data  =
       dataGTWR.dat.f,studentize = F)
```

Table 5: Breusch-Pagan test

Years	Breusch-Pagan value	p-value
Full	49.604	$4.367x10^{-10}$ **
2011	16.403	0.002524 **
2012	13.845	0.007808 **
2013	12.208	0.015870 **
2014	9.0890	0.058900 *
2015	10.718	0.029930 **

(**) significant at the level 5%, (*) significant at the level 10%

Figure 2 shows that gross regional domestic product value increasing each year. Differences in each year indicate temporal heterogeneity. The syntax of the R program as follows:

```
abaru<-cbind(temp,y)
```

```
boxplot(y~temp,xlab="years",
        ylab="GRDP values (bilion rupiah)")
```

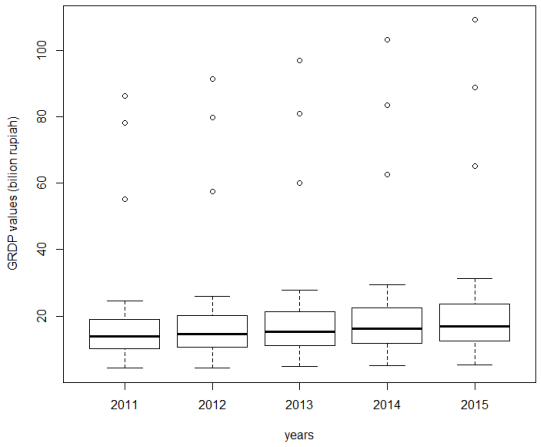

Fig. 2 Temporal of heterogeneity with boxplot

Boxplot width describes the diversity of data. In Figure 2 shows that each year the boxplot size tends to widen. Outliers in the boxplot also increase every year. There are three outliers in the plot. So it can be explained that there are three districts / cities that have the highest gross regional domestic product value compared with other regions. The further the value of the outlier with the midpoint of the boxplot indicates an increase in the gross regional domestic product value. This explains that the diversity of gross regional domestic product in the districts / cities of Central Java province tend to increase within 5 years.

4.4 GTWR Modeling

Table 6 describes the summary of parameter estimators using the GTWR method. The X_4 variable has the highest average value compared to other variables. Therefore, it can be concluded that the human development index (X_4) has a greater influence on economic growth for each district / city in Central Java. Syntax R program as follows:

```
bandwith.GWR.a<-bwd.gwr.twr(y=y,x=x,
                long.lat= coord,spasial=TRUE,
                kernel="gaussian")
taw.GTWR.a<-taw.cv.gtwr(y=y,x=x,
            bwd=bandwith.GWR.a,long.lat=coord,
            waktu=temp,spasial=T,
            kernel="gaussian")
lambda.miu.a<-lamda.cv.gtwr(y=y,x=x,
              bwd=bandwith.GWR.a,taw=taw.GTWR.a,
              long.lat=coord,waktu=temp,
              spasial=TRUE, kernel="gaussian")
bandwith.GTWR.a<-bwd.gtwr(y=y,x=x,
lambda=lambda.miu.a[[1]],miu=lambda.miu.a[[2]],
long.lat=coord,waktu=temp,spasial=TRUE,
kernel="gaussian")
```

```
output1.a<-prog.gtwr(y=y,x=x,long.lat=coord,
           waktu= temp, lambda=lambda.miu.a[[1]],
           miu=lambda.miu.a[[2]],
           bwd=bandwith.GTWR.a,kernel="gaussian")
```

Table 6: Summary of GTWR model parameters

Variables	Minimum	Maximum	Mean	Standard Deviation
Constant	-28.640	85.030	19.740	10.575
X_1	-47.831	71.353	1.9844	13.257
X_2	-16.528	81.170	4.4530	8.4402
X_3	-29.547	25.420	-1.846	5.2834
X_4	-106.43	54.614	6.1280	17.607

The best model obtained based on the value of R^2 which is bigger than other model and also generate smaller RMSE and AIC values. Table 7 shows that modeling economic growth in Central Java using GTWR gives better results than modeling using global regression.

Table 7: Comparison of models

Method	R^2	RMSE	AIC
Global Regression	0.640	11.747	2233.68
GTWR	0.998	0.8403	1051.98

4.5 Map of Parameter Estimator

The estimation values of GTWR model parameters for each location and time are shown in Figures 3 through 6. This figure explains the magnitude of each explanatory variable on economic growth that has different effects on each space and time.

