Spatial Data Panel Analysis for Poverty in East Java Province 2012-2017

A M Yolanda*, K Yunitaningtyas, and Indahwati
Department of Statistics, IPB University, Jln. Meranti Wing 22 Level 4 Kampus IPB Dramaga, Bogor 16680, Jawa Barat, Indonesia

Corresponding author: anne_mudya@apps.ipb.ac.id

Abstract. Poverty is a phenomenon of spatial heterogeneity, which is usually indicated by the tendency of poor people to group in a certain area. It means that poverty is influenced by the spatial effect. Geographical conditions and economic inequality that occurred in various regions in Indonesia, especially East Java reinforced the reason for adding spatial effects in the analysis. In addition, as an effort to obtain an increasingly accurate model, poverty that occurs is highlighted in a period of time. The aim of this study is to determine the factors that influence the percentage of poor people according to regency/city in East Java province in 2012-2017 by a spatial data panel analysis. The best model for this case is Spatial Autoregressive (SAR) model using the concept of distance. It shows that poverty in East Java is related to spatial effect. Therefore, each regency/city has different constants, depending on its spatial weighting matrix. Thus, the percentage of poor people by Regency/City in East Java province was influenced by Human Development Index (HDI), Poverty Gap Index (P1), Poverty Severity Index (P2) and also the spatial effect.

1. Introduction
Poverty is still a crucial problem in a developing country, including Indonesia. It is true that the national poverty rate tends to decline year to year. But, in some region, the poverty rate is higher than the national rate. Publication of Statistics Indonesia 2017 claim that the highest poverty rate was in East Java province, even though East Java as is of the business and industrial center in Indonesia. The percentage of poor people in East Java is around 11.77 in 2017. Poverty is a big chore. That is why the government are rightly concerned in eradicating poverty.

The high percentage of poverty is also related to geographical conditions and economic inequality. Henninger & Snel [1] in Rahayu [2] reveals that poverty is a phenomenon of spatial heterogeneity, which is usually indicated by the tendency of poor people to group in a certain area. Geographical variations in poverty and the magnitude of poverty are often caused by factors with spatial dimensions, such as the contribution of natural resources and access to services such as health and education. According to The First Law of Geography: everything is related to everything else, but near things are more related than distant things [3]. It means that basically, the near area has the same characteristics. That is why Crandall & Weber [4] signifies that reducing poverty in a place will affect and be influenced by other places around it. In other words, it can assume that poverty has spatial effects. In this way, simple regression will not enough to solve the problem, because the observation contains spatial information [5].

Furthermore, there are 38 regencies/cities in East Java with varied regional contour. According to publication from BPS-Statistics Indonesia, the percentage of poor people is quite diverse in every
Regency/City in East Java province. These facts give a strong reason for adding spatial effects in the model.

In addition, as an effort to obtain an increasingly accurate model, poverty that occurs is highlighted in a period/time period. Based on Elhorst [6], the benefit of adding spatial panel is that can handle both spatial and time effects. Spatial observation are seemingly to differ in their background variables, which are usually space-specific time-invariant variables that have an effect on the response variable, but hard or exhausting to get. Mostly, the previous study present that spatial panel data panel obtained larger $R^2$.

It is evident that the Spatial Data Panel is a good method to analyze poverty in East Java.

Problems regarding poverty in East Java have been studied in previous studies namely Setiawati and Setiawan [7] in the Modelling of Poor Population in East Java with the Econometric Approach Spatial Panels, and Metyopandi et al. [8] in the study entitled Spatial Panel Regression Method in Modelling Poverty Levels in the regency/city in East Java in 2014. In addition, cases using the Spatial Panel studied by Nandori [9] to examines whether economic growth (or recession) has an effect on poverty and if so, what kind of effect it is, Karbasi and Najkar [10] to examine provinces variation in poverty rates in Iran, and many others study related to Spatial Panel Data Analysis [11, 12, 13, 14, 15, 16, 17, 18]. The aim of this study is to determine the factors that influence the percentage of poor people by regency/city in East Java in 2012-2017 by a spatial data panel analysis.

2. Research Method

The Panel data is a combination of time series and cross-section data, i.e. time series data is from an object over a period of time, while cross-section data is for several objects for a given period [19,20]. Panel data analysis is used when observing an event is not enough to only be carried out on observation units at a given time, but also observing these units at various time periods [21]. Spatial panel data is combination of location, time series, and cross-section data. Simply, this model can capture spatial interactions across observation and over time [22]. Moreover, spatial dependence happens when observations do not seem to sampled at random, that is why spatial panel is more reliable than classic panel model [23].

