A Spatial Analysis of Crime in East Java Province in 2019

C Ummah, Rini Rahani
STIS Polytechnic of Statistics, Jakarta, Indonesia

*Corresponding author's e-mail: 211709601@stis.ac.id

Abstract. Crime is one of the consequences of fluctuations in the economic condition of a country. Crime incidents harm many parties. The number of criminal acts increased in 2019, especially in Sumatra and Java Island. Most provinces experienced an increasing number of criminal acts, one of them was East Java. East Java contributed more than a quarter of the number of crimes throughout Java Island. The number of criminal acts is count data with overdispersion because its variance is higher than its average. This study aims to analyze the number of criminal acts by applying Geographically Weighted Negative Binomial Regression (GWNBR). The result shows that GWNBR formed two regional groups based on significant variables. The four independent variables namely the unemployment rate, the number of poor people, the Gini ratio, and the police population ratio have a significant effect on all districts/cities. However, the mean year of schooling shows a significant effect only in certain districts/cities. The GWNBR is the best model in modelling the number of criminal acts in East Java.

1. Introductions

Humans have a series of basic needs that must be fulfilled, one of them is the need for security. Security is one of the human rights guaranteed by law. Due to this reason, the government is responsible for ensuring community security and safety and controlling society. To achieve this, the government has launched a series of targets and policies in the 2015-2019 RPJMN, namely realizing a safe, peaceful, and united Indonesia. One of them is to protect and protect the community, prevent crime, and solve the crime.

Crime is one of the consequences of fluctuations in the economic condition of a country. The state is estimated to be capable to increase the population's per capita income growth by around 1 percent per year if it can reduce the crime rate to 10 percent [1]. Meanwhile, the ability of each individual to meet the needs of life is at different levels. It potentially creates society gaps, one of which is that individuals with insufficient income to meet their daily needs will take other ways to earn additional income, including using illegal methods.

The Police Headquarters Operations Control Bureau noted that there were 269,324 criminal acts in 2019. The crimes occurred in Java and Sumatra, ranging from 2.8 percent to 11.9 percent. While the rest spread to other regions ranging from 0.3 percent to 2.8 percent. One of the areas experiencing an increase in the number of criminal acts is East Java Province. In 2019, 26,985 crimes occurred in the province. This number shows an increase of 690 cases compared to the previous year, even contributing 29.32 percent of criminal acts on the island of Java. This increase in the number of criminal acts causes the risk of people becoming victims of crime to increase and the time interval for the occurrence of criminal acts to increase rapidly.

Based on data from the Central Statistics Agency for East Java Province in 2019, the number of poor people is 4,112.25 thousand people, a decrease of 0.61 percent compared to the previous year. Not only

that, but the Open Unemployment Rate (TPT) in East Java also decreased by 0.08 points to 3.92 percent [2]. These two indicators are a signal of an improvement in the socio-economic conditions of the community. The number of criminal acts in the same year increased. This phenomenon shows that some factors have influenced the number of criminal acts in East Java.

A study [3] researched criminal acts in East Java Province in 2015 using the GWNBR method. The results of the study divided 38 districts/cities in East Java into three groups. The first group consists of 14 districts/cities where the variables TPT, Gini ratio, and Mean Year Schooling (MYS) significantly affect the crime numbers. Furthermore, the second group that consists of seven regencies/cities, the TPT variable, the percentage of the poor population, the Gini ratio, and the RLS have a significant effect on the crime's number. Meanwhile, in the third group (17 other districts/cities) only the open unemployment rate and the Gini ratio were significant in the model. However, the research [3] does not consider the existence of multicollinearity. The violation of non-multicollinearity causes the estimated regression coefficient to be inaccurate and it produces a large standard error. [4] A similar study, focusing on property crimes also conducted using two methods, namely GWPR and GWNBR. There are eight regional groups formed from the GWPR model and two regional groups from the GWNBR model based on the significant independent variables. In terms of performance, the GWNBR model (AIC = 369.21) is better than the GWPR model (AIC = 614.37). However, this research only focuses on property crimes so that it is not able to capture other criminal phenomena. In addition, both previous studies only involved factors that motivated the perpetrators. Whereas other factors can lead to crime, namely the absence of an effective guardian as described in Routine Activity Theory [5].

The data on the crime that occurred is the count data or count data. The model that becomes the benchmark for the count data is the Poisson model [6]. The use of the Poisson regression model allows for overdispersion, which is a condition where the variance value is greater than the average. One technique to overcome the overdispersion state is the negative binomial model. Negative binomial modelling will produce a global regression, even though criminal acts between regions can vary due to different socio-demographic characteristics. Global statistics are usually single values. Meanwhile, local statistics have multiple values where the statistical values can differ between regions [7]. Therefore, this study aims to describe the crimes in East Java Province, find out the suitable spatial model for modelling crimes in East Java Province, and find out the variables that influence it.

2. Theoretical Review

2.1 Crime

Crime is synonymous with unlawful acts that harm the victim. In addition to having an impact on victims, crime also affects social life. A high crime rate indicates an unsafe area that can cause unrest. Therefore, it is necessary to monitor the development of the crime rate from time to time as the basis for formulating well-targeted policies.

The crime rate resulting from the recording of criminal acts describes the condition of security and public order (Kamtibmas) and the level of vulnerability of an area.The crime can be classified based on some characteristics such as the target of the crime, the seriousness of the crime committed and how the crime is done [8]. Based on these criteria, crimes are grouped into crimes against life, crimes against the physical, crimes against decency, crimes against the liberty of people, crimes against property rights/goods with violence, crimes against property rights/goods, crimes related to narcotics, crimes related to fraud, embezzlement, and corruption, and crimes against public order.

The crime that occurs in a country is an accumulation of criminal acts committed by a person. A person has the opportunity to commit a crime as a result of failure to meet the needs of life with income earned from daily work. For that, someone will seek additional income through criminal activity. In addition, the existence of inequality creates opportunities for criminal acts to occur. [9] argues that inequality leads to crime by placing low-income individuals near high-income individuals who have things worth taking away.

