Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 16(1) 2023: 47-58
http://ejournal.undip.ac.id/index.php/media_statistika

GEOGRAPHICALLY WEIGHTED PANEL LOGISTIC REGRESSION SEMIPARAMETRIC MODELING ON POVERTY PROBLEM

Aliyah Husnun Azizah[1], Nurjannah[1], Adji Achmad Rinaldo Fernandes[1], Rosita Hamdan[2]
[1] Department of Statistics, University of Brawijaya, Indonesia
[2] Department of Development Economics, University Malaysia Serawak, Malaysia

e-mail: *fernandes@ub.ac.id*

DOI: 10.14710/medstat.16.1.47-58

Article Info:
Received: 24 July 2023
Accepted: 20 September 2023
Available Online: 4 October 2023

Keywords:
Geographically Weighted Regression; Geographically Weighted Panel Logistic Regression Semiparametric; Poverty Gap Index.

Abstract: Regression analysis is a statistical method used to investigate and model the relationship between variables. Furthermore, a regression analysis was developed that involved spatial aspects, namely Geographically Weighted Regression (GWR). GWR modeling consists of various types, one of which is Geographically Weighted Logistic Regression Semiparametric (GWLRS), an extension of the Logistic GWR model that produces local and global parameter estimators. In this study, it is proposed to combine the GWLRS model using panel data or Geographically Weighted Panel Logistic Regression Semiparametric (GWPLRS). The case study used in this research is the problem of poverty in 38 regions/cities in East Java, Indonesia, in 2018 – 2022 as seen from the Poverty Gap Index. The weights used in this research are the adaptive gaussian kernel weighting functions. The results of the parameter significance test show that the Human Development Index as global variable has a significant effect on each region/city.

1. INTRODUCTION

Poverty is a condition that results in limitations and occurs not at the will of the concerned person. A population is said to be poor if it is marked by a low level of education, work productivity, income, health, nutrition, and well-being, showing a circle of powerlessness. In Indonesia, poverty remains a problem that must be resolved. This is similar to the poverty levels in East Java. East Java is one of Indonesians province. East Java is located between 7.12’ South Latitude - 8.48’ South Latitude and between 111.0’ East Longitude - 114.4’ East Longitude. By the end of 2021, East Java Province was divided into 29 regencies and nine cities (BPS, 2022a). The poverty rate can be seen in the Poverty Gap Index, which is the average measurement of the expenditure gap of each poor population against the Poverty Line. The poverty line reflects the rupiah value of the minimum expenditure needed by a person to meet both food and non-food basic needs for a month.

Regression analysis is a statistical method used to investigate and model the relationship between variables. There are various kinds of regression analysis, starting from the simplest, namely, simple linear regression analysis, multiple regression, logistic regression, Poisson regression, robust regression, and others. The regression analysis that has been mentioned before does not involve the effect of the location of the observations.

Therefore, a regression analysis was developed which also involved spatial or locational aspects, namely Geographically Weighted Regression (GWR).

GWR modeling consists of various types, including Geographically Weighted Poisson Regression (GWPR), Geographically Weighted Negative Binomial Regression (GWNBR), Geographically Weighted Logistic Regression (GWLR), Geographically Weighted Logistic Semiparametric Regression (GWLRS), and others. The use of various types of GWR modeling was adjusted based on the characteristics of the data used. If the response variable used is qualitative or categorical data, then GWLR or GWLRS modeling can be used. GWLR is a combination of the GWR and logistic regression (Atkinson et al., 2003). GWLRS is an extension of the GWLR model, which produces local and global parameter estimators (Nakaya et al., 2009).

GWR modeling has been widely used in various fields. Maggri & Ispriyanti (2013) applied GWR modelling to poverty data. Other researchers have combined GWR modeling and kriging interpolation to predict stunting risk factors by considering the proximity of the area (Pramoedyo et al., 2020). Taufiq et al. (2019) modeled the GWR in the Cox Survival Analysis for data with a Weibull distribution using a Bayesian approach.

