ENTHUSIASTIC
INTERNATIONAL JOURNAL OF APPLIED STATISTICS AND DATA SCIENCE
Volume 3, Issue 2, October 2023, pp. 189-201

Application of Geographically Weighted Regression Method on the Human Development Index of Central Java Province

Devi Octaviani Hasibuan [a,1,*], Heribertin Pau Teku [b,2], Maria Fatima Drostela Putri [b,3], Yudi Setyawan[b,4], Rokhana Dwi Bekti[b,5]

[a,b] Department of Statistics, IST AKPRIND Yogyakarta, Indonesia
[1] hasibuandevioctaviani@gmail.com*; [2] heribertin44@gmail.com; [3] mariafatimaputri1@gmail.com; [4] setyawan@akprind.ac.id; [5] rokhana@akprind.ac.id
* Corresponding author

ARTICLE INFO

Keywords
Geographically weighted regression
Human development index
Tableau

ABSTRACT

Spatial data are data containing information on the location or geography of a region on the representation of objects on earth. Geographically Weighted Regression (GWR) is a development of the Ordinary Least Square (OLS) theory into a weighted regression model that considers spatial effects, resulting in a parameter estimation that can only be used to predict each location where the data are observed. The Human Development Index (HDI) is an essential indicator for measuring success in efforts to build human quality of life. HDI data regencies/cities in Central Java are interconnected, so it is said to be spatial data and there are spatial effects in it. Therefore, the GWR method was applied to obtain faculties affecting HDI in Central Java Province. The data used were secondary data in 2020. The determination coefficients of the GWR model ranged between 76.09% and 87.16%. If the variable values of population density and Gross Regional Domestic Product (GRDP) increase by one unit in each district/city in Central Java Province, the HDI variable value increases. These results were visualized on a dashboard providing information about the characteristics of HDI and independent variables, GWR parameter estimates, and the significance of independent variables in each regency/city.

1. Introduction

Over the past few decades, spatial analysis has evolved into two major research fields: spatial data analysis and spatial modeling. Spatial statistics is one of the areas of interest within geography-based statistics. Using data, the existence of spatial effects is something that frequently occurs between two regions or the geographical location of a place.

Spatial data are data comprising information on the location or geography of a region. Hence, it does not only contain what is measured. One method that can be used in spatial analysis is Geographically Weighted Regression (GWR), which is a development of the Ordinary Least Square (OLS) theory into a weighted regression model by paying attention to spatial effects. Thus, parameter estimation that can only be used to predict each point or location where the data are observed and inferred. GWR analysis is a method used to process spatial data. The GWR model is a model that pays attention to geographical factors as free variables that affect response variables. It will generate a local model parameter estimator for each point or location where the data are observed.

The achievement of human development is measured by paying attention to three essential aspects: longevity and healthy living, knowledge, and a decent standard of living. Human Development Index (HDI) is an indicator used to see the development in the long term. In general, Indonesia's human development continued to progress during the period from 2010 to 2020. HDI is an indicator used to see the development in the long term. In general, Indonesia's human development continued to progress during the period from 2010 to 2020.

According to the United Nations Development Programme (UNDP), HDI is defined as a process of enlarging the choice of people. It measures the achievement of development results from an area/region in three basic dimensions of development: longevity, knowledge/education level, and a decent standard of living. According to the Central Statistics Agency of Central Java Province, human development in Central Java advanced in 2020, as indicated by an increase in the Central Java HDI. Despite the impact of the COVID-19 outbreak, the HDI of Central Java in 2020 was still able to grow positively by 0.14 points, from 71.73 points in 2019 to 71.87 points in 2020.

Geographically Weighted Regression (GWR) is a statistical technique that applies bias to data with spatial effects in order to model the multiplicity of relationships in the visualization of spatial dimensions. In contrast to global regression, the GWR method can model spatial distance-weighted relationships. Spatial effects that occur between regions can be divided into two types, namely spatial dependence and spatial heterogeneity [1]. The fundamental thing of the GWR method is the proximity between regions which is shown by the weighting matrix. The closer the distance between regions, the greater the weight value will be. As a result, the GWR method will provide a more precise statistical analysis of the spatial relationship between multiple variables, as it can overcome the issue of space diversity.

Other research employing the GWR method has studied factors affecting the reading literacy activity index in Indonesia. In the study, the results showed that the best modeling of several regression methods is the GWR model because it has greater model goodness than the linear regression model, which was 92.46%. There was a significant influence of the factors of the literacy index figures in Indonesia. In grouping variables, a significant influence on the literacy activity index was obtained by 11 groups. Whereas, in group 1, there was only one variable that had a significant influence on the literacy activity index, namely the percentage of Latin literacy, which was found in Papua Province. Meanwhile, in group 11, all independent variables had a significant effect on the literacy activity index. This was evident in Jambi Province, South Sumatra Province, and Lampung Province [2].