2.2 Poisson Regression

Poisson regression is the standard model that becomes the reference for modelling count data [6]. The Poisson regression model uses a log link function, namely $\boldsymbol{ln}\ (\boldsymbol{\mu_i})\ = \boldsymbol{\eta_i}$. The relationship between the independent variable and the dependent variable in Poisson regression is defined as follows:

$$ln\ (\mu_i) = \beta_0 + \beta_i x_i \tag{1}$$

with β_i as regression's coefficient.

2.3 Overdispersion

Overdispersion is a condition when the resulting variance value is greater than the average value in the Poisson regression model. The overdispersion can be identified through the resulting dispersion parameter, namely *theta* (θ). A theta value equal to 1 indicates an equidispersion, a theta value less than one indicates an underdispersion, and a theta value greater than one indicates an overdispersion. The dispersion can be tested by dividing the deviance and Pearson Chi-Square values by the degrees of freedom.

2.4 Negative Binomial Regression

Negative binomial regression is one of the regression models that can be used to model count data. The negative binomial regression model uses a dispersion parameter so that the variance value is greater than the average value. The negative binomial model is a combination of the Gamma distribution and the Poisson distribution, so we need an $\boldsymbol{\alpha}$ dispersion parameter to obtain a negative binomial regression model. With a dispersion parameter value used, this model can overcome the symptoms of overdispersion that occur in Poisson regression. The relationship between the independent variable and the dependent variable in negative binomial regression is defined as follows:

$$y_i = \mu_i + \varepsilon_i = e^{x_i'\beta} + \varepsilon_i \tag{2}$$

with ε_i as *error* on the *i*-th observation and β as the value of regression's coefficient.

2.5 Spatial Autocorrelation

Spatial dependencies reflect the observation value in an area depending on the observation value of the nearest neighbouring region. The relationship between these observation values can be calculated by spatial autocorrelation. The spatial autocorrelation can be tested using the Moran Index. Moran index is the oldest index used to identify spatial autocorrelation formulated as follows [10]:

$$I = \frac{n\sum_i\sum_j \boldsymbol{W}_{ij}(x_i - \bar{x})(x_j - \bar{x})}{W\sum(x_i - \bar{x})^2} \tag{3}$$

with
n : the number of observations
x_i: the observation value of the i-th region
x_j: the observation value of the $\boldsymbol{j}$-th region
$\bar{x}$: the average value of $\boldsymbol{n}$ regions
$\boldsymbol{W_{ij}}$: weighting matrix of the $\boldsymbol{i}$-th and $\boldsymbol{j}$-th region
$\boldsymbol{W}$: the sum of all values in the weighting matrix

2.6 Spatial Heterogeneity

Spatial heterogeneity is a common problem related to changes in spatial structure caused by differences/instability of independent variables in each unit of spatial analysis. Spatial heterogeneity is also associated with the heteroscedasticity that occurs due to missing variables or other forms of misspecification. It causes errors with non-constant variances [11]. This Research applied the Breusch-Pagan test for testing the spatial heterogeneity with the following formula:

$$BP = \frac{1}{2} f^T Z(Z^T Z)^{-1} Z^T f \sim \chi^2_{(k-1)} \tag{4}$$

with $\boldsymbol{f} = (\boldsymbol{f_1}, \boldsymbol{f_2}, \dots, \boldsymbol{f_n})^{\boldsymbol{T}}$ obtained from $\boldsymbol{f_i} = \left(\frac{e_i^2}{\sigma^2} - \mathbf{1}\right)$ and $\boldsymbol{e_i} = \boldsymbol{y_i} - \hat{\boldsymbol{y}}_{\boldsymbol{i}}$
where
$\boldsymbol{e_i^2}$: the squared error for $\boldsymbol{i}$-th observation
$\boldsymbol{Z}$: $\boldsymbol{n} \times \boldsymbol{p}$ sized matrix containing standardized vector for each observation's region
$\hat{\sigma}^{\boldsymbol{2}}$: the dependent variable's variance

2.7 Optimum Bandwidth
The Optimum bandwidth selection affects the resulting parameter coefficients. Bandwidth is a "smoothing parameter" because the higher the bandwidth value, the greater the smoothing. Oversmoothed models will produce homogeneous parameter values, while under smoothed models tend to yield heterogeneous parameter values [7]. Three ways to get the optimum bandwidth value are user-supplied bandwidth, minimizing the CV value, or minimizing the AIC value.

2.8 Geographically Weighted Negative Binomial Regression (GWNBR)
Geographically Weighted Negative Binomial Regression (GWNBR) is a statistical method used to model overdispersion count data and contain spatial heterogeneity. The GWNBR model will produce local parameter estimates so that each observation area has a different parameter value. The formula for the GWNBR model is as follows [12]:

$$E[y_i] = \hat{\mu}_i = exp\left\{\beta_0(u_i, v_i) + \sum_{i=1}^{p} \beta_k(u_i, v_i) x_{ik} + \theta(u_i, v_i)\right\} \quad (5)$$

where:
y_i : the value of i-th observation
x_{ik}: the value of k-th independent variable observation at the observed location (u_i, v_i)
(u_i, v_i): the point location of i-th observation
$\beta_k(u_i, v_i)$: regression coefficient of the k-th independent variable for each location (u_i, v_i)
$\theta(u_i, v_i)$: dispersion parameter for each location $(\boldsymbol{u_i}, \boldsymbol{v_i})$

3. Methodology
This study covers 38 districts/cities in East Java Province in 2019. The data used in this study is secondary data sourced from various publications compiled by the Central Statistics Agency of East Java Province. Based on the literature review and the data sources used, the variables used in this study are as follows:

Table 1. The variable and data source

Code	Variable	Data Source
Y	Crime numbers	Statistik Politik dan Keamanan Provinsi Jawa Timur 2019
X_1	The open unemployment rate	Laporan Eksekutif Keadaan Angkatan Kerja Provinsi Jawa Timur 2019
X_2	The number of poor people (in thousands)	Provinsi Jawa Timur Dalam Angka Tahun 2020
X_3	Mean year schooling	Statistik Pendidikan Provinsi Jawa Timur 2019
X_4	Gini ratio	Website BPS Provinsi Jawa Timur
X_5	Population to police ratio	Statistik Politik dan Keamanan Provinsi Jawa Timur 2019

The descriptive analysis is presented using graphs and tables to explore statistics and thematic maps to describe the distribution of crime and the variables that influence it. The inferential analysis is done by forming a spatial model that can explain the influence of the independent variables in this study. The steps taken are as follows:

1. Describing the suspected variables that influencing the number of criminal acts in East Java Province in 2019.
2. Testing for multicollinearity.
3. Establishing a Poisson regression model.
4. Performing dispersion testing.
5. Establishing a negative binomial regression model.
6. Identifying spatial effects in the form of spatial autocorrelation and spatial heterogeneity.
7. Determining the optimum bandwidth and form a spatial weighting matrix.
8. Forming the GWNBR model.
9. Testing the significance of the parameters simultaneously.
10. Testing the significance of the parameters partially.
11. Grouping districts/cities according to significant variables.
12. Evaluating the model results.

4. Results and discussion

4.1 The Crimes Overview in East Java

The number of crimes (total crimes) by district/city in East Java Province is presented in Figure 1. Based on Figure 1, information is obtained that the three areas with the highest number of crimes are Surabaya City (3,980 incidents), Malang Regency (3,337 incidents), and Jember Regency (1,721 incidents). Meanwhile, the three areas with the lowest number of crimes are Pacitan Regency with 93 incidents, followed by Madiun City with 162 incidents, and Mojokerto City with 202 incidents.

The figure shows that most of the crimes occurred in the central and eastern regions, indicated by the red color which tends to be dense, indicating areas with high to very high crime incidence. This is because the central region is an area with high economic activity and many factories such as Surabaya, Gresik, Sidoarjo, Kertosono, Jombang, and Malang. The eastern regions such as Jember, Banyuwangi, Pasuruan, Probolinggo, Situbondo, and Bondowoso land distribution routes connecting Java Island to Bali and other parts of Central Indonesia. Hence, those areas are at high risk of criminal acts.

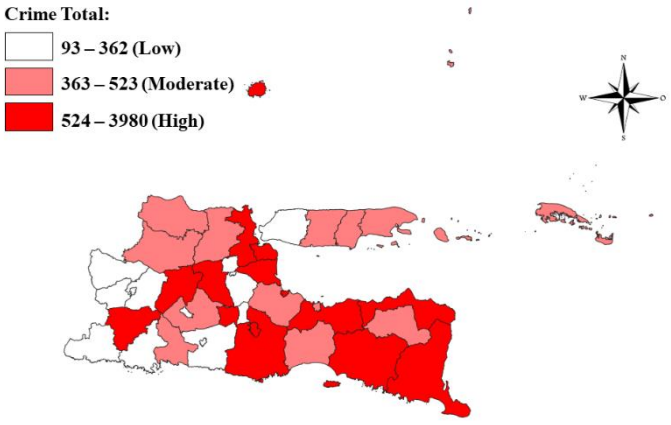


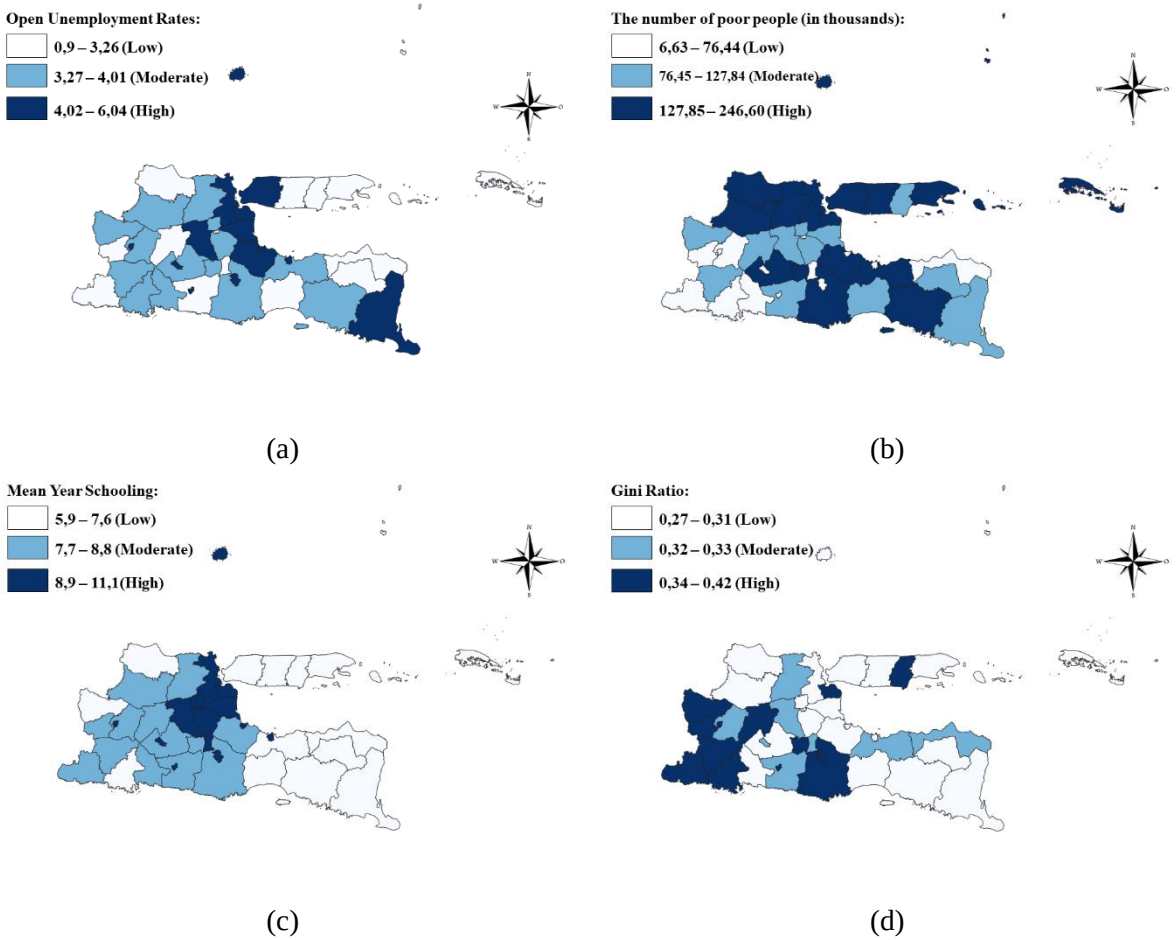
Figure 1. Thematic map of the distribution of crime in East Java Province in 2019