Statistical modeling was also performed based on the type of data. The types of data obtained in the field are diverse, including time-series data, cross-section data, and panel data. Panel data are known to have many advantages, so some researchers combine it with other analyses, such as GWR with panel data regression. The advantages of panel data include being able to control the heterogeneity of individual observations, the available data is more informative, can reduce collinearity between variables, and is more efficient (Baltagi, 2005). In this study, a new method is proposed that combines the GWLRS model using panel data or Geographically Weighted Panel Logistic Regression Semiparametric (GWPLRS). The contribution of this research is expected to be able to produce models for data analysis that have response variables in the form of categorical data, estimate local and global parameters, and obtain information from spatial and temporal aspects. The case study used in this research is the problem of poverty in 38 regencies/cities in East Java Province, Indonesia, from 2018 to 2022.

2. LITERATURE REVIEW

2.1. Geographically Weighted Logistic Regression Semiparametric

Geographically Weighted Logistic Regression Semiparametric (GWLRS) is an extension of the GWLR model that produces local and global parameter estimators (Nakaya et al., 2009). In the GWLRS model, the response variable is predicted with a predictor variable, each of which has a regression coefficient $\beta(u_i, v_i)$, depending on the location of the data, and a constant regression coefficient γ_m (Caraka & Yasin, 2017). The GWLRS model is expressed as follows:

$$\pi(x_i) = \frac{\exp\left[\sum_{j=0}^{k^*} \beta_j(u_i, v_i)x_{ij} + \sum_{m=k^*+1}^{k} \gamma_m x_{im}\right]}{1 + \exp\left[\sum_{j=0}^{k^*} \beta_j(u_i, v_i)x_{ij} + \sum_{m=k^*+1}^{k} \gamma_m x_{im}\right]} \tag{1}$$

Information x_{ij} is observed value of the j-th predictor variable at location (u_i, v_i); (u_i, v_i) is geographic coordinates of the i-th observation location; $\beta_j(u_i, v_i)$ is regression coefficient for each location (u_i, v_i); γ_m is constant regression coefficient; x_{im} is the observed value of the m-th predictor variable; k is the number predictor variables; and k^* is the number of local predictor variables.

2.1.1. Multicolinearity Test

To detect multicollinearity, it can be done by calculating the Variance-Inflating Factor (VIF) value (Gujarati, 2004).

$$VIF = \frac{1}{1 - R_j^2} \tag{2}$$

R_j^2 Is the coefficient of determination of the j-th predictor variable with other predictor variables.

2.1.2. Determination of Distance and Bandwidth

In GWPLR model, longitude (v_i) and latitude (u_i) (spatial coordinates) are used to determine the distance between locations (d_{ij}). This can be calculated using the Euclidean distance (Fotheringham et al., 2002).

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{3}$$

To determine the optimum bandwidth value, it can be done by finding the minimum Cross Validation (CV) value (Fotheringham et al., 2002).

$$CV = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq i}(b)\right)^2 \tag{4}$$

y_i is the response variable in the i-th observation, and $\hat{y}_{\neq i}(b)$ is the estimated value of y_i with bandwidth b, but the observation for point i is omitted from the estimation process.

2.1.3. Weighting Function

In this study, the weighting function used is the adaptive weighting function. The adaptive weighting function is a function where bandwidth can adjust its value based on variations in data density. The Adaptive Gaussian Kernel weighting function will decrease the further the distance (d_{ij}) is. Meanwhile, the Adaptive Bisquare Kernel and Adaptive Tricube Kernel weighting functions provide a continuous weighting function up to the distance from the regression point and then give zero weight to any data point outside b_i (bandwidth). In this study, we used an Adaptive Gaussian Kernel weighting function (Fotheringham et al., 2002).

$$w_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b_i}\right)^2\right] \tag{5}$$

2.1.4. Spatial Heterogeneity Test

The data related to a particular spatial or location tends to be non-homogeneous and varies between locations, areas, or other characteristics of the spatial unit. Spatial heterogeneity refers to variations in relationships between spaces. Spatial heterogeneity can be determined by carrying out the Breusch Pagan test (Anselin, 1998).

$$BP = \left(\frac{1}{2}\right)\boldsymbol{f}'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-\mathbf{1}}\boldsymbol{Z}'\boldsymbol{f} + \left(\frac{1}{T}\right)\left[\frac{\boldsymbol{e}'\boldsymbol{W}\boldsymbol{e}}{\sigma^2}\right]^{\mathbf{2}} \tag{6}$$

$\boldsymbol{Z}$ is a standardized observation matrix, $\boldsymbol{W}$ is a weighting matrix, $\boldsymbol{e}$ is an error vector, and $f_i = \frac{e_i^2}{\hat{\sigma}^2} - 1$.