Research with similar methods on the pneumonia cases in East Java Province has been carried out. In this study, there was an aspect of spatial heterogeneity in pneumonia cases in East Java in 2016. As a result, it was deemed necessary to analyze it with the GWR method. The results of the GWR analysis indicated that a value of the sum of squares of GWR model errors was smaller than that of the sum of squares of multiple linear regression model. This finding suggested that the GWR model was more feasible to describe the pneumonia cases that occurred in East Java in 2016 [3].

Research with similar methods was also carried out on the analysis of social vulnerability and its effect on social problems in Semarang City. The results showed that the GWR model yielded a positive relationship between social problems (Y) and population density (X1), the number of unemployed (X4), and the average length of schooling (X5), with the negative relationship to the sex ratio (X2) and the life dependence rate (X3). The GWR model showed a degree of significance only on the population density factor with a value $t_{hitung} = 2.065 \geq t_{(0.025;452)} = 2.059$, and did not differ significantly from the global regression model. However, the GWR model provided a better model with a higher coefficient of determination value of R^2 of 0.326 and a lower residual sum of squares (RSS) of 15.733 [4].

p-ISSN 2798-253X
e-ISSN 2798-3153

The reason for using the GWR model is that the HDI case data in Central Java shows a spatial effect. Visualization of the data will also help facilitate interpretation. Visualization is defined as a method of presenting data or problems in a graphic format or image form that is easy to understand. The use of data visualization presented by researchers through the utilization of diverse graphs or interactive visuals will make it easier for readers to understand information quickly and effectively.

The research in this case applied the GWR method which was expected to be able to produce the right HDI model in each district/city in Central Java. Furthermore, the data were also visualized on the Tableau dashboard to provide more interesting and easy-to-understand illustrations.

2. Method

GWR is a spatial method involving the geographical conditions of each region as one of the factors suspected of influencing dependent variables. GWR develops by adding a geographic at each location point for each parameter. In general, this development is based on the concept of nonparametric regression applied to regression model. The obtained GWR model was used to predict the magnitude of the response variable with the resulting parameters where each parameter was obtained from the location of the object. The fundamental thing of the GWR method is the proximity between regions which is shown by the weighting matrix. The closer the distance between regions, the greater the weight value will be. The general equation of GWR is as in the following equation:

$$y_i = \beta_0(u_i, v_i) + \sum_{t=1}^{p} \beta_j(u_i, v_i)x_{ij} + \varepsilon_i \quad (1)$$

$i = 1,2, \dots, n$

where:

y_i	: the value of the variable bound to the *i*th observation
x_{ij}	: the value of the *j*th free variable on the *i*th observation
$\beta_0(u_i, v_i)$	: constants on the *i*th observation
$\beta_j(u_i, v_i)$	: the value of the free variable function *xj* on the *i*th anniversary
p	: number of free variables
(u_i, v_i)	: coordinate point of the *i*th observation location
ε	: random error

The estimated parameters at each location *i* in equation (1) via Weighted Least Square (WLS) are:

$$\hat{\beta}(i) = (X^T W(i) X)^{-1} X^T W(i) y \quad (2)$$

where *X* is data matrix of independent variables, *y* is the vector of dependent variables, and *W(i)* is weighting matrix.

GWR model with adaptive Gaussian kernel weighting appeared to be more suitable for modeling cases of malnutrition of children under five in West Java than the OLR model and the GWR model with fixed Gaussian kernel weighting. It can be seen from the sum of the residual squares of the GWR model with the adaptive Gaussian kernel weighting and the value of the coefficient of determination of the GWR model with the adaptive Gaussian kernel weighting [5].

Based on the values of sum of squares error (SSE), R^2, and Akaike information criterion (AIC), the good model was the near neighborhood kernel weighting. Whereas if it was based on a significant value with α=5%, a good model would be a bi-square kernel weighting by using 1 free variable i.e. variable IPM because there were 17 regencies/cities. Thus, it can be concluded that the best model in this research was a bi-square kernel weighting using 1 free variable, namely the HDI variable as there was no significant location in the GWR model using a near neighborhood kernel weighting [6].

The data used in this study were secondary data obtained from the Staistics Indonesia of Central Java Province. The data used were data on the HDI, population density, percentage of poor people, and gross regional domestic product on a constant price basis in 2020. Each variable contains 35 observation data. The variables used are presented in Table 1.