The thematic map of the variables that affect the number of criminal acts in East Java Province is presented in Figure 2. The distribution of the variable open unemployment rate can be observed in Figure 2a. The solid dark blue represents the area with high value of the open unemployment rates. This area covers districts/cities with high economic activity and industrial centres such as Gresik, Surabaya, Sidoarjo, and Pasuruan. The three regions with the highest open unemployment rates are Malang City (6.04 percent), Surabaya City (5.87 percent), and Bangkalan Regency (5.84 percent). Meanwhile, the

three regions with the lowest open unemployment rates are Pacitan Regency (0.95 percent), Sumenep Regency (2.17 percent), and Pamekasan Regency (2.32 percent).

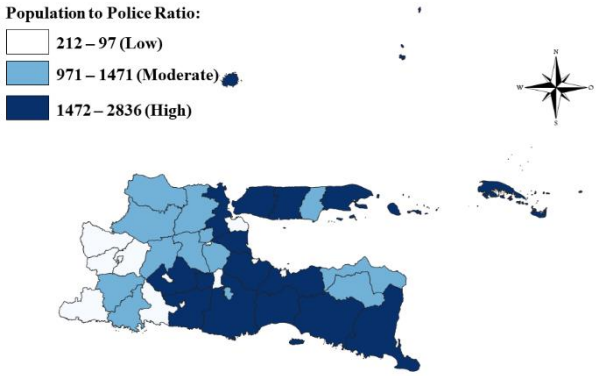
Figure 2b shows the distribution of the number of poor people in East Java Province. Areas with a high number of poor people spread across the north, south, and Madura Island. The three regions with the highest number of poor people are Malang Regency with 246,600 people, Jember Regency with 226,570 people, and Sumenep Regency with 211,980 people. While the areas with the lowest number of poor people are occupied by urban areas, namely Mojokerto City with 6,630 inhabitants, Madiun City with 7,690 inhabitants, and Batu City with 7,890 inhabitants. This visualization shows a different pattern in the distribution of criminal events described previously. In some areas, the high number of poor people is not always accompanied by a high crime rate.

Furthermore, Figure 2c shows the distribution pattern of the mean year schooling (MYS) for the population of East Java province which tends to be high in urban areas. The three regions with the highest RLS are Madiun City (11.08 years), Mojokerto City (10.68 years), and Malang City (10.46 years). Meanwhile, areas with low MYS are the horseshoe area and Madura Island. The three regions with the lowest MYS are Sampang Regency with 5.86 years, followed by Sumenep Regency with 6.11 years, and Bangkalan Regency with 6.28 years.



(a) (b)

(c) (d)



(e)

Figure 2. Variable thematic map: a) Open Unemployment Rate; b) Number of Poor People (in thousands); c) Mean Year Schooling; d) Gini Ratio; e) Population to Police Ratio

The Gini ratio shows various values for each district/city (Figure 2d). Observing deeply at the Gini ratio, a fairly high inequality still occurs in some areas in the western part of East Java Province. The highest Gini ratio is in Nganjuk Regency at 0.42, followed by Surabaya City at 0.41, and Pacitan Regency at 0.40. While the lowest Gini ratio is in three regions, namely Mojokerto Regency, Sampang Regency, and Probolinggo City with a Gini ratio value of 0.27.

The distribution of the population to police ratio is presented in Figure 2e. The population to police ratio in the central and western regions is lower than in the east. The three areas with the highest population to police ratio are Malang Regency with 1:2.185, Jember Regency with 1:2.098, and Probolinggo Regency with 1:2.022. Meanwhile, the three regions with the lowest population to police ratio are Mojokerto City with 1:212, Madiun City with 1:236, and Blitar City with 1:259.

4.2 Modelling the variables influencing crime incidents in East Java

Multicollinearity checking was carried out as an initial step in modelling to ensure that the variables used were not correlated with each other. The results of multicollinearity checking show the value of VIF < 10 for all variables. The VIF value for each independent variable is presented in Table 2.

Table 2. VIF value of the independent variable

X_1	X_2	X_3	X_4	X_5
1.598762	2.71171	1.64982	4.58364	4.76973

Source: R-studio

This result proves that there is no significant linear relationship between the independent variables. Thus, it can be concluded that the non-multicollinearity assumption is satisfied and the five independent variables can be used in the analysis.

The next step is modelling the relationship between the dependent variable and the independent variables. As already mentioned, the dependent variable is count data, so the first model used is the Poisson regression model. The test results of the Poisson regression parameters are presented in table 3.