2.2. Geographically Weighted Panel Logistic Regression Semiparametric

This research proposes a combination of the Geographically Weighted Logistic Regression Semiparametric (GWLRS) model using panel data or the Geographically Weighted Panel Logistic Regression Semiparametric (GWPLRS). The difference between the GWPLRS and GWLRS models is that there is a t-index to indicate the year observed.

$$\pi(x_{it}) = \frac{\exp\left[\sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm}\right]}{1 + \exp\left[\sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm}\right]} \tag{7}$$

The logit form of the GWPLRS model is as follows:

$$logit[\pi(x_{it})] = \sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm} \tag{8}$$

Information x_{itj} is observed value of the j-th predictor variable at location (u_{it}, v_{it}) at t-th time; $\beta_j(u_{it}, v_{it})$ is regression coefficient for each location (u_{it}, v_{it}); γ_m is constant regression coefficient; x_{im} is the observed value of the m-th predictor variable at t-th time; k is the number of predictor variables; and k^* is the number of local predictor variables.

2.2.1.Estimation of GWPLRS Parameters

The GWPLRS model parameters were estimated using the Maximum Likelihood Estimation method. The first step was to form a likelihood function.

$$L(\beta(u_{it}, v_{it})\gamma_m) = \left\{\exp \sum_{i=1}^{n} y_{it} \left(\sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})\, x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm}\right)\right\} \left\{\prod_{i=1}^{n} \left(1 + \exp\left(\sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})\, x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm}\right)\right)^{-1}\right\} \tag{9}$$

Next, look for the lnlikelihood form of equation (9).

$$\ln L(\beta(u_{it}, v_{it})\gamma_m) = \sum_{i=1}^{n} y_i \left(\sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})\, x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm}\right) - \sum_{i=1}^{n} \ln\left(1 + \exp\left(\sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})\, x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm}\right)\right) \tag{10}$$

In the GWPLRS model, the weighting function used is the geographic location factor; therefore, to obtain the GWLRS model, it is given a weight on its lnlikelihood function. GWPLRS model weighting is the same as GWR weighting in general.

$$\ln L^*(\beta(u_{it}, v_{it})\gamma_m) = \sum_{i=1}^{n} y_{it} w_{itj}(u_{it}, v_{it}) \left(\sum_{j=0}^{k^*} \beta_j(u_{it}, v_{it})\, x_{itj} + \sum_{m=k^*+1}^{k} \gamma_m x_{itm}\right) - \sum_{i=1}^{n} w_{itj}(u_{it}, v_{it}) \tag{10}$$

$$\ln\left(1+\exp\left(\sum_{j=0}^{k^*}\beta_j(u_{it},v_{it})\,x_{itj}+\sum_{m=k^*+1}^{k}\gamma_m x_{itm}\right)\right)$$

The composition of the weighting matrix for GWPLRS modeling is as follows. The weighting matrix is a matrix of $Nt \times Nt$ size, where the diagonal of the matrix is the geographical weight of each location that is repeated the number of years observed. The weighted matrix elements, other than the main diagonal, were zero.