Table 1. Research Variables

Variable Type	Variable Name	Operational Definition	Data Scale
Dependent	HDI	Human Development Index	Ratio
Independent	KP	Population Density Figures	Ratio
	PPM	Percentage of Poor Population	Ratio
	GRDP	Gross Regional Domestic Product on a Constant Price Basis	Ratio

The methods used in this study were descriptive analysis and spatial analysis of GWR with the following stages:

a. Selecting variables believed to influence the HDI to be involved in shaping the model.
b. Identifying descriptive analysis and spatial patterns of HDI variables to determine the characteristics of the research data used through thematic maps.
c. Conducting linear regression analysis (Ordinary Least Square, OLS) by testing assumptions, namely:
   - Estimation of OLS model parameters
   - Parameter testing with simultaneous and partial tests.
   - Assumption of normality using the Kolmogorov-Smirnov test.
   - Assumption of homoscedasticity using the Glejser test.
   - Assumption of multicollinearity based on VIF values.
   - Autocorrelation assumptions using the Durbin-Watson test.
d. Conducting a GWR analysis with the following stages:
   - Determining u_i and v_j based on latitude and longitude for each regency/city in Central Java Province obtained from the shp map with software.
   - Creating a spatial weighting with a queen contiguity weight.
   - Conducting spatial effect testing, that was, spatial dependency effect with Moran's I test.
   - Calculating Euclidean distances between observational locations based on geographical position/avalanche point and latitude.
   - Determining the optimum bandwidth value using cross-validation (CV).
   - Assessing the parameters of the GWR model by using the WLS method.
   - Conducting a goodness of fit test of the GWR model.
   - Testing the significance of GWR model parameters as well as mapping the significance of GWR model parameters.
   - Interpretating the GWR model.
e. The selection of the best models of the OLS model and the GWR model was based on the values of R^2 and AIC.

3. Results and Discussion

3.1. Description of Data and Spatial Pattern Distribution

Table 2 shows the averages, minimum, maximum, and standard deviation values of each variable. The average HDI in 35 regencies/cities in Central Java was 72.51. The region with low HDI was the Brebes Regency (66.11), one with a high HDI was Salatiga City (83.14).

Table 2. Descriptive Statistical Analysis of Variables

Variable	Average	Minimum Value	Maximum Value	Standard Deviation
HDI	72.51	66.11	83.14	4.42
KP	2,092.51	490.00	11,353.00	2,417.86
PPM	11.01	4.34	17.59	3.52
GRDP	27,543.47	6,314.05	137,951.30	25,179.86

The average KP in 35 regencies/cities in Central Java was 2,092.51. The area with a low KP was Blora Regency (490 people/km^2). The area that with a high KP was Surakarta City (11,353 people/km^2). The average PPM in 35 regencies/cities in Central Java was 11.01. The area with a low PPM was Semarang City (4.34%). The area with a high PPM was Kebumen Regency (17.59%). The average Gross Regional Domestic Product (GRDP) in 35 regencies/cities in Central Java was

p-ISSN 2798-253X
e-ISSN 2798-3153

27,543.47. The area with a low GRDP was Magelang City (6,314.05 billion rupiah). The area with a high GRDP was Semarang City (13,7951.30 billion rupiah).

The visualization results of spatial patterns of HDI data of regencies/cities in Central Java Province are presented in Fig. 1. The achievement of the HDI in Central Java regencies/cities in 2020 was divided into four groups: very high (HDI≥80), high (70≤IPM≤80), medium (60≤IPM≤70), and low (HDI<60). It can be known that regencies/cities with a high HDI were in the northern Central Java. It indicates that there was a spatial effect on the HDI data of regencies/cities in Central Java Province.

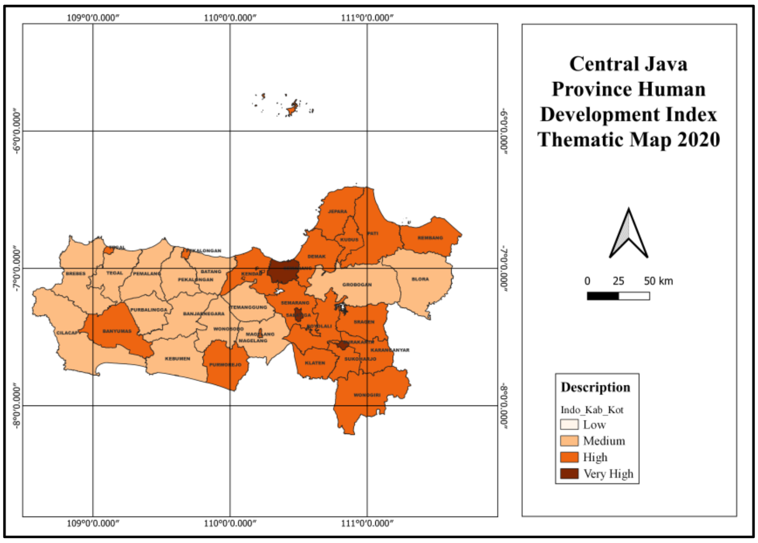

Fig. 1 Central Java HDI thematic map in 2020.

3.2. Ordinary Least Square

The results of estimating the parameters of the model OLS regression analysis produce the parameter values in Table 3.

Table 3. Output of Multiple Linear Regression

Variable	Estimation	*p*-Value (t-test)	*p*-Value (F-test)
Intercept	76.13	$< 2 \times 10^{-16}$	
KP	0.0008343	0.000448	1.699×10^{-7}
PPM	-0.5611	0.000854	
GRDP	0.00002957	0.133094	

Based on Table 3, the obtained a multiple linear regression model was IPM=76.13 + 0.0008343KP -0.5611PPM + 0.00002957GRDP.

