Identification of Spatial Autocorrelation in the Poverty Level in West Pasaman Regency with Moran Index

F M Sari¹, H Frananda², S Fransiska¹

¹Mathematics Department, Universitas Negeri Padang, Padang, Indonesia
²Geography Department, Universitas Negeri Padang, Padang, Indonesia

*corresponding author: fransiskasonia74@gmail.com

Abstract. West Pasaman Regency is one of the 50 largest disadvantaged regions in Indonesia. One indicator of disadvantaged areas is there are still many poor people in the area. The level of poverty in an area is estimated to be influenced by the poverty of the surrounding area. Relationships between these regions can be known by calculating spatial autocorrelation. This research aims to determine spatial autocorrelation in the poverty level in West Pasaman using the Moran Index. The results of this study indicate that the poverty level in West Pasaman have positive spatial autocorrelation, but the correlation is weak because the Moran Index value is close to 0, which is 0.0765.

1. Introduction

West Pasaman Regency is one area in the province of West Sumatra that has various regional characteristics such as: forests, beaches, and land. Natural Resources in the highlands with mountains and hills in the East of West Pasaman Regency, lowlands with agricultural areas and coastal and marine areas with coastlines of approximately 152 km are the capital and power to improve the regional economy. The potential resources of West Pasaman Regency are very promising, such as the economic potential in the fields of mining, forestry, plantation, food crops, animal husbandry, fisheries as well as tourism and other potentials. However, this potential has not been fully utilized. According to the evaluation of the Ministry of Villages, Development of Underdeveloped Regions and Transmigration (Kemendes PDTT), Pasaman Barat Regency is among the top 50 disadvantaged areas in Indonesia, which is the 33rd position. One indicator of disadvantaged areas is that there are still many poor people.

Poverty is a problem which until now is still a central issue in various parts of the world, including Indonesia. BPS (The Indonesian Statistics Agency) states that in September 2017 the number of poor people (residents with per capita expenditure per month below the Poverty Line) in Indonesia reached 26.58 million people or 10.12% of the total population of Indonesia. Many factors affect poverty fluctuations in Indonesia, including: regional topography (coastal, valleys, or land), geographical location of villages (located around forests or outside forest areas), supporting facilities for transportation (types of paved, graveled, or land roads), educational facilities (number of kindergarten, elementary and junior high schools), health facilities (presence of puskesmas / auxiliary puskesmas, use of own latrines, frequency of posyandu activities), economic facilities (presence of shops), income of residents (farm laborers), and land use (especially paddy fields). These factors directly or indirectly occur in villages and are interconnected between one village and another. In other words, these factors
are spatial. Therefore, the identification of spatial patterns and spatial autocorrelation (spatial dependency) testing of poverty in West Pasaman will be carried out.

Identification of the spatial autocorrelation is done using the Moran Index method while in determining the weighting matrix using the Queen Contiguity method. The data used was the poverty level in West Pasaman Regency in 2017 taken from the “Kabupaten Pasaman Barat dalam Angka Tahun 2018” [1].

2. Literature review

2.1. Spatial statistics
Spatial statistics are statistical methods used to analyze spatial data. Spatial data is data that has georeferenced in which various attribute data are located in various spatial units[2]. Recently, spatial data has become an important medium for development planning and sustainable management of natural resources in continental, national, regional and local areas.

The first law about geography was stated by W Tobler. Tobler in Anselin [4] argues that, “everything is related to everything else, but near things are more related to each other”. This law is the pillar of regional science studies. It can be concluded that spatial effects are a natural occurrence between one region and another.

2.2. Spatial data analysis
According to De Mers in Budiyanto [3], spatial analysis is a technique or process that involves several or a number of calculation functions and evaluation of mathematical logic that can be performed on spatial data, in order to obtain added value, extraction and new information with spatial aspects. Spatial analysis consists of three groups; visualization, exploration, and modeling. Spatial analysis results are presented through visualization. Statistical methods are used to process data in exploration. The concept of causal relations to predict the existence of spatial patterns is demonstrated through modeling [5]. The location of the spatial data must be measured in order to know the spatial effects that occur. According to Kosfeld [6], information of location can be known from two thins, i.e:

2.2.1. Neighborhood.

Spatial neighborhood units are usually formed based on maps. It is expected to reflect a high degree of spatial dependence when compared to spatial units that are located far apart.