Table 3. Poisson regression parameter estimation coefficient

Variable	Coefficient	*Std. Error*	*Z-value*	*P-value*
(Constant)	-1,251	0,00847	-14,77	<2e-16*
X_1	0,02644	0,00069	37,85	<2e-16*
X_2	0,006326	0,00001	32,12	<2e-16*
X_3	0,0286	0,00077	36,67	<2e-16*
X_4	8,993	0,01862	48,30	<2e-16*
X_5	0,000462	0,000002	22,36	<2e-16*

Source: processing results with R-studio
Note: *) Significant at the 5 percent significance level

Based on Table 1, it is known that the five variables have a z-value greater than the value $Z_{\alpha/2} = 1{,}96$ and p-value less than 0.05. These results indicate that the five independent variables have a significant effect on crime in East Java Province in 2019. After the Poisson regression model is formed, overdispersion checks are carried out. The results of overdispersion checking show the value of the dispersion parameter, the ratio between the deviance and degrees of freedom, resulting in a value of 171.925 and a p-value of 0.00002709. The calculation results show that with a confidence level of 95 there is overdispersion. Thus, the Poisson regression model is no longer appropriate to use.

One alternative model that can be applied to model overdispersion of count data is negative binomial regression. The negative binomial regression model has certain dispersion parameters so that it can overcome the overdispersion case. The results of modelling with negative binomial regression are as follows:

Table 4. Estimated coefficient of negative binomial regression model parameter

variable	coefficient	*Std. Error*	*Z-value*	*p-value*
(Constant)	1.588050	1,12745	1,409	0,1590
X_1	0.362824	0,08780	4,132	0,0000*
X_2	0.080400	0,00251	0,834	0,4045
X_3	0.003207	0,09645	1,276	0,2021
X_4	5.483709	2,41120	2,274	0,0230*
X_5	0.000477	0,00028	1,674	0,0942

Source: processing results with R-studio
Note: *) Significant at the 5 percent significance level

The variable open unemployment rate (TPT) and the Gini ratio variable (GINI) have a z-value greater than the value $Z_{\alpha/2} = 1{,}96$ and a p-value less than 0.05. These results indicate that the two variables have a significant effect on crime in East Java Province in 2019, while the other variables are not significant.

The next stage is a spatial exploration to identify the existence of spatial effects in the form of spatial autocorrelation and spatial heterogeneity. Spatial autocorrelation examination with Global Moran's I test resulted in a value of 0.181 with a p-value of 0.037. These results indicate that there is a spatial autocorrelation and a positive value for the variable number of crimes in East Java Province. After testing the spatial autocorrelation, it is continued with spatial heterogeneity testing to test the spatial variation of crime in East Java Province. Tests were carried out using the Breusch-Pagan test. The test resulted in a test statistic of 17.756 with a p-value of 0.003268. The value of the resulting test statistic is greater than the value of $\chi^2_{(0,05;5)}$=11,0705 with the p-value is smaller than 0.05, so it can be concluded that there is a spatial effect in the form of spatial heterogeneity between regions. Thus, the negative binomial global regression model is not appropriate to use. Due to this problem, the GWNBR is applied to accommodates the different parameter values in each region

To perform modelling with GWNBR, a spatial weighting matrix is required. A spatial weighing matrix is obtained by determining the optimum bandwidth value. The optimum bandwidth value is obtained by minimizing the CV value. The kernel weighting function used in this study is a fixed bisquare kernel because the observed area is uniformly distributed. After the spatial weighting matrix is formed, a model is formed using the GWNBR method. For the 38 districts/cities observed, the modelling produces 190 parameter coefficients $\left(\hat{\beta}_j(u_i, v_i)\right)$.

After the model is formed, the simultan and part test are applied. Simultaneous testing was carried out by comparing the deviance value of the GWNBR model of 7,442.46 to the value of $\chi^2_{(0,05;5)}$=11,0705. Thus, at the 5 percent significance level, it can be concluded that there is at least one parameter that significantly influences the number of criminal acts in East Java Province in 2019. After simultaneous testing shows significant results, then partial testing is carried out. The test is carried out by comparing the resulting z-value to the value of $Z_{\alpha/2} = 1{,}96$. If the resulting z-value is greater than 1.96, it means that the variable has a significant effect on the number of criminal acts partially. The test results show that the variables of the open unemployment rate, the number of poor people (in thousands), the Gini ratio, and the ratio of the population to the police have a significant effect in all regions, while the mean year schooling variable has a significant effect in certain areas, but has no significant effect in other areas. A summary of the parameter coefficients for the GWNBR model is presented in Table 5.

Table 5. GWNBR model parameter summary

	Minimum	Maximum	Average
Theta	4,214445	4,218581	4,216119
(Constant)	0,017318	0,018208	0,017711
The open unemployment rate (X_1)	-0,055210	-0,051417	-0,053092
The number of poor people (in thousand) (X_2)	0,000206	0,000295	0,000250
Mean year schooling (X_3)	0,139038	0,141319	0,140145
Gini Ratio (X_4)	0,005730	0,005956	0,005858
Population to police ratio (X_5)	0,000152	0,000158	0,000155

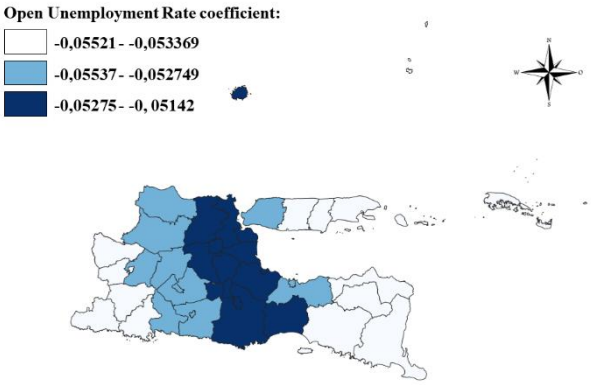
Source: Processing results with R-studio

The variable open unemployment rate significantly influences the number of criminal acts in all districts/cities in East Java Province. On average, every one percent increase in the open unemployment rate will reduce the number of criminal acts by $exp\ (-0{,}053092) = 0{,}9482$ times cases. Coefficients with high values are found in the west and east regions, while for the middle region they have moderate to low coefficient values as shown in Figure 3a. The open unemployment rate variable harms crime, meaning that when the open unemployment rate increases, it decreases crime. This is because unemployment in East Java Province is dominated by educated unemployed and aged 15-24 years [2]. Educated unemployed are the high school and college graduates who are still unemployed. Some of the unemployed activities in this group are by choice, not by compulsion. They decide to be unemployed because of a preference to get jobs that match their skills. They generally come from prosperous families so that they can achieve a high level of education. Due to this fact, they still get guarantees, the family economic support, satisfying their basic needs while unemployed. Based on the age characteristics, unemployment is dominated by young people (15-24 years). In this age group, most of them are not yet the head or the breadwinner of the family. So that, there is no strict responsibility to fulfil the economic needs of the family.