On the F-test, it is known that the coefficient was not feasible to enter the model with a 95% confidence level and H0 is rejected if the *p*-value<α. Therefore, that it can be inferred from the overall independent variable obtained *p*-value <α then H0 is rejected. So, the coefficient was feasible to enter the model with a confidence level of 95% with a confidence level of 95%.

In the t-test, it was found that independent variables (KP, PPM, GRDP) did not have a significant effect on the dependent variable (HDI) partially with a confidence level of 95% and H0 was rejected if the *p*-value <α. Hence, it can be concluded that the KP and PPM variables obtained *p*-value <α, then H0 was rejected. Thus, the KP and PPM variables had a significant effect on the dependent variable (HDI) partially with a confidence level of 95%. Meanwhile, if the GRDP variable obtain *p*-value >α, then H0 is not rejected. Thus, the GRDP variable did not have a significant effect on the dependent variable (HDI) partially with a confidence level of 95%.

p-ISSN 2798-253X
e-ISSN 2798-3153

Table 4. Residual Normality

Test	Statistical D	p-Value
Kolmogorov-Smirnov	0.064646	0.9965

From Table 4, it is known that H0 was normally distributed residual data with a confidence level of 95% and H0 was rejected if the p-value <α. Thus, if P p-value =0.9965>α=0.05, then H0 was not rejected. It indicated that residual data were normally distributed with a 95% confidence level.

Table 5. Heteroscedasticity

Test	BP Statistics	p-Value
Breusch Pagan	5.2597	0.1537

From Table 5, tt is known that H0 was not heteroscedasticity of data with a confidence level of 95% and H0 was rejected if p-value <α. So, if p-value = 0.1537>α = 0.05 could be obtained, then H0 was not rejected. In indicated that there was no heteroscedasticity of data with a 95% confidence level.

Table 6. Autocorrelation

Test	DW Statistics	p-Value
Durbin-Watson	1.2373	0.005296

From Table 6, it is known that H0 was not correlated with a 95% confidence level and H0 was rejected if p-value <α. Hence, if p-value = 0.005296 <α = 0.05, then H0 was rejected. It suggested that there was an autocorrelation with a 95% confidence level.

To identify multicollinearity, the Variance Inflation Factor (VIF) value must be examined. A VIF value below 10 indicates the absence of multicollinearity in the data. The results of the multicollinearity test are shown in Table 7 using the VIF values as follows:

Table 7. Multicollinearity

Variable	VIF	Conclusion
KP	1.244563	No multicollinearity
PPM	1.353183	No multicollinearity
GRDP	1.098557	No multicollinearity

3.3. Spatial Effects Testing

The spatial dependency test used in this study was Moran's I test with the output can be seen in Table 8.

Table 8. Output Moran's I

Test	Moran I Statistics	p-Value	Expectations
Moran's I	0.2379	0.01948	-0.0294

From this table, it is known that H0 had no spatial autocorrelation in the HDI data with a confidence level of 95%. H0 was rejected if the p-value <α=0.05. The obtained H0 was rejected, meaning that there was spatial autocorrelation in the HDI data with a confidence level of 95%.

The value of I = 0.2379 was greater than E(I) = -0.0294, meaning that there is a positive, but significant autocorrelation. A positive coefficient value signifies that a HDI in one area will lead to high HDI in nearby areas as well.

3.4. Geographically Weighted Regression

From the results of the OLS regression model, it is known that that GWR modeling was carried out since there was an occurrence of autocorrelation. The steps performed on the GWR were determining u_i and v_i location points, conducting queen contiguity spatial weighting, determining the optimum bandwidth value, estimating the GWR Model Parameters, conducting the model conformity test, and conducting the GWR model parameter significance test.

p-ISSN 2798-253X
e-ISSN 2798-3153

The location points or u_i and v_i in this study were determined based on the southern latitude and east longitude for each district and city in Central Java Province. The function of the longitude and latitude values was to map the variable characteristics of each district and city. The results of the values u_i and v_i can be seen in Appendix 1.

The specification of the weighting matrix is to represent information on the scope and intensity of the spatial effects of a unit of location within a geographic system. In this study, queen contiguity weighting was used to determine neighboring relationships or neighboring locations, which showed a higher spatial dependency relationship than those that are more distant (Law I Tobler). Distance weighting was obtained from the latitude and longitude coordinates of a point or area. The output of spatial weighting queen contiguity in this study is in Appendix 2.

The weighted formula is a location that is side by side or the point of the blade meets the location of the concern given weighting $W_{ij} = 1$, while for other locations it is $W_{ij} = 0$. Determination of the optimum bandwidth value (b) coefficient of variation (CV) criteria. Adaptive kernel weighting will generate bandwidth values that will be different for each location. Table 9 is a table of bandwidth values using the adaptive Gaussian kernel weighting function.