2.2.2. Distance.

Location which is located in a certain space with the existence of latitude and longitude becomes a source of information. The distance between points needed in space is calculated using this information. It is expected that the spatial dependence will decrease according to the distance.

The most important thing in spatial analysis is the existence of weights or often referred to as spatial weighting matrices. Spatial weighting matrices are used to determine the weights between locations that are observed based on neighborhood relations between locations. Spatial weighting matrix is also called a matrix that illustrates the strength of interaction between locations. Figure 1 shows the proximity (contiguity) of the position or location of a location to other locations. According to Anselin (1995), weighting matrices can be divided into three, including:

1) Rook contiguity, the area of observation is determined based on the sides that intersect and the angle is not taken into account.
2) Bishop contiguity, the area of observation is determined based on the intersecting angles and sides are not taken into account.
3) Queen contiguity, the observation area is determined based on the intersecting sides and the angle is also taken into account.
Spatial weighting matrix W can be obtained in two ways, namely standardized contiguity matrix ($W_{\text{standardize}}$) and unstandardized contiguity matrix ($W_{\text{unstandardize}}$). The standardized contiguity matrix is a weighting matrix obtained by giving equal weight to the nearest neighbour location and the others is zero, whereas the unstandardized contingency matrix is a weighting matrix that is obtained by giving one to the nearest neighbor location and the other zero.

2.3. Spatial autocorrelation
Spatial autocorrelation is an estimate of the correlation between observed values related to the spatial location of the same variable. Characteristics of spatial autocorrelation revealed by Kosfeld [6], i.e:

a) If the observed variables have a spatial distribution pattern, then there is a spatial autocorrelation.

b) If the distance between regions is closer, then there is positive spatial autocorrelation (tend to be in groups).

c) Asymmetrical patterns of neighborliness indicate negative spatial autocorrelation (tend to spread).

d) The absence of spatial autocorrelation is indicated by random patterns.

One of the autocorrelation measurements for spatial data can be calculated using the Moran’s Index method. This method can be used to detect the beginning of spatial randomness. This spatial randomness can indicate patterns that cluster or form trends in space [6]. According to Kosfeld [6], the calculation of spatial autocorrelation by the Moran Index method can be done in two ways, i.e:

1) Moran index with an unstandardized spatial weighting matrix $W^*$

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^* (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

where: $S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^*$

$w_{ij}^*$: elements of the unstandardized weights between regions i and j

2) Moran index with a standardized spatial weighting matrix W

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

The range of values from the Moran’s Index in the case of standardized spatial weighting matrices is $-1 \leq I \leq 1$. Negative spatial autocorrelation is indicated by $-1 \leq I < 0$, while positive spatial autocorrelation is indicated by values $0 < I \leq 1$, the Moran’s Index value is zero indicate not in groups. If the weighting matrix used is unstandardized weighting, then the measurement accuracy cannot be
guaranteed using the Moran Index. To identify whether or not spatial autocorrelation was performed, the Moran Index significance test was performed. Hypothesis testing for the Moran’s Index is as follows:

\( H_0: \) There is no spatial autocorrelation  \\
\( H_1: \) There is spatial autocorrelation

Test Statistics:

\[
Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \approx N(0,1)
\]

where:

\[
E(I) = \frac{-1}{n-1}
\]

\[
S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}
\]

\[
S_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} (w_{ij} + w_{ji})^2
\]

\[
S_2 = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ji} \right)^2
\]

Test Criteria:

Reject \( H_0 \) at the significance level \( \alpha \) if \( Z(I) > Z_{1-\alpha} \) with \( Z_{1-\alpha} \) is \((1-\alpha)\) quantile from the standard normal distribution