Parameters in panel data can be estimated by three approaches, namely Common Effect, Fixed Effect, and Random Effect [24]. Common Effect is not paying attention to individual units or time. It assumed that it is the same in various time periods. The model in fixed effect changes for each individual and time reflected through intercept. Random Effect checks for errors, there may be a correlation between time series and cross-section data. Chow and Hausman’s test are used for selecting which approaches should be used to estimate.

Breusch-Pagan test is widely used to check time series effect. Spatial effects can be divided into spatial dependence and spatial heterogeneity. Spatial dependence occurs due to dependence between areas and spatial heterogeneity occurs due to the diversity between areas [25]. Spatial Heterogeneity is tested by Breusch-Pagan test, while Moran’s Index for Spatial Dependence.

Once specifying the interaction between spatial units, the model might contain dependence on spatial lag or a spatial autoregressive in the error term, known as the spatial lag and spatial error model [26]. This analysis requires panel data with one response variable and several explanatory variables are observed in a certain time period by adding spatial effect in the model.

The relationship between one response variable and one or more explanatory variable(s) can be expressed in a linear regression model [27]. Yet, spatial, time series, and cross section are in the model, thus General model of spatial regression panel data as shown in equation

$$y_{it}^* = \rho \sum_{j=1}^{N} w_{ij} y_{jt}^* + \beta_0 + \sum_{j=1}^{k} \beta_j x_{jit} + \mu_i + \phi_{it}$$

and

$$\phi_{it} = \lambda \sum_{j=1}^{N} w_{ij} \phi_{jt} + \epsilon_{it}$$

with
\begin{equation}
    y_{it}^* = \beta_0 + \sum_{j=1}^{k} \beta_j x_{jit} + \phi_{it}
\end{equation}

\begin{equation}
    y_{it}^* = \rho \sum_{j=1}^{N} w_{ij} y_{jt}^* + \beta_0 + \sum_{j=1}^{k} \beta_j x_{jit} + \epsilon_{it}
\end{equation}

The best model is the model with higher $R^2$. $R^2$ or coefficient of determination is usually used to see the compatibility of regression models [28]. $R^2$ describes the proportion of total variance in the response variables which is explained by the model. The greater the value of $R^2$, variance of the response variables can be explained by the model better.

3. Data Description

Data in this study is about Poverty in East Java Province, collected from BPS-Statistics Indonesia [29]. There are 38 regencies/cities in East Java are included from 2012-2017. The variables used are summarized from various previous studies on poverty.

BPS-Statistics Indonesia [30] has used the concept of a basic needs approach to measuring poverty. Therefore, poverty is called as an economic inability to fulfill food and nonfood basic needs which are measured by consumption/expenditure. An individual whose expenditure/capita/month is below the poverty line is considered to be poor. Definition of all variable in this study according to BPS-Statistics Indonesia [30] shown below.

\begin{table}[h]
\centering
\caption{Variables for Analysis}
\begin{tabular}{|c|c|}
\hline
\textbf{Variable} & \textbf{Description} \\
\hline
\textbf{Response Variable} & \textbf{Percentage of poor people} \\
& The percentage of population under the poverty line, counted by Head Count Index \\
\hline
\textbf{Explanatory Variable} & \textbf{Human Development Index (HDI)} \\
& a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable, and have a decent standard of living \\
\hline
\textbf{Explanatory Variable} & \textbf{Poverty Gap Index (P1)} \\
& Measures the extent to which individuals fall below the poverty line (the poverty gaps) as a proportion of the poverty line. The higher value of the index shows that the gap between the average expenditure of the poor and the poverty line is wider. \\
\hline
\end{tabular}
\end{table}
### Poverty Severity Index (P2)
Describes inequality among the poor. The higher value of the index shows that inequality among the poor is higher

### Results and Discussion

The percentage of poor people in East Java in the period 2012-2017 is shown in the diagram below.

**Figure 1.** Percentage of Poor People in East Java 2012-2017

![Figure 1](image)

Figure 1 shows that in the past six years, the percentage of poor people in this province, in general, has decreased. The percentage of poor people by regency/city in East Java in 2017 appears as follows.

**Figure 2.** Boxplot of Percentage of Poor People in East Java 2012-2017

Based on the boxplot above, it appears that diversity is relatively homogeneous during 2012-2017. Besides that, it also shows that the data tends to be symmetric. There is an outlier during 2012-2015, it is Sampang district.
According to figure 3, it can be seen that the administrative areas whose colors are getting yellow have a greater percentage of poor people from other regions, namely Sampang district and the lowest one is Malang. Figure 3 also shows that regency/city with a high percentage of poor people close to each other.