Table 9. Rated Bandwidth

No	Districts/Cities	Bandwidth	No	Districts/Cities	Bandwidth
1	Cilacap District	0.7795	19	Kudus District	0.6085
2	Banyumas District	0.5894	20	Jepara District	0.7749
3	Purbalingga District	0.5084	21	Demak District	0.4925
4	Banjarnegara District	0.4649	22	Semarang District	0.3989
5	Kebumen District	0.6475	23	Temanggung District	0.3716
6	Purworejo District	0.6540	24	Kendal District	0.4721
7	Wonosobo District	0.4523	25	Batang District	0.5281
8	Magelang District	0.4170	26	Pekalongan District	0.5359
9	Boyolali District	0.4127	27	Pemalang District	0.4680
10	Klaten District	0.4466	28	Tegal District	0.5375
11	Sukoharjo District	0.5411	29	Brebes District	0.7686
12	Wonogiri District	0.8067	30	Magelang City	0.4440
13	Karanganyar District	0.5955	31	Surakarta City	0.4476
14	Sragen District	0.5144	32	Salatiga City	0.3923
15	Grobogan District	0.4777	33	Semarang City	0.5017
16	Blora District	0.7734	34	Pekalongan City	0.5626
17	Rembang District	0.9471	35	Tegal City	0.6617
18	Pati District	0.7083			

Estimation of the GWR model with adaptive Gaussian kernel weighting functions for each district and city in Central Java Province can be seen in Table 10.

Table 10. GWR Model Estimation

Districts/Cities	Intercept	KP	PPM	GRDP
Cilacap District	72.60773	0.000895	-0.36768	0.000023
Banyumas District	71.6122	0.00093	-0.30211	0.000024
Purbalingga District	71.47199	0.000955	-0.30716	0.000033
Banjarnegara District	72.61336	0.000959	-0.37796	0.000044
Kebumen District	74.84792	0.00089	-0.48828	0.000030
Purworejo District	76.56969	0.000863	-0.583	0.000031
Wonosobo District	74.89867	0.000941	-0.50501	0.000042
Magelang District	78.46321	0.000856	-0.69932	0.000026
Boyolali District	80.38886	0.000794	-0.77973	0.000014
Klaten District	80.09247	0.000769	-0.74393	0.000018
Sukoharjo District	79.56144	0.000799	-0.71178	0.000018
Wonogiri District	78.35587	0.000838	-0.65224	0.000024
Karanganyar District	79.13147	0.000825	-0.68613	0.000019
Sragen District	79.24297	0.000839	-0.6934	0.000016
Grobogan District	78.31503	0.000912	-0.64821	0.000017

p-ISSN 2798-253X
e-ISSN 2798-3153

Districts/Cities	Intercept	KP	PPM	GRDP
Blora District	76.94073	0.000926	-0.55954	0.000025
Rembang District	76.23765	0.00093	-0.52578	0.000030
Pati District	76.08449	0.000961	-0.51872	0.000030
Kudus District	76.24582	0.000972	-0.53424	0.000029
Jepara District	75.75509	0.000926	-0.51892	0.000035
Demak District	77.1397	0.000962	-0.61157	0.000026
Semarang District	79.77886	0.000851	-0.78866	0.000018
Temanggung District	76.64983	0.000949	-0.63073	0.000035
Kendal District	76.05245	0.000895	-0.59269	0.000040
Batang District	74.08366	0.000878	-0.48614	0.000049
Pekalongan District	72.62239	0.000907	-0.40498	0.000049
Pemalang District	70.32666	0.000987	-0.274	0.000043
Tegal District	70.21111	0.000992	-0.26114	0.000033
Brebes District	72.15604	0.000904	-0.36615	0.000031
Magelang City	78.01721	0.000866	-0.67443	0.000029
Surakarta City	80.20704	0.000777	-0.74076	0.000014
Salatiga City	80.15338	0.000827	-0.80149	0.000016
Semarang City	77.20354	0.000905	-0.63959	0.000031
Pekalongan City	73.02878	0.000876	-0.43682	0.000052
Tegal City	71.45625	0.000933	-0.3422	0.000037

Suppose a model will be formed in Tegal Regency, then the GWR model is $y_{(Kab.Tegal)}$=70.21111+0.000992KP-0.26114PPM+0.000033GRDP.

Based on the model, if the variable population density (KP), percentage of poor people (PPM), and GRDP were of constant value, then the value of the HDI variable was 70.21111. Furthermore, if the population density (KP) variable experienced an increase in one unit and other independent variables were of constant value, then the HDI variable increased by 0.000992. Then, if the percentage variable of the poor population (PPM) experienced an increase in one unit and other independent variables were of constant value, the HDI variable experienced a decrease of 0.26114. Then, if the GRDP variable experienced an increase in one unit and other independent variables were of constant value, the HDI variable had an increase of 0.000033. Likewise with modeling other counties and cities.

