Modeling Spatial Variation of Money Laundering Crime in Indonesia Using Geographically Weighted Multinomial Logistic Regression

A Yulia*, S Astutik and U Sa’adah
Faculty of Mathematics and Natural Sciences, Brawijaya University, Indonesia

*Corresponding author: anisyulia7@gmail.com

Abstract. This document presents the mapping of crime in Indonesia. Crime is a real threat that needs to be optimally handled which increases significantly and it is very detrimental to the country's economy. The crime that occurred created another more serious crime, namely Money Laundering Crime. The impact of Money Laundering Crime is the flow of illicit fund through the payment system to financial service providers as well as providers of goods and services. The perpetrator of Money Laundering crime is a player who is shrewd and familiar with the state financial system, so that the flow of funds does not involve a region but will expand to other areas in a country. This study analyses the spatial variations in the factors causing the Money Laundering crime in Indonesia by using a Geographically Weighted Multinomial Logistic Regression (GWMLR) technique. In this document, the GWMLR method is used to model the Money Laundering Crime and its explanatory variables by considering the effect of geographic location using the Adaptive Kernel weighting function. The response variables were divided into four categories with nominal scales by combining the general criminal case and money laundering crime.

1. Introduction
Regression analysis is a statistical analysis tool that utilizes the relationship between two or more variables [1]-[4]. The aim is to make unreliable estimates (predictions) for the value of a variable (response variable) if the value of another related variable (variable prediction) is known. In spatial data, regression analysis began to be developed by considering the effect of location or geographic location which was then called the model called Geographically Weighted Regression (GWR) [5]-[7]. GWR is starting to develop and not only uses simple spatial regression analysis, but can be applied to many other regression models, such as a logistic regression model called Geographically Weighted Multinomial Logistic Regression (GWMLR) [8]-[10].

Money laundering is a dual crime. By handling cases thoroughly in the existing of two crimes, namely predicate crime and money laundering crime, should be carried out more efficiently and effectively. In general, every perpetrator of a criminal act will try to hide or disguise the origin of assets that are the result of a crime in various ways so that the assets resulting from the crime are difficult to trace by law enforcement officials, so that he can freely use these assets for legal or illegal activities. Therefore Indonesian Law no. 8/2010 concerning the Prevention and Eradication of the Crime of Money Laundering states that the crime of money laundering not only threatens the stability and integrity of the economic system and the financial system, but can also endanger the joints of society, nation and state based on the 1945 Constitution of the Republic of Indonesia.

The impact of the crime of money laundering is the flow of illicit fund that is not only involves a region but also will develop into other areas in a country. Therefore, an area map is needed to find out criminal acts in each region in Indonesia to facilitate handling of crimes that occur.
2. Research Model

2.1 Data Source

In this research, the data used are data on factors that influence criminal acts obtained from the Central Statistics Agency (BPS) in 34 Provinces in Indonesia in 2019. The variable used in this research consists of one response variable ($Y$) with four categories with a scale nominal and five predictor variables ($X$). The response variable ($Y$) used is the number of cases of predicate crime and money laundering in the province. While the predictor variables used are Poverty Percentage ($X_1$), Open Unemployment Rate ($X_2$), Population Density ($X_3$), Higher Education Enrolment Rates ($X_4$), Per Capita Expenditure ($X_5$), Human Development Index ($X_6$), and Percentage of Households Have Access to Decent Housing ($X_7$).

2.2 Analysis Method

The method used in this research is Geographically Weighted Multinomial Logistic Regression using the best weight based on the smallest AIC value. The conducted assumption test is dependency spatial test and heterogeneity spatial test.

3. Result and Discussion

3.1 Descriptive Analysis

The variable dependent used in this research is criminal cases in Indonesia which are divided into 4 (four) categories. Descriptively, the causes of this crime could be explained in Figure 1.

![Figure 1. Distribution of Criminal Cases in Indonesia.](image)

Based on the distribution of criminal cases in Indonesia, it can be seen that as many as 7 provinces (20.6%) are included in category 1, namely the number of cases of Predicate Crime and Money Laundering in the Province which is greater than the average, 1 province (2.9%) which is included in category 2, namely the number of cases of predicate crime in large provinces and money laundering in small provinces, 4 provinces (11.8%) are included in category 3, namely the number of cases of predicate offense in small provinces, small regencies / cities and money laundering in big regencies / cities, and 22 provinces (64.7%) are included in category 4, namely the number of cases of predicate crime and money laundering in districts / cities that are small. Because there is only a fairly far comparison of the number of provinces between categories, in this study criminal cases are
divided into 2 categories. Descriptively, the research variable, the independent variable could be explained in Table 1.