To obtain an estimate of parameter β that maximizes $L(\beta(u_i,v_i)\gamma_m)$, equation (12) is derived from $\beta_j(u_i,v_i)$, and γ_m then equals zero.

$$\frac{\partial \ln L^*(\beta(u_{it},v_{it}),\gamma_m)}{\partial \beta_j(u_{it},v_{it})}=\sum_{i=1}^{n} w_{itj}(u_{it},v_{it})y_{it}x_{itj} -\sum_{i=1}^{n}\pi(x_{it})x_{itj}w_{itj}(u_{it},v_{it})=0 \tag{13}$$

$$\frac{\partial \ln L^*(\beta(u_{it},v_{it}),\gamma_m)}{\partial \gamma_m}=\sum_{i=1}^{n} w_{itj}(u_{it},v_{it})y_{it}x_{itm} -\sum_{i=1}^{n}\pi(x_{it})x_{itm}w_{itj}(u_{it},v_{it})=0 \tag{14}$$

Because the functions in equations (13) and (14) are in implicit form, Newton–Raphson iterations are performed in the following general form.

$$\begin{aligned}&\left(\boldsymbol{\beta}^{(t)}(u_i,v_i),\boldsymbol{\gamma}^{(m)}\right)\\ &\left(\boldsymbol{\beta}^{(t+1)}(u_{it},v_{it}),\boldsymbol{\gamma}^{(m+1)}\right)\\ &=-\left(\boldsymbol{H}^{(t)^{-1}}\left(\boldsymbol{\beta}^{(t)}(u_{it},v_{it}),\boldsymbol{\gamma}^{(m)}\right)\right)\boldsymbol{g}^{(t)}\left(\left(\boldsymbol{\beta}^{(t)}(u_{it},v_{it}),\boldsymbol{\gamma}^{(m)}\right)\right)\end{aligned} \tag{15}$$

The initial stage of iteration is to determine the initial values of $\hat{\beta}$ and $\hat{\gamma}$. For example, determine the value of $\hat{\beta}^{(0)}=0$ and $\hat{\gamma}^{(0)}=0$. The iteration stops when a convergent state is reached. Defined $\boldsymbol{\alpha}'=[\beta_0 \quad \beta_1 \quad \cdots \quad \beta_k \quad \gamma_1 \quad \gamma_2 \quad \cdots \quad \gamma_k]$. So, the convergent state is when $\left\|\boldsymbol{\alpha}^{(t+1)}-\boldsymbol{\alpha}^{(t)}\right\|\le\varepsilon$, where ε is a positive number with very small value, for example 10^{-12}. In addition, if it does not meet the convergent conditions, the maximum iteration is determined, which is 100 iterations.

2.2.2.GWPLRS Model Parameter Testing

GWPLRS model parameter testing is performed by partially testing the parameters. This test was conducted to determine the local and global predictor variable parameters that significantly influenced the response variable. To determine whether the parameters of predictor variables that are local in nature have a significant effect, hypothesis testing was carried out as follows.

$$H_0 : \beta_j(u_{it},v_{it})=0$$

$$H_1 : \beta_j(u_{it},v_{it})\neq 0$$

Test statistic :

$$Z=\frac{\hat{\beta}_j(u_{it},v_{it})}{Se\left(\hat{\beta}_j(u_{it},v_{it})\right)} \tag{16}$$

H_0 is rejected if $|Z_{score}| > Z_{\alpha/2}$.

To find out the parameters of global predictor variables have a significant effect, hypothesis testing is carried out as follows

$$H_0 : \gamma_m = 0$$

$$H_1 : \gamma_m \neq 0$$

Test statistic:

$$Z = \frac{\hat{\gamma}_m}{Se(\hat{\gamma}_m)} \tag{17}$$

H_0 is rejected if $|Z_{score}| > Z_{\alpha/2}$.

2.3. Poverty Problem

Poverty Gap Index (P1). The Poverty Gap Index (P1) is a measure of the average expenditure gap of each poor person against the poverty line. The higher the index value, the farther the average population expenditure from the poverty line (BPS, 2023).

$$P_\alpha = \frac{1}{n}\sum\nolimits_{i=1}^{q} \left[\frac{z - y_i}{z}\right]^\alpha \tag{18}$$

where α = 1, z is poverty line, y_i is the average expenditure per capita per month for the population who is below the poverty line, q is the number of people who are below the poverty line, and n is the total population.

Human Development Index (HDI). Human development is the expansion of people's choices to live freely and with dignity and the expansion of capabilities to fulfill aspirations. Human development is used as a way to measure human quality in an area. IPM explains how residents can access development results in obtaining income, health, education, and other aspects of life. HDI achievements between regions can be seen by grouping HDI into several categories, namely low (HDI<60), moderate (60≤HDI<70), high (70≤HDI<80), and very high (HDI≥80) (BPS, 2022b). Based on previous research, HDI has a significant effect on the Poverty Gap Index (Agustin et al., 2019).