After obtaining the results of GWR global and local modeling, it was continued with a model comparison test, which was to determine that there was no significant difference between the global and local models. In Table 11, conclusions can be obtained regarding the model suitability test.

Table 11. Goodness of Fit Model GWR

	df	Sum Square	Mean Square	*F*-Value
OLS Residuals	4	223.52		
GWR Improvement	7.6922	116.04	15.0860	
GWR Residuals	23.3078	107.48	4.6111	3.2716

After testing the suitability of the GWR model with the adaptive function of the Gaussian kernel, it was found that the value of F_{count} = 3.2716 and the value of $F_{(0.05,3.35)}$ = 2.91, meaning that H0 is rejected. It indicated that there was a significant difference between the OLS regression model and the GWR. Hence, it can be concluded that the GWR model has better goodness of fit than the global regression model.

The model parameter significance test was carried out to see which independent variables affected the HDI in each district and city in Central Java Province. With critical areas stating that H0 was rejected if $|t_{count}|>t_{(0.025;31)}$=2.03951. The KP Variable was significant in 29 regencies and 6 cities in Central Java Province. The PPM variable was significant in 25 regencies and 6 cities in Central Java Province. The GRDP variable was significant in 6 regencies and 1 city in Central Java Province.

 p-ISSN 2798-253X
e-ISSN 2798-3153

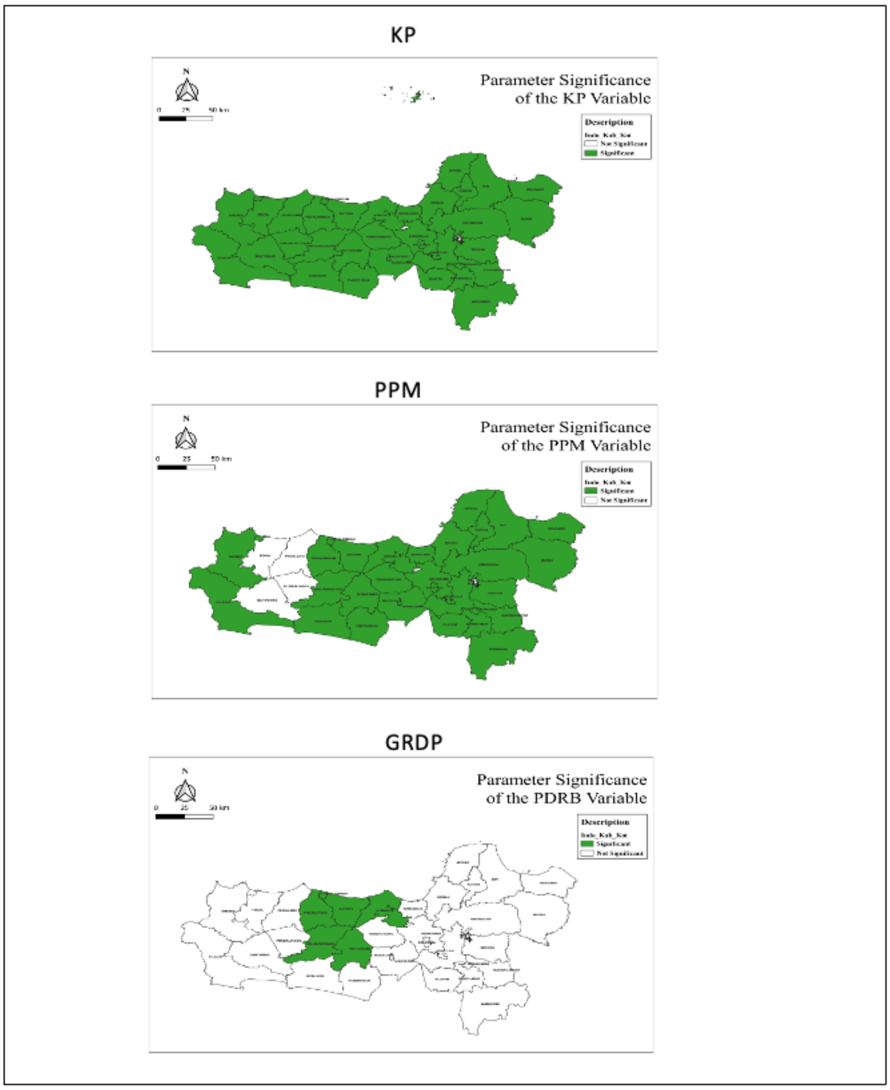

Fig. 2 Significance of GWR model parameters in each district and city.

Fig. 2 presents the significance of the GWR model parameters in each district and city of Central Java Province. The green color represents the significant, while white represents insignificant. It can be known that each significant variable in each location was different. It showed that there was a spatial effect on variables affecting the HDI of regencies/cities in Central Java Province in 2020.