Figure 3 shows influence of workforce (X_1) degree to gross regional domestic product. The darker the colors of each location indicate that Workforce has a major impact on economic growth. In figure 3 there is an area that has a darker color every year. Therefore, concluded that the influence of Workforce variables on economic growth increases every year in some areas in Central Java.

Figure 4 shows that the effect of Locally-generated revenue on economic growth tends to decrease at each location over time. Locally-generated revenue degree impact on economic growth also tends to be smaller than other variables. This can be seen from a lighter color than other variables.
Figures 5 and Figure 6 show colors that tend to be darker than other figures. It can be concluded that the District Minimum Wage and Human Development Index variables have a considerable influence on economic growth.

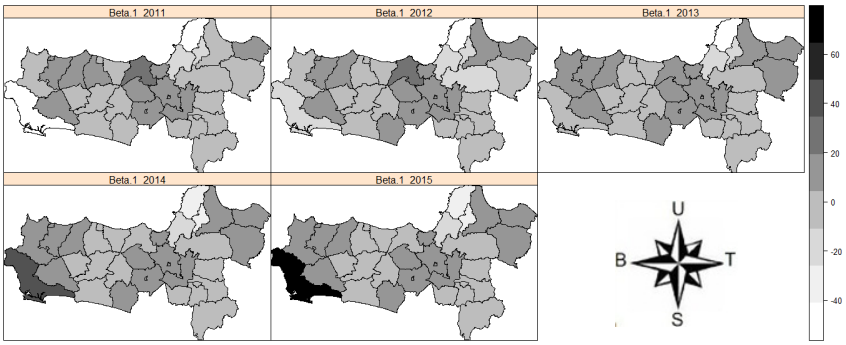

Fig. 3 Map of parameter estimator heterogeneity for X_1

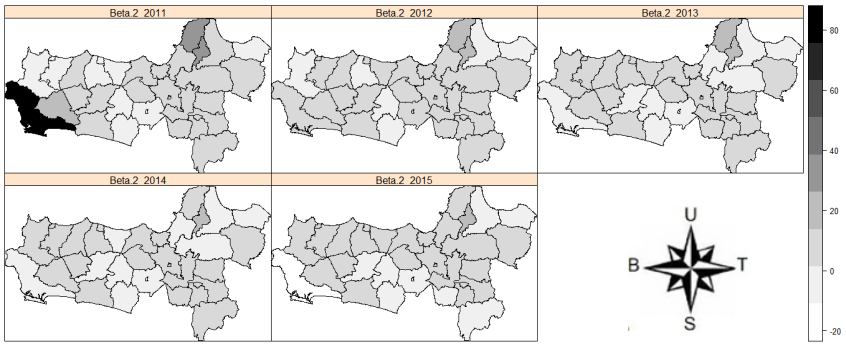

Fig. 4 Map of parameter estimator heterogeneity for X_2

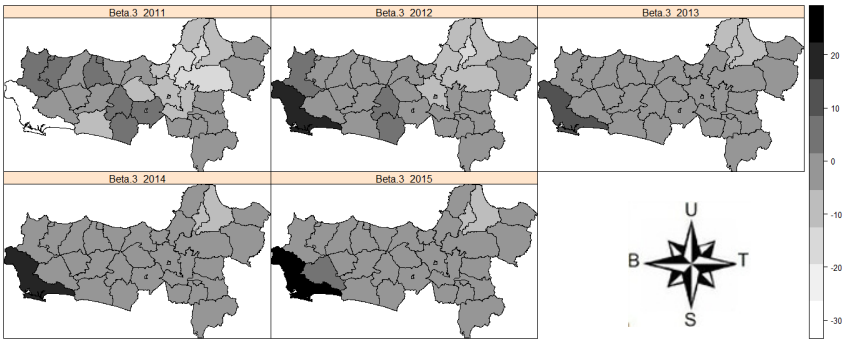

Fig. 5 Map of parameter estimator heterogeneity for X_3

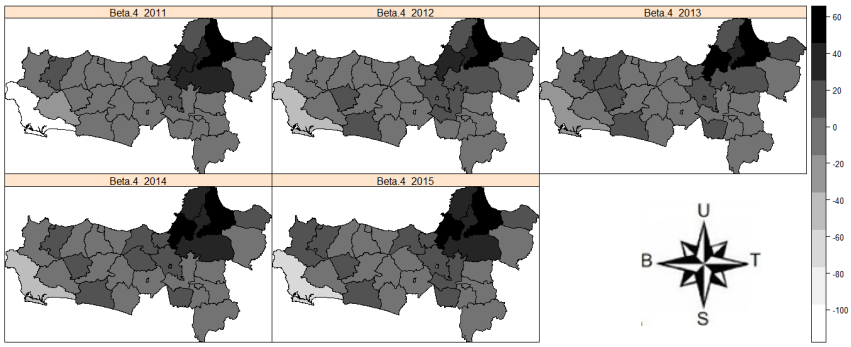

Fig. 6 Map of parameter estimator heterogeneity for X_4

5. Conclusion

The GTWR method is better than the global regression method on economic growth modeling in Central Java in 2011-2015. GTWR model is able to explain the R^2 value of 99.8% with the RMSE and AIC values of 0.8403 and 1051.98. The human development index (X_4) variables give greater influence to economic growth at each regency / city in Central Java during the period of 2011-2015.