Open Unemployment Rate (OUR). The Open Unemployment Rate is the percentage of the number of unemployed to the total labor force. The population included in the labor force is the population of working age, namely those aged 15 years and over who are working, or have a job but are temporarily not working and are unemployed. The unit of the Open Unemployment Rate is percent. Based on previous research, the Open Unemployment Rate has a significant effect on the Poverty Depth Index (Apriliani, 2023).

Minimum Wage. The minimum wage is a minimum standard used by workers or industry players to provide wages to their workers. The unit of the minimum wage is rupiah. Based on previous research, the minimum wage has a negative effect on the Poverty Depth Index (Wiriarsa, 2015).

3. MATERIAL AND METHOD

3.1. Data Sources

The data used is secondary data taken from the website of the Central Bureau of Statistics (https://bps.go.id/). The research involved 38 regencies/cities in East Java from 2018 to 2022. Table 1 shows a list of regencies/cities used in this study.

Table 1. List of Regencies/Cities in East Java

ID	Regency/City	ID	Regency/City	ID	Regency/City	ID	Regency/City
1	Pacitan	11	Trenggalek	20	Gresik	30	Probolinggo
2	Ponorogo	12	Tulungagung	21	Surabaya City	31	Jember
3	Magetan	13	Blitar City	22	Sidoarjo	32	Banyuwangi
4	Madiun City	14	Blitar	23	Pasuruan City	33	Bondowoso
5	Madiun	15	Kediri City	24	Pasuruan	34	Situbondo
6	Ngawi	16	Kediri	25	Batu City	35	Bangkalan
7	Bojonegoro	17	Jombang	26	Malang City	36	Sampang
8	Tuban	18	Mojokerto City	27	Malang	37	Pamekasan
9	Lamongan	19	Mojokerto	28	Lumajang	38	Sumenep
10	Nganjuk			29	Probolinggo City		

3.2. Research Variables

The variables used in this study are Poverty Gap Index (Y), Human Development Index (X_1), Open Unemployment Rate (X_2), and Minimum Wage (X_3).

3.3. Data Analysis Method

The following are the steps taken in this research:

1. Perform multicollinearity testing by looking at the VIF according to equation (2).
2. Calculate the Euclidean distance using equation (3).
3. Calculating the CV to determine the optimum bandwidth using equation (4).
4. Determine the weighting matrix using equations (5).
5. Perform a BP test to test the assumption of spatial heterogeneity using equation (6).
6. Estimating GWPLRS model parameters.
7. Test the parameters of the GWPLRS model.

4. RESULTS AND DISCUSSION

The difference between this research and previous research lies in the method used. The method used in this study was GWPLRS, which is a combination of GWLRS using panel data. This study uses secondary data, with the Poverty Gap Index (P1) as the response variable, the Human Development Index (IPM), the Open Unemployment Rate (OUR), and the Minimum Wage as predictor variables. The response variable used was categorical data, where the Poverty Gap Index (P1) was classified into 0 and 1. If P1 < average P1, then coded as 0 (low), and if P1 ≥ average P1, then coded as 1 (high) (Hendayanti & Nurhidayati, 2020). The data used is in the form of panel data, where 38 regencies/cities in East Java are cross-section units and 2018 – 2022 are time series units.

4.1. Multicollinearity Testing

Multicollinearity testing is used to see whether there is a linear relationship between the independent variables. The VIF value for each independent variable is <10, so it can be concluded that there is no multicollinearity between the independent variables.

Table 2. VIF value of each Predictor Variable

Variable	X_1	X_2	X_3
VIF	1.4604	1.8670	1.5008

4.2. Euclidean Distance Calculation

In this study, the location of each province is expressed in terms of latitude and longitude, which are then projected into the coordinates of location points using the Universal Transverse Mercator (UTM) coordinate system. The projection produces easting and northing coordinates in metres (Fitriani & Efendi, 2019). Distance determination was performed using the Euclidean distance. This distance is required to determine the weight matrix for GWPLRS modeling. The Euclidean distances for several regions/cities in East Java were show in Table 3.