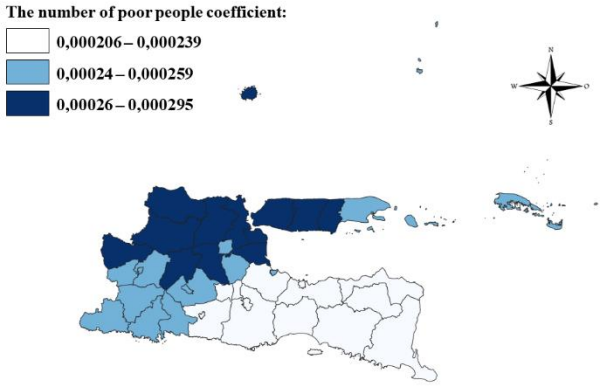
The variable number of poor people significantly affects the number of criminal acts in all districts/cities in East Java Province. On average, every 1,000 people increase in the number of poor people will increase the number of criminal acts by $exp\ (0{,}00025) = 1{,}0002$ times cases. The high-value coefficients generally can be found in the northern region and parts of Madura Island. However, the moderate-value coefficients dominate the southern region, while the low-values coefficients dominate the eastern part of East Java province, as shown in 3b. This result is in line with the previous

study [13], in which when the number of poor people increases in a given year, it will lead the number of criminal acts to increase. The crime motivated by a person's poverty aims to meet the needs. This is also similar to the research [14] that states that short term poverty will not lead someone into crime, but in the long term, it will increase crime. The sudden and short-term poverty will not make someone commit a crime, because they are still trying to earn an income for some time. However, long-term poverty is possible to cause stress so that individuals will adopt criminal acts.

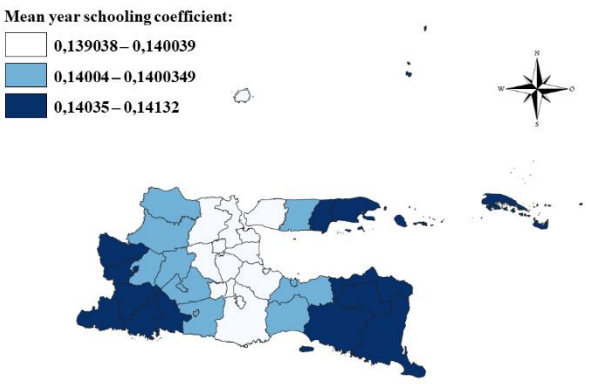
The mean year of schooling is significant only in some areas. On average, every one-year increase in the mean year schooling will increase the number of criminal acts by $exp\,(0{,}140145) = 1{,}1504$ times. Figure 3c shows that the western and eastern regions have a high-value coefficient, while the south-central regions have a moderate-value coefficient, and the north-central regions mostly have a low-value coefficient. This fact shows that crime begins to involve residents who are pursuing higher education both as victims and as perpetrators of criminal acts, as found in the research [15]. Supporting this phenomenon, a theory [16] explained that as far as schools can educate their pupils to become good citizens and treat people well, education can reduce the tendency of a person to commit crimes. Opposite to this theory, there are many "white-collar" crimes such as fraud, forgery, and embezzlement done by highly educated people in Indonesia. This phenomenon is motivated by the declining moral education, considering that schools tend to emphasize their pupils to master certain scientific practices and skills without equipped them with an understanding of morality [17].



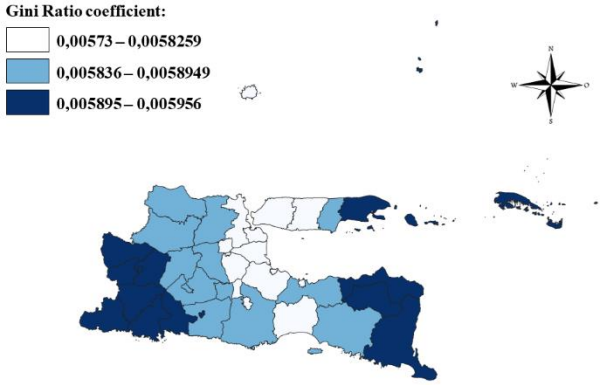
(a)



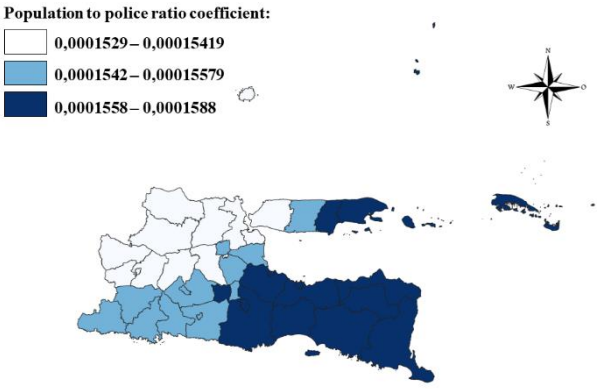
(b)



(c)



(d)



(e)

Figure 3. Thematic map of the distribution of parameter coefficient estimates: a) Open unemployment rate; b) Number of poor people (in thousands); c) Mean year schooling; d) Gini ratio; e) Population to police ratio

The mean year of schooling is significant only in some areas. On average, every one-year increase in the mean year schooling will increase the number of criminal acts by $exp\,(0{,}140145) = 1{,}1504$ times. Figure 3c shows that the western and eastern regions have a high-value coefficient, while the south-central regions have a moderate-value coefficient, and the north-central regions mostly have a low-value coefficient. This fact shows that crime begins to involve residents who are pursuing higher education both as victims and as perpetrators of criminal acts, as found in the research [15]. Supporting this phenomenon, a theory [16] explained that as far as schools can educate their pupils to become good citizens and treat people well, education can reduce the tendency of a person to commit crimes. Opposite to this theory, there are many "white-collar" crimes such as fraud, forgery, and embezzlement done by highly educated people in Indonesia. This phenomenon is motivated by the declining moral education, considering that schools tend to emphasize their pupils to master certain scientific practices and skills without equipped them with an understanding of morality [17].