This study is limited to the Gaussian kernel function in constructing a weighted matrix. In the next study can be done with other kernel functions to get better results.

References

[1] K. Aisyiah, Sutikno, and I.N. Latra, "Pemodelan Konsentrasi Partikel Debu (PM_{10}) pada Pencemaran Udara di Kota Surabaya dengan Metode Geographically-Temporally weighted Regression", Jurnal Sains dan Seni Pomits, Vol. 2, No.1, 2014, pp. 2337-3520.

[2] Bank Indonesia, "Bersinergi Menawal Stabilitas, Mewujudkan Reformasi Struktural", Laporan Perekonomian Indonesia, 2015. [Online]: Available at http://www.bi.go.id/id/ publikasi/ laporan-tahunan/ perekonomian/ Documents/ LPI_2015_web_final.pdf.

[3] Badan Pusat Statistik, "Berita Resmi Statisik: Pertumbuhan Ekonomi Indonesia Triwulan I-2015", 2015. [Online]: Available at https://www.bps.go.id/ pressrelease/2015/05/05/1143/ pertumbuhan-ekonomi-indonesia- triwulan-i-2015-tumbuh-4-71-persen.html.

[4] Badan Pusat Statistik, "JawaTengah dalam Angka: Jawa Tengah in Figures 2015", Jawa Tengah (ID): BPS.

[5] Boediono, Teori Pertumbuhan Ekonomi, Yogyakarta: BPFE, 1999.

[6] N.R. Draper, and H. Smith, Applied Regression Analysis, Second Edition, Canada : John Wiley & Sons, Inc, 1998.

[7] A.S. Fotheringham, R. Crespo, and J. Yao, "Geographically and Temporal Weighted Regression (GTWR)", Geographical Analysis, The Ohio State University, 2015, pp. 1-22.

[8] B. Huang, B. Wu, and M. Barry, "Geographically and Temporally Weighted Regression for Modeling Spatio-Temporal Variation in House Prices", International Journal of Geographical Information Science. Vol. 24, No.3, 2010, pp. 383-401.

[9] J. Liu, Y. Yang, S. Xu, Y. Zhao, Y. Wang, and F. Zhang, "A Geographically Temporal Weighted Regression Approach with Travel Distance for House Price Estimation", Article Entropy MDPI, Vol. 303, No. 18, 2016, pp. 1-13.

[10] C.I. Mei, "Geographically Weighted Regression Technique for Spatial Data Analysis", School of Science Xi'an Jiaotong University, 2005.

[11] E.W. Pambudi, "Analisis Pertumbuhan Ekonomi dan Faktor-Faktor yang Mempengaruhi (Kabupaten/Kota di Provinsi Jawa Tengah)", skripsi, Universitas Diponegoro, Semarang, Indonesia, 2013.

[12] Y.A. Rahman and A.L. Chamelia, "Faktor-faktor yang Mempengaruhi PDRB Kabupaten/Kota Jawa Tengah Tahun 2008-2012", Journal of Economics and Policy, Vol. 8, No. 1, 2015, pp. 88-99.

[13] D. Yu, Y.D Wei, and C. Wu, "Modeling Spatial Dimensions of Housing Prices in Milwaukee", Environment and Planning B: Planning and Design, Vol. 34, 2007, pp. 1085-1102.

Mr. M. Sholihin master student in Department of Statistics, Bogor Agricultural University. His main interests is on spatial analysis and statistical computation.

Dr. A.M. Soleh Currently worked as a lecture in Department of Statistics, Bogor Agricultural University. His main interests is on statistical computation and data mining.

Dr. A. Djuraidah Currently worked as a lecture in Department of Statistics, Bogor Agricultural University. Her main interests is on spatial analysis and statistical probability teory.