Table 3. Euclidean Distance for Each Regency/City

Regency/City	Pacitan	⋯	Sumenep
Pacitan	0	⋯	378668.20
Ponorogo	41354.68	⋯	338292.10
Magetan	54886.23	⋯	345433.50
Kota Madiun	67193.57	⋯	326265.40
Madiun	75674.47	⋯	313536.90
Ngawi	78069.80	⋯	342475.40
⋮	⋮	⋮	⋮
Sumenep	378668.20	⋯	0

Based on Table 3, the Euclidean distance for each region/city was obtained in metres. The farther the location between the regions, the greater the Euclidean distance. For example, the distance between Ponorogo and Pacitan was smaller than that between Sumenep and Pacitan. Meanwhile, the distance between the region/city and itself is zero.

4.3. Geographically Weighted Panel Logistic Regression Semiparametric (GWPLRS)

In this sub-chapter, Geographically Weighted Panel Logistic Regression Semiparametric (GWPLRS) modeling is carried out using the same weighting function as GWPLR modeling in the previous chapter, namely the Adaptive Gaussian Kernel. The GWPLRS model included local and global variables. If the variables are local, they have different regression coefficients at each location. When the variable is global, it means that the regression coefficient for this variable has the same value for all locations. Data processing was carried out using GWR 4.0 software, which uses the global-to-local variable selection option for selecting variables to become local and global variables (Nakaya et al., 2009).

4.3.1.Determination of the GWPLRS Optimum Bandwidth

Bandwidth is a circle with a radius from a central location point that is used to determine the weight of each observation. In this study, an Adaptive Gaussian Kernel weighting function is used. Table 4 shows the optimum bandwidth value of the GWPLRS model for each region/city in East Java, using the Adaptive Gaussian Kernel weighting function.

4.3.2.Determination of Weighting Matrix

In this study, we used an Adaptive Gaussian Kernel weighting function. This weighting function was chosen because, even though one location is very far from another, it is still possible to obtain a weight even though the value is very small. This weighting function was obtained after searching for the optimum distance and bandwidth values. In this study, the location of each regency/city is expressed in terms of latitude and longitude.

Latitude and longitude are projected into coordinates of location points using the Universal Transverse Mercator (UTM) coordinate system. The projection produces the easting and northing coordinates in meter. The following presents weights from one of the regencies/cities, namely, the Madiun Regency. Table 5 shows the weighted value for each region/city using the Adaptive Gaussian Kernel weighting function for the Madiun Regency.

Table 4. Optimum Bandwidth for Each Regency/City

Regency/City	Optimum Bandwidth
Pacitan	24408.76577
Ponorogo	24392.07047
Magetan	24190.78799
Kota Madiun	24137.18683
Madiun	24066.31982
⋮	⋮
Sumenep	24581.88953

Table 5. Weighting Value for Madiun Regency

Regency/City	Optimum Bandwidth
Pacitan	0.0082
Ponorogo	0.3050
Magetan	0.4161
Kota Madiun	0.8655
Madiun	1.0000
Ngawi	0.2441
⋮	⋮
Sumenep	4.71×10^{-36}

Based on Table 5, it can be seen that the Adaptive Gaussian Kernel weighting function has a different value for each region/city. In this study, we used this weighting function because even though one location is very far from another, it is still possible to obtain a weight even though the value is very small. It is different from the other two weighting functions, namely, the Adaptive Bisquare Kernel and Adaptive Tricube Kernel, which give zero weighting if the distance is greater than the bandwidth value. In other words, locations with great distances are considered to have no effect on point i.

The weighted value of a location was equal to one. For example, the Madiun Regency; the weight for the Madiun Regency itself is one. Regions located closer to the Madiun Regency receive greater weighting than areas farther away. For example, the weighting for Madiun City, which is adjacent to the Madiun Regency, has a higher value than the weighting for Bondowoso Regency for Madiun Regency, which is located further away.