3.5. Best Model

The selection of the best model was used to determine the best model that was good in estimating the opportunities of each model from the existing data. In this study, two models (OLS and GWR) were compared based on R^2 and AIC values. The selection of the best model can be seen in Table 12.

Table 12. Comparative Result Value

Criterion	OLS Models	GWR Models
R^2	0.6317	0.8385
AIC	189.221	147.8095

Based on Table 12, GWR model had a greater value of R^2 than the OLS model, which was 0.8385, and a smaller AIC compared to the OLS model, which was 147.8095. Thus, it is proven that the GWR model is good to use as a model in estimating model parameters.

p-ISSN 2798-253X
e-ISSN 2798-3153

3.6. Visualization Results

The visualization that has been built is presented in Fig. 5 or accessible at https://tabsoft.co/3b8juRi Dashboard, which displays the scatter plot of each variable that affects the HDI of Central Java Province in 2020. Scatter plots in visualization were used to observe relationships between variables.

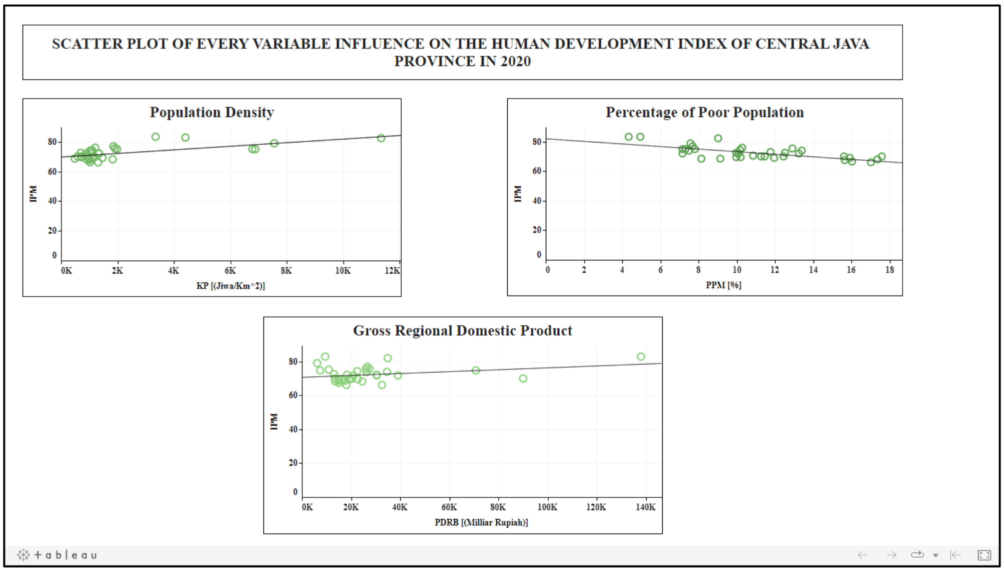

Fig. 3 Visualization results on the tableau dashboard.

4. Conclusion

A model is considered good if it has the smallest AIC value with a larger criterion of R^2. The GWR model with adaptive Gaussian kernel weighting appeared to be more suitable for modeling the HDI in Central Java in 2020 than the OLS model. It is evident from the AIC value of 147.8095 and the coefficient of determination value of 83.85% that the GWR model with adaptive Weighting of the Gaussian kernel is appropriate.

In the model parameters' significance test, it can be seen that the KP variable was significant in 29 regencies and 6 cities in Central Java Province. The PPM variable was significant in 25 regencies and 6 cities in Central Java Province. Significant GRDP variable in 6 Regencies and 1 City in Central Java Province.

Acknowledgment

The authors would like to thank the Institut Sains and Teknologi AKPRIND Yogyakarta, Indonesia, who has supported this research through grant funding and facilities to carry out this research.

References

[1] D.S. Susanti, A.S. Lestia, and Y. Sukmawaty, "Pemodelan Tingkat Kesejahteraan Peduduk Propinsi Kalimantan Selatan dengan Pendekatan Geographically Weighted Regression (GWR)," in *Prosiding Seminar Nasional MIPA 2016*, Oct. 2016, pp. 184–192.

[2] D.T. Hapsery Alfisyahrina, "Aplikasi Geographically Weighted Regression (GWR) untuk Pemetaan Faktor yang Mempengaruhi Indeks Aktivitas Literasi Membaca Di Indonesia," *Jurnal Riset dan Aplikasi Matematika*, vol. 5, no. 2, pp. 80–90, Oct. 2021, doi: 10.26740/jram.v5n2.p80-91.

[3] M. Anjas A, I.K.G. Sukarsa, and I.P.E.N. Kencana, "Penerapan Metode Geographically Weighted Regression (GWR) pada Kasus Penyakit Pneumonia di Provinsi Jawa Timur," *E-Jurnal Matematika*, vol. 8, no. 1, pp. 27–34, Feb. 2019, doi: 10.24843/MTK.2019.v08.i01.p231.