The Gini ratio variable significantly influences the number of criminal acts in all districts/cities in East Java Province. On average, every 0.1 unit increase in the Gini ratio will increase the number of criminal acts by $exp\,(0{,}005858) = 1{,}00058$ cases. Coefficients with high values are found in the west and east regions, while the middle region is dominated by moderate to low coefficient values as shown in Figure 3d. The Gini ratio reflecting inequality is the main problem that triggers crime. The previous study [18] explains that personal inequality happens because of welfare inequality between regions, economic discrimination in the political sphere, and poverty. Income inequality in an area will cause social jealousy, which triggers people to commit crimes [19].

The population to police ratio variable significantly influences the number of criminal acts in all districts/cities in East Java Province. On average, each increase in the ratio of the population to the police by one unit will increase the number of criminal acts by $exp\,(0{,}000155) = 1{,}000155$ times cases. The population to police ratio is the ratio of one police officer to a certain number of residents. If the higher the value of this ratio, the greater the responsibility of the police. The population to police ratio has a significant effect on the incidence of crime in East Java Province. This result is motivated by the fact that most districts/cities do not have the ideal ratio of 1:350, except Kediri, Blitar, Mojokerto, and Madiun [20]. A previous study [21] concluded the same result in which there was a positive relationship between the number of police officers and crime, which is motivated by the insufficient number of police officers to secure and eradicate crimes. Crime generally occurs not only in one area but also can move to other areas. This fact sometimes causes difficulties for the police to arrest the criminals, as stated in [22] that the number of existing police personnel is still insufficient to reduce and prevent crime.

Table 6. Grouping of districts/cities based on significant variables

No	Districts/cities	Significant variables
1	Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Pasuruan, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Gresik, Bangkalan, Sampang, Pamekasan, Sumenep, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu	X_1, X_2, X_3, X_4, X_5
2	Kediri, Malang, Lumajang, Jember, Banyuwangi, Bondowoso, Situbondo. Probolinggo, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo	X_1, X_2, X_4, X_5

As the final stage of the analysis, Table 6 presents the results of regional grouping based on the parameter coefficient values and the partial test. Variables that have a significant effect in all regions are the open unemployment rate, the number of poor people in thousands, the Gini ratio, and the population to police ratio. Meanwhile, the variable of mean year schooling is significant only in some areas.

Regional grouping depends on the mean years of schooling which performs a significant difference for each area. Figure 4 shows that the mean years of schooling has no significant effect in eastern regions. A former study [23] found that the effect of education on crime follows an inverted U pattern which means that the highest chance of a person committing a crime occurs when a person reaches secondary education, but a lower or higher education level reduces a person's chances of committing a crime.

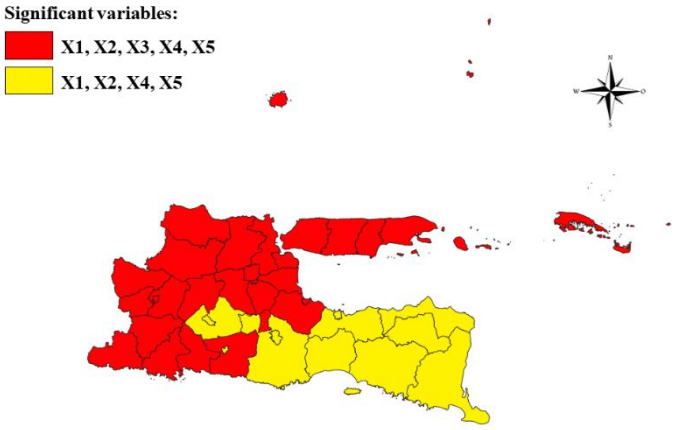


Figure 4. Thematic map of the grouping of districts/cities in East Java Province according to significant variables.

The formation of two regional groups shows that the factors that influence crime in each region are different. Thus, the government needs to consider this in making policies to reduce crime in its territory.

To find out a better model in explaining the effect of the independent variables used on the number of crimes in East Java Province, a comparison of the AIC values was carried out. The comparison of the AIC values of each model is as follows:

Table 7. Comparison of AIC values

Model	AIC
Poisson Regression	5818,60
Negative binomial regression	540,44
GWNBR	540,05

Table 7 shows that the GWNBR model has a smaller AIC value than the negative binomial global regression model and the Poisson global regression model. Thus, it can be concluded that the GWNBR model can explain the effect of the variable open unemployment rate, the number of poor people, the mean year schooling, the Gini ratio, and the population to police ratio on the number of crimes better than the Poisson global regression and the negative binomial global regression.

5. Conclusion

Based on the results and discussion, the following conclusions are obtained:

1. The Crime in East Java Province tends to form a cluster but varies between regions, in which the middle and eastern parts of this province are dominated by a moderate to a high number of crimes. However, the western areas of East Java province, mostly have a low number of crimes.
2. GWNBR model forms two groups based on the significant variables, the mean years of schooling. Mean Years of Schooling has no significant effect in the eastern part of East Java Province, which is dominated by lower Mean Years of Schooling. The U pattern of education effect on crime stated that the opportunity of commiting a crime reduces for lower or higher eduation level.
3. Based on the value of AIC, the GWNBR model has the best performance in modelling the crimes in East Java 2019.