Table 1. Descriptive Independent Variable.

| Variable                                | Minimum | Maximum | Average | Deviation Standard |
|-----------------------------------------|---------|---------|---------|--------------------|
| Poverty Percentage ($X_1$)              | 3.42    | 26.55   | 10.24   | 5.52               |
| Open Unemployment Rate($X_2$)           | 1.52    | 8.11    | 4.77    | 1.51               |
| Population Density($X_3$)               | 9.00    | 15900.00| 742.03  | 2708.76            |
| Higher Education Enrolment Rates($X_4$)| 14.27   | 73.14   | 33.59   | 10.75              |
| Per Capita Expenditure($X_5$)           | 7336.00 | 18527.00| 10915.59| 2213.47            |
| Human Development Index($X_6$)          | 60.84   | 80.76   | 71.04   | 3.91               |
| Percentage of Households Have Access to  | 26.16   | 81.61   | 54.14   | 12.42              |
| Decent Housing($X_7$)                   |         |         |         |                    |

3.2 Assumption Test

The assumption test that needs to be done is dependency spatial test and heterogeneity spatial test.

3.2.1 Dependency Spatial Test

The assumption dependency test that is carried out to determine whether the residual variable respond in a region is related to the surrounding region which is carried out by using Moran’s I test.

Table 2. Moran’s I test

|                     |       |
|---------------------|-------|
| Observed            | 0.1062|
| Expected            | -0.02702|
| Sd                  | 0.051193|
| P-Value             | 0.009256|

Based on Table 2, it is obtained a p-value of 0.009256 or less than $\alpha$, so that $H_0$ is rejected and it can be concluded that the dependency assumption is met.

3.2.2 Heterogeneity Spatial Test

Testing of the heterogeneity spatial assumption is carried out to determine variance of the residuals of different respond variable in each location or there is one observation location that has different residual variance. Testing of the Spatial Heterogeneity is carried out by using Breusch Pagan test.

Table 3. Breusch-Pagan Test.

| Breusch Pagan     | 13.728 |
|-------------------|--------|
| df                | 6      |
| P-Value           | 0.0362 |

The critical area rejects $H_0$ if the value of $BP > \chi^2(\alpha, p)$. Obtained chi-square value $\chi^2(0.05, 6)$ of 3.325 and p-value is less than $\alpha$. So, it can be concluded that there is $\sigma_1^2 \neq \sigma_2^2$ or heterogeneity occurs. So that the assumption of spatial heterogeneity is fulfilled.
3.3 The testing result of Geographically Weighted Multinomial Logistics Regression

3.3.1 Selection of the best weighted

The first step in modeling a criminal case is to determine the weights to be used with the kernel function weights. In determining the best weight, what must be done is to determine the smallest AIC.

Table 4. Comparison of AIC Value Each Weighting Kernel Function

| Kernel Weighted         | AIC     | $R^2$  |
|-------------------------|---------|--------|
| Adaptive Gaussian Kernel| 268.562 | 0.688  |
| Adaptive Bisquare Kernel| 261.473 | 0.778  |

Based on Table 4, it is found that the appropriate weight is the Adaptive Bisquare Kernel because this weight has the smallest AIC value which is 261.473 with the highest $R^2$ value among other weights namely 77.8%. The estimation of the criminal act model will be carried out using the GWMLR approach with adaptive weighting Bisquare Kernel.

3.3.2 Parameter Testing

Furthermore, the variables that have a significant effect on criminal cases will be classified based on significant variables in each province in Indonesia according to Table 5.

Table 5. Grouping of Provincial Based on Predictor Variables with Significant Influence.

| Province                               | Significant Variable                                                                 |
|----------------------------------------|--------------------------------------------------------------------------------------|
| Aceh, West Sumatera, Riau              | 1. $X_3$ (Population Density)                                                        |
|                                        | 2. $X_5$ (Per Capita Expenditure)                                                    |
|                                        | 3. $X_7$ (Percentage of Households Have Access to Decent Housing)                    |
| North Sumatera, West Nusa Tenggara, South Kalimantan | 1. $X_3$ (Population Density)                                                        |
| Jambi, South Sumatera, Bangka           | 1. $X_3$ (Population Density)                                                        |
| Belitung Islands, Central Java, Yogyakarta, Bali, East Nusa Tenggara, West Kalimantan, Central Kalimantan, North Kalimantan, North Sulawesi, West Sulawesi, Maluku | 2. $X_7$ (Percentage of Households Have Access to Decent Housing)                    |
| Bengkulu, Lampung, Riau Islands, DKI Jakarta, West Java, East Java, Banten | 1. $X_2$ (Open Unemployment Rate)                                                    |
|                                        | 2. $X_3$ (Population Density)                                                        |
|                                        | 3. $X_7$ (Percentage of Households Have Access to Decent Housing)                    |
| Central Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, North Maluku, West Papua, Papua | 1. $X_3$ (Population Density)                                                        |
|                                        | 2. $X_6$ (Human Development Index)                                                   |
|                                        | 3. $X_7$ (Percentage of Households Have Access to Decent Housing)                    |
The final step is to make a description of the distribution of predictor variables that have a significant effect on criminal cases in each province in Indonesia using thematic maps. The use of this thematic map will make it easier to visually describe groupings of regions. The distribution of areas according to significant variables in each province can be presented in Figure 2, as follows.