3. Results

3.1. Spatial Data Exploration of poverty levels in West Pasaman Regency

The poverty level data in West Pasaman Regency can be described spatially on an area basis through exploration of the data that you want to observe on a map of the area as shown in the figure 2. By presenting in quantile form as shown in Figure 2 it can be seen that the high poverty rate (indicated in red) in West Pasaman District tends to be clustered, as well as for medium (yellow) and low (green). Graphically the spatial pattern is not very visible, even it can still be called a pattern of spreading randomly meaning it does not indicate a spatial pattern. To be able to ensure whether there is a spatial pattern, especially spatial dependencies, further testing is needed.
3.2. Spatial Autocorrelation Testing

In this paper a queen contiguity spatial weighting matrix approach is used based on angles and sides based on a map of the West Pasaman administrative region. The selection of the queen contiguity type spatial weighting matrix assumes that the boundaries of the area between districts / cities can be passed by the population. The spatial weighting matrix formed like (unstandardize queen contiguity) in the table 1.

|   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 2 | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 3 | 1   | 1   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| 4 | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   |
| 5 | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 6 | 0   | 0   | 0   | 0   | 1   | 0   | 1   | 1   | 0   | 0   | 0   |
| 7 | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 1   | 0   | 0   | 0   |
| 8 | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 0   |
| 9 | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 10| 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 11| 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |

Whereas the standardized spatial weighting matrix (standardize queen contiguity) in the table 2.

|   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 0   | 0.5 | 0.5 | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 2 | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 3 | 0.33| 0.33| 0   | 0   | 0.33| 0   | 0   | 0   | 0   | 0   | 0   |
| 4 | 0   | 0   | 0   | 0   | 0.5 | 0.5 | 0   | 0   | 0   | 0   | 0   |
| 5 | 0   | 0   | 0.5 | 0.5 | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
After obtaining the weighting matrix, the identification of spatial autocorrelation is carried out using the Moran Index method as follows:

Hypothesis testing for the Moran Index is as follows:
H₀: there is no spatial autocorrelation
H₁: there is spatial autocorrelation

Moran Index values obtained as follows:

The test criteria are reject H₀ at the significance level of α if the p-value is more than α. By using a significance level of 95%, a p-value of 0.2437 is obtained so that it is known that the p value is more than α, then the decision fails to reject H₀, meaning that there is no spatial autocorrelation at the poverty level in West Pasaman Regency in 2017.

From the Moran's I test above, it was concluded that at a significance level of 5% it was stated that there was no spatial autocorrelation of HDI figures between regencies / cities in West Papua Province in 2012. However, Moran's I statistical value of 0.0765 was in the range of 0<I ≤ 1 which means it shows a positive spatial autocorrelation pattern.

4. Conclusion
The results of the statistical analysis concluded that between sub-districts in West Pasaman District there was no spatial autocorrelation of poverty levels in 2017. However, it is necessary to further study what aspects affect poverty levels related to regional proximity. It should be noted that geographically West Pasaman Regency has various regional characteristics such as: forests, beaches, and land, so it is possible that economic and educational activities depend on the ease of transportation. Therefore, it would be better if the selection of spatial weights was also determined based on the level of difficulty of access between regions.

The calculation results obtained by the Moran Index value of 0.0765. This value indicates that the spatial autocorrelation that occurs at the poverty level in West Pasaman Regency is a positive spatial autocorrelation but the correlation is weak because it is close to zero. This means that there is no spatial autocorrelation at the poverty level in West Pasaman District, so the pattern formed is random.

References
[1] Badan Pusat Statistik (BPS) 2018 Kabupaten Pasaman Barat dalam Angka. Pasaman Barat: Kantor Badan Pusat Statistik Pasaman Barat
[2] Banerjee S 2004 *Hierarchical Modeling and Analysis for Spatial Data* Boca Raton: Chapman and Hall/CRC

[3] Budiyanto E 2010 *Sistem Informasi Geografis dengan ArcView GIS* Yogyakarta: Penerbit Andi

[4] Anselin L 1993 *Exploratory Spatial Data Analysis and Geographic Information Systems*, National Center for Geographic Information and Analysis of California Santa Barbara: CA93106

[5] Pfeiffer D, *et al.* 2008 *Spatial Analysis in Epidemiology* New York: Oxford University Press

[6] Kosfeld R 2006 *Spatial Econometric* URL: http://www.scribd.com