This research provides some suggestions as follows:

1. Although the open unemployment rate can reduce the number of criminal acts in 2019, it must be supervised and controlled wisely. The local governments in each district/city must strive to provide job opportunities that are suitable with the competence of highly educated graduates.
2. Higher education does not guarantee a person not to commit a crime. Government efforts are still needed to improve public education quality, especially moral education in areas where the mean years of schooling has a significant effect on increasing the number of criminal acts, such as Surabaya City, Sidoarjo Regency, the western part of East Java Province, and Madura Island
3. Based on the security point of view, it is necessary to distribute police personnel evenly and proportionally to the total population in each district/city, to reduce the population to police ratio. In addition, intensive patrols are needed to provide wider protection for the community.
4. Future research may enhance the analysis by applying other methods that can overcome the symptoms of overdispersion, for example, the generalized Poisson regression. In addition, further exploration and analysis of the spatial effects on the effect of lower education in crime numbers due to the U pattern effect.

References

[1] Goulas, E., & Zervoyianni, A. (2015). Economic growth and crime: Is there an asymmetric relationship? *Economic Modelling*, **49**, 286-295.

[2] BPS Provinsi Jawa Timur. (2019a). *Laporan Eksekutif Keadaan Angkatan Kerja Provinsi Jawa Timur 2019*. Surabaya: Badan Pusat Statistik Provinsi Jawa Timur.

[3] Johannes, H.V., Padmadisastra, S., & Tantular, B. (2017). Model Geographically Weighted Negative Binomial Regression (GWNBR) untuk Data Kejahatan. *Prosiding Seminar Statistika FMIPA UNPAD 2017*, 260-267.

[4] Priambodo, B. W. Y., & Irhamah, I. (2019). Pemetaan Jumlah Property Crime di Provinsi Jawa Timur Menggunakan Metode Geographically Weighted Negative Binomial Regression (GWNBR) dan Geographically Weighted Poisson Regression (GWPR). *Inferensi*, **2(2),** 53–62.

[5] Felson, M. & Cohen, L.E. (1980). Human ecology and crime: A routine activity approach. *Human Ecology,* **8(4)**, 389-406.
[6] Cameron, A. Colin, Trivedi, Pravin K. (1998). *Regression analysis of count data.* Cambridge: Cambridge University Press
[7] Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression - the analysis of spatially varying relationships*. John Wiley & Sons, Ltd.
[8] BPS. (2020). Statistik Kriminalitas 2020. In *BPS-Statistics Indonesia*. Jakarta: Badan Pusat Statistik.
[9] Becker, G. S. (1986). Crime and Punishment: an Economic Approach. In *The Economic Dimensions of Crime* (pp. 15–68). London: Palgrave Macmillan.
[10] Anselin, L. (1995). Local indicators of spatial organization -LISA. *Geographical Analysis*, **27(2),** 93–115.
[11] Anselin, L. (1998). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
[12] Mustika, Diva Arum, Nooraeni, Rani. (2019). Kajian Efek Spasial Pada Kasus Difteri dengan Metode Geographically Weighted Negative Binomial Regression. *Indonesian Journal of Statistics and Its Application*, **3(1),** 91-104.
[13] Dulkiah, M., & Nurjanah. (2018). Pengaruh Kemiskinan Terhadap Tingkat Tindak Kriminalitas Di Kota Bandung. *Jurnal Ilmu Sosial Dan Ilmu Politik*, **8(2),** 36-57.
[14] Khan, N., Ahmed, J., Nawaz, M., & Zaman, K. (2015). The Socio-Economic Determinants of Crime in Pakistan: New Evidence on an Old Debate. *Arab Economic and Business Journal*, **10(2),** 73–81.
[15] Lispani, N. M. L., Sumarjaya, I. W., & Sukarsa, I. K. G. (2018). Pemodelan Jumlah Tindak Kriminalitas Di Provinsi Jawa Timur Dengan Analisis Regresi Spatial Autoregressive and Moving Average. *E-Jurnal Matematika*, **7(4),** 346-356.
[16] Lochner, L. (2020). Education and crime. In *The Economics of Education: A Comprehensive Overview* (pp. 109–117).
[17] Ervina, M. E. (2020). Hubungan Rata-rata Lama Sekolah, Gini Ratio, dan Pengeluaran Per Kapita dengan Tingkat Kejahatan Tahun 2011-2018. *Euclid*, **7(2),** 97-107.
[18] Edwart, A. O., & Azhar, Z. (2019). Pengaruh Tingkat Pendidikan, Kepadatan Penduduk dan Ketimpangan Pendapatan Terhadap Kriminalitas di Indonesia. *Jurnal Kajian Ekonomi Dan Pembangunan*, **1(3),** 759–768.
[19] Steviani, E., Handra, H., & Maryati, S. (2020). Faktor Sosial-Ekonomi Yang Mempengaruhi Tindak Kejahatan Di Provinsi Sumatera Barat. *Menara Ilmu*, **XIV(01),** 42–51.
[20] BPS Provinsi Jawa Timur. (2019b). *Statistik Politik dan Keamanan Jawa Timur 2019*. Surabaya: Badan Pusat Statistik Provinsi Jawa Timur.
[21] Nadilla, U., & Farlian, T. (2018). Pengaruh PDRB Perkapita, Pendidikan, Pengangguran, dan Jumlah Polisi Terhadap Kriminalitas di Provinsi Aceh. *Jurnal Ilmiah Mahasiswa*, **3(1),** 110–118.
[22] Nugroho, A. H., & Harmadi, S. H. B. (2016). Analisis Spasial Kriminalitas Harta Benda di Wilayah Jadetabek. *Jurnal Ekonomi Dan Pembangunan Indonesia*, **15(2),** 158.
[23] Jonck, P., Goujon, A., Testa, M. R., & Kandala, J. (2015). Education and crime engagement in South Africa: A national and provincial perspective. *International Journal of Educational Development*, **45**, 141–151.