![Figure 2. The Factors that Effect Significantly to the Criminal Act in each Province.](image)

Figure 2 shows the distribution of the predictor variables that have a significant effect on each province, areas that tend to be visible, with the same factors. Provincial grouping that shows the existence of significant influencing factors, for example occurs in Central Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, North Maluku, West Papua, and Papua, with factors that have a significant effect are population, Human Development Index, and the Percentage of Households Have Access to Decent Housing.

With parameter estimation that obtained, the Model GWMLR formed. One of the Model GWMLR in East Java Province as follows:

\[
g_1(x) = 1.231 + 0.77X_2 + 4.029X_3 - 4.136X_7
\]

\[
g_2(x) = 2.231 + 0.623X_2 + 2.098X_3 - 3.222X_7
\]

\[
g_3(x) = 0.445 + 0.854X_2 + 3.097X_3 - 3.109X_7
\]

Criminal cases in East Java Province are influenced by three factors, that is Open Unemployment Rate \((X_2)\), Population Density \((X_3)\), and Percentage of Households Have Access to Decent Housing \((X_7)\). With calculating the odd ratio, interpretation of the first Logit Model is the higher the open unemployment rate in the area, the chances of the high number of cases of predicate crime and money laundering will increase 2.16 times compared to the chances of having a low number of predicate crimes and money laundering cases.

4. Conclusion

This study examines whether the Poverty Percentage \((X_1)\), Open Unemployment Rate \((X_2)\), Population Density \((X_3)\), Higher Education Enrolment Rates \((X_4)\), Per Capita Expenditure \((X_5)\), Human Development Index \((X_6)\), and Percentage of Households Have Access to Decent Housing \((X_7)\) spatially affects the number of cases of predicate crime and money laundering in the province \((Y)\). The results of the analysis using the Geographically Weighted Multinomial Logistics Regression show that
the Population Density variable ($X_3$) affects the number of cases of Origin and Money Laundering in the Province ($Y$). While the variable Open Unemployment Rate ($X_2$), Per Capita Expenditure ($X_5$), Human Development Index ($X_6$), and the Percentage of Households who Have Access to Decent Housing($X_7$) have an effect on several provinces in Indonesia.

References

[1] Stancu I and Rece D 2009 The relationship between economic growth and money laundering—a linear regression model *Theoretical and Applied Economics* 9 (9) 20-29
[2] Stanley T D and Jarrell S B 2005 Meta-regression analysis: a quantitative method of literature surveys *Journal of economic surveys* 19 (3) 299-308
[3] Kaneko H, Arakawa M and Funatsu K 2008 Development of a new regression analysis method using independent component analysis *Journal of chemical information and modeling* 48 (3) 534-541
[4] Chao S H and Chen Y H 2019 A novel regression analysis method for randomly truncated strong-motion data *Earthquake Spectra* 35 (2) 977-1001
[5] Brunsdon C, Fotheringham A S and Charlton M E 1996 Geographically weighted regression: a method for exploring spatial nonstationarity *Geographical analysis* 28 (4) 281-298
[6] Yu H, Fotheringham A S, Li Z, Oshan T, Kang W and Wolf L J (2020) Inference in multiscale geographically weighted regression *Geographical Analysis* 52 (1) 87-106
[7] Mennis J 2006 Mapping the results of geographically weighted regression *The Cartographic Journal* 43 (2) 171-179
[8] Luo J and Kanala N K 2008 Modeling urban growth with geographically weighted multinomial logistic regression *Geoinformatics 2008 and Joint Conference on GIS and Built Environment: The Built Environment and Its Dynamics* 7144 71440
[9] Mustafa A, Heppenstall A, Omrani H, Saadi I, Cools M and Teller J 2018 Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm *Computers, Environment and Urban Systems* 67 147-156
[10] Laili I 2018 *Estimasi parameter model mixed Geographically Weighted Multinomial Logistic Regression (GWMLR) dengan Pembobot Adaptive Bisquare Kernel* Doctoral dissertation Universitas Islam Negeri Maulana Malik Ibrahim,Indonesia