The weighting matrix for each location was in the form of a diagonal matrix of the weights obtained, as shown in Table 4. The weighting value is repeated for each year. In this study, five years of observation were used (2018 to 2022), to form a diagonal matrix measuring 190×190. In addition to the main diagonal, the elements of matrix are zero. The following is the Adaptive Gaussian Kernel weighting matrix for Madiun Regency.

$$\boldsymbol{W}(u_5, v_5) = diag[0.0082 \quad \cdots \quad 4.7 \times 10^{-36}] \tag{19}$$

4.3.3.Spatial Heterogeneity Test

In the GWPLRS analysis, it is necessary to test spatial heterogeneity using the Breusch–Pagan test with test statistics, as shown in equation (9). This test is carried out after obtaining the weighting matrix. Based on the test results, the statistical value of the BP test was 196.5793 and the p-value of the BP test was $2.04 \times 10^{-41} < 0.05$. It can be concluded that at a significance level of 5%, the decision is to reject H_0, and it can be said that there is spatial heterogeneity, meaning that there is variation in the relationship between locations. Because there is spatial heterogeneity based on the results of the BP test, GWPLRS modeling can be continued.

4.3.4.Estimation and Significance of GWPLRS Model Parameters

The results of the parameter estimation of the GWPLRS model vary at each observation location. Based on the results of the analysis, the variable X_1 or HDI was obtained as a global variable. This means that the coefficient value for the HDI is the same for all observation locations. The variables X_2 (OUR) and X_3 (Minimum Wage) are local variables with different regression coefficients in each location. The following are the results of the parameter estimation of the GWPLRS model for each region/city, using the Adaptive Gaussian Kernel weighting function. In addition, a parameter significance test for each of the variables studied was also carried out. In this discussion, only one of the 38 regencies/cities in East Java, the Madiun Regency, is discussed. Table 6 presents the results of the estimation and odds ratio of the parameters of the GWPLRS model for Madiun Regency.

Table 6. Parameter Estimation Results and the Odds Ratio Value

	Parameter			
	$\hat{\beta}_0$	$\hat{\gamma}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
Estimation	35.11673	−0.5217	0.264619	0,0000
	(0.0318)*	(0.0295)*	(0.3568)	(0.5372)
Odds Ratio		0.5935	1.3029	1.0000

*Significant at $\alpha = 5\%$

The following is the GWPLRS model formed with the Adaptive Gaussian Kernel weighting function for Madiun Regency.

$$\hat{\pi}(x_{itj}) = \frac{\exp(35.1167 - 0.5217X_{5t1} + 0.2600X_{5t2} + 0{,}0000X_{5t3})}{1 + \exp(35.1167 - 0.5217X_{5t1} + 0.2600X_{5t2} + 0{,}0000X_{5t3})} \tag{20}$$

The logit form of equation (20) is as follows.

$$logit\ \hat{\pi}(x_{itj}) = 35.1167 - 0.5217X_{5t1} + 0.2600X_{5t2} + 0{,}0000X_{5t3} \tag{21}$$

The odds ratio for the HDI estimator was 0.5935<1, indicating that there was a negative relationship between HDI and P1. Each addition of one HDI unit caused a 52.17% decrease in P1. The odds ratio for the OUR estimator is 1.3029> 1, indicating a positive relationship between OUR and P1. The minimum wage variable estimator has an odds ratio value close to 1, meaning that there is no relationship between minimum wage and P1. Based on the parameter significance test, HDI has a significant effect on P1 in all regions, while OUR and minimum wages have no significant effect on P1 in any region.

Table 7 presents the classification results from GWPLRS modeling. The classification accuracy was calculated using the 1-APER formula. Based on the calculation results, the classification accuracy was 76.32%.