[4] I.M. Hida, A. Sukmono, and H.S. Firdaus, "Analisis Kerentanan Sosial dan Pengaruhnya terhadap Masalah Sosial dengan Geographically Weighted Regression (GWR) (Studi Kasus: Kota Semarang)," *Jurnal Geodesi Undip*, vol. 9, no. 1, pp. 237–246, Dec. 2019, doi: 10.14710/jgundip.2020.26168.

[5] A. Maulani, N. Herrhyanto, and M. Suherman, “Aplikasi Model Geographically Weighted Regression (GWR) untuk Menentukan Faktor-Faktor yang Mempengaruhi Kasus Gizi Buruk Anak Balita di Jawa Barat,” *Jurnal EurekaMatika*, pp. Vol. 4, No. 1, pp. 46–63, May 2016, doi: 10.17509/jem.v4i1.10454.

[6] A.F.J.P. Dao and Kartiko, “Analisis Geographically Weighted Regression Menggunakan Pembobot Adaptive Bisquare dan Near Neighbourhood Kernel (Studi Kasus: Kemiskinan di Provinsi Sulawesi Tenggara tahun 2017),” *Jurnal Statistika Industri dan Komputasi*, vol. 4, no. 2, pp. 75 –93, Jul. 2019, doi: 10.34151/statistika.v4i2.1950.

 p-ISSN 2798-253X
e-ISSN 2798-3153

Appendix

Appendix 1. Data

Districts/Cities	HDI	KP	PPM	GRDP	Ui	Vi
Cilacap District	69.95	915	11.46	89,934.72	108,890.41	-7,49096
Banyumas District	71.98	1331	13.26	39,121.62	109,175.58	-7,45515
Purbalingga District	68.97	1474	15.9	17,174.55	109,407.32	-7,32292
Banjarnegara District	67.45	994	15.64	15,045.88	109,657.42	-7,35188
Kebumen District	69.81	1114	17.59	19,526.36	109,617.41	-7,65527
Purworejo District	72.68	705	11.78	13,132.49	109,966.04	-7,69959
Wonosobo District	68.22	896	17.36	13,569.63	109,907.22	-7,41554
Magelang District	69.87	1179	11.27	22,861.47	110,246.32	-7,5012
Boyolali District	74.25	1054	10.18	22,399.52	110,650.65	-7,41856
Klaten District	75.56	1915	12.89	27,482.91	110,619.9	-7,6864
Sukoharjo District	76.98	1856	7.68	26,616.94	110,834.63	-7,68086
Wonogiri District	70.25	582	10.86	20,561.6	111,000.4	-7,9202
Karanganyar District	75.86	1202	10.28	26,142.87	111,025.08	-7,61623
Sragen District	73.95	1038	13.38	26,367.26	110,978.58	-7,38779
Grobogan District	69.87	722	12.46	19,379.68	110,927.1	-7,11688
Blora District	68.84	490	11.96	17,464.95	111,387.64	-7,07596
Rembang District	70.02	727	15.6	13,409.63	111,461.41	-6,77556
Pati District	71.77	889	10.08	30,545.61	111,041.41	-6,74342
Kudus District	75	1997	7.31	70,662.04	110,869.72	-6,78907
Jepara District	71.99	1119	7.17	20,969.88	110,783.95	-6,58371
Demak District	72.22	1338	12.54	18,374.56	110,632.01	-6,91112
Semarang District	74.1	1108	7.51	34,687.62	110,476.42	-7,27534
Temanggung District	69.57	943	9.96	14,890.75	110,135.63	-7,25786
Kendal District	72.29	911	9.99	30,443.69	110,156.03	-7,03778
Batang District	68.65	1017	9.13	15,030.58	109,861.47	-7,0213
Pekalongan District	69.63	1157	10.19	16,047.51	109,620.42	-7,05678
Pemalang District	66.32	1316	16.02	18,146.6	109,394.98	-7,03654
Tegal District	68.39	1823	8.14	24,502.62	109,158.4	-7,03109
Brebes District	66.11	1040	17.03	32,640.97	108,927.52	-7,05932
Magelang City	78.99	7567	7.58	6,314.05	110,220.13	-7,47714
Surakarta City	82.21	11353	9.03	34,827.19	110,823.39	-7,55808
Salatiga City	83.14	3353	4.94	9,503.16	110,498.43	-7,33827
Semarang City	83.05	4424	4.34	137,951.3	110,389.54	-7,02042
Pekalongan City	74.98	6788	7.17	7,337.83	109,677.89	-6,89301
Tegal City	75.07	6901	7.8	10,953.33	109,115.77	-6,86882

p-ISSN 2798-253X
e-ISSN 2798-3153

Appendix 2. Spatial Weighting Queen Contiguity

Lokasi	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
2	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
3	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
4	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
5	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0
9	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
10	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0
22	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	1	1	0	0
23	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0
25	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0
26	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0
27	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
28	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1
29	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
30	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0

p-ISSN 2798-253X
e-ISSN 2798-3153