Table 7. Classification Table

Actual Value	Prediction	
	P1 = 0	P1 = 1
P1 = 0	97	21
P1 = 1	27	48
Accuracy	76,32%	

5. CONCLUSION

The model formed is GWPLRS using Adaptive Gaussian Kernel weighting function. GWPLRS is a combination of GWPLR analysis using panel data. Each region/city in East Java has a different model, according to the regression coefficient that has been obtained. In the GWPLRS model there are local and global variables. Based on the results of the analysis, the variable X_1 or HDI was obtained as a global variable. This means that the coefficient value for the HDI is the same for all observation locations. The variables X_2 (OUR) and X_3 (Minimum Wage) are local variables with different regression coefficients in each location. HDI has a significant and negative influence on the Poverty Gap Index (P1) in each region or city. The classification accuracy of the GWPLRS model formed was 76.32%

REFERENCES

Agustin, N., Hidayat, M. S., & Umiyati, E. (2019). Analisis Pengaruh Pertumbuhan Ekonomi, Indeks Pembangunan Manusia (IPM) dan Upah Minimum Provinsi (UMP) terhadap kemiskinan di Kabupaten Merangin. *E-Jurnal Perspektif Ekonomi dan Pembangunan Daerah*, *8*(1), 53–64. https://doi.org/10.22437/pdpd.v8i1.7346

Anselin, L. (1998). Spatial Econometrics Methods and Models *Studies in Operational Regional Science* (SORS) volume 4. New York: Springer.

Apriliani, A. R., Safrida, & Fajri (2023). Pengaruh Pengangguran terhadap Indeks Kedalaman Kemiskinan di Provinsi Aceh. *Journal Ilmiah Mahasiswa Pertanian*, *8*(1), 73–81.

Atkinson, P. M., German, S. E., Sear, D. A., & Clark, M. J. (2003). Exploring The Relations between Riverbank Erosion and Geomorphological Controls using Geographically Weighted Logistic Regression. *Geographical Analysis*, *35*(1), 58–82. https://doi.org/10.1111/j.1538-4632.2003.tb01101.x

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data* (Issue 1). John Wiley & Sons, Ltd.

BPS (2022a). *Provinsi Jawa Timur dalam Angka 2022*. Badan Pusat Statistik.

BPS (2022b). *Statistik Indonesia Statistical Yearbook of Indonesia 2022*. Badan Pusat Statistik. https://www.bps.go.id/publication/2020/04/29/e9011b3155d45d70823c141f/statistik-indonesia-2020.html

BPS (2023). *Poverty and Inequality*. https://www.bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html

Fitriani, R. & Efendi, A. (2019). *Ekonometrika Spasial Terapan dengan R*. Malang: UB

Press.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression the Analysis of Spatially Varying Relationships*. John Wiley & Sons, Ltd.

Gujarati, D. N. (2004). *Basic Econometrics* (4th ed.). McGraw-Hill/Irwin. https://doi.org/10.1596/1813-9450-8096

Hendayanti, N. P. N. & Nurhidayati, M. (2020). Regresi Logistik Biner dalam Penentuan Ketepatan Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi-provinsi di Indonesia. *Sainstek : Jurnal Sains dan Teknologi*, *12*(2), 63. https://doi.org/10.31958/js.v12i2.2483

Maggri, I. & Ispriyanti, D. (2013). *Pemodelan Data Kemiskinan Di Provinsi Sumatera Barat Dengan Metode Geographically Weighted Regression (GWR)*. 37–49.

Nakaya, T., Fotheringham, A. S., Charlton, M., & Brunsdon, C. (2009). *Semiparametric Geographically Weighted Generalised Linear Modelling in GWR 4.0*.

Pramoedyo, H., Mudjiono, M., Fernandes, A. A., Ardianti, D., & Septiani, K. (2020). Determination of Stunting Risk Factors Using Spatial Interpolation Geographically Weighted Regression Kriging in Malang. *Mutiara Medika: Jurnal Kedokteran dan Kesehatan*, *20*(2), 98–103. https://doi.org/10.18196/mm.200250

Taufiq, A., Astuti, A. B., & Fernandes, A. A. R. (2019). Geographically Weighted Regression in Cox Survival Analysis for Weibull Distributed Data with Bayesian Approach. *IOP Conference Series: Materials Science and Engineering*, *546*(5). https://doi.org/10.1088/1757-899X/546/5/052078

Wiriarsa, M. (2015). Minimum Wage and Poverty : Empirical Study on District Panel Data in Java, Indonesia. *Thesis*. International Institute of Social Studies.