Figure 3. The percentage of poor people by regency/city in East Java 2017

Figure 4. Poverty Depth Index (P1)

Figure 5. Poverty Severity Index (P2)
Figure 4 shows that Bangkalan district has the highest Poverty Depth Index (P1) and Malang city has the lowest index, whereas from figure 5, Pasuruan city has the smallest Poverty Severity Index (P2) and Bangkalan regency has the highest index.

Table 2. Breusch-Pagan Test

| Effect            | Breusch-Pagan Test | df | p-value |
|-------------------|--------------------|----|---------|
| Time and Spatial  | 7.0088             | 2  | 0.7047  |
| Spatial           | 6.8652             | 1  | 0.008789|
| Time              | 0.14358            | 1  | 0.7047  |

Data obtained by Breusch-Pagan test and it presented that location and time effect are statistically significant.

Table 3. P-values of Chow and Hausman Test

|                 | Chow Test | Hausman Test |
|-----------------|-----------|--------------|
|                 | 2.2e-16   | 2.2e-16      |

Using p-values in the tables shown above can conclude that with a 5% significant level, the best analysis for this panel data is Fixed Effect.

This model is expressed as

\[ Poverty_{it} = -0.280092 \text{HDI}_{it} + 2.636556 P1_{it} - 5.718554 P2_{it} + \mu_t \]

The spatial dependence test is carried out by using Moran's Index for data of which are taken from 38 regencies/cities in East Java from 2012 to 2017. The results indicate that the percentage of poor people according to regencies/cities in East Java is influenced by spatial effects or in other words the dependences among areas existed.

Spatial regression of panel data was analyzed in three conditions, namely the General Spatial Model, Spatial Autoregressive (SAR) and Spatial Error Model (SEM). All three are analyzed by entering the fixed effect panel data model.

Table 4. General Spatial Model

| Effect | Coefficient | P-Values |
|--------|-------------|----------|
| \( \rho \) | -0.77539 | 0.04594 |
| \( \lambda \) | 0.56885 | 1.391e-06 |
| HDI | -0.081890 | 0.06536 |
| P1 | 2.581107 | 1.786e-13 |
| P2 | -5.623708 | 2.720e-11 |

As can be seen from table 3, at the 5% significant level, the HDI variables significantly affected the model. If this explanatory variable is not included in the model, then the results of the variables P1, P2, spatial lag, and spatial errors are statistically significant. That is, there are lag dependencies and spatial errors simultaneously on the data.

Table 5. Spatial Autoregressive (SAR)

| Coefficient | P-Values |
|-------------|----------|
| \( \rho \) | 0.31086 | 0.02491 |
| HDI | -0.175915 | 0.001609 |
| P1 | 2.611407 | 7.071e-13 |
Coefficient | P-Values
--- | ---
P2 | -5.661162 | 1.249e-10

Table 4 shows that all variables and spatial lag are significant influence are statistically significant with a confidence level of 95%.

**Table 6. Spatial Error Method (SEM)**

| | Coefficient | P-Values |
| --- | --- | --- |
| \(\lambda\) | 0.22503 | 0.2276 |
| HDI | -0.282594 | 2.2e-16 |
| P1 | 2.569741 | 3.154e-12 |
| P2 | -5.566849 | 4.226e-10 |

Based on table 5, all explanatory variables are statistically significant at a significant level of 5%. But, the parameter of the error of the dependent variable is bigger than p-values (0.05), there is sufficient evidence to refute the existence of spatial error dependencies on the data. Thus, it is not included in the selection of the best models.

**Table 7. R² Comparison**

| Model | R² |
| --- | --- |
| Panel Regression with Fixed Effect | 59.63% |
| General Spatial Model with Fixed Effect | 99.44% |
| Spatial Autoregressive Model (SAR) with Fixed Effect | 99.46% |

Based on the R², it can be stated that the best model is the Spatial Autoregressive (SAR) with Fixed Effect. The R² is 99.46%, mean that variance of the percentage of poor population by regency/city in East Java can be explained by the model and the remainder is explained by other factors outside the model.

SAR with Fixed Effect can be expressed as

\[
\hat{Poverty}_{it} = 0.31086 \sum w_{ij}y_{jt} - 0.175915HDI_{it} + 2.611407P1_{it} - 5.566849P2_{it}
\]

The results from the case just discussed suggest that the percentage of poor people by regency/city in East Java is influenced by the Human Development Index (HDI), Poverty Gap Index (P1), Poverty Severity Index (P2), and also spatial effects.

The Human Development Index and Poverty Severity Index are negative, meaning that these two variables show a negative correlation on the percentage of poor people. When the Human Development Index rises by one unit, the percentage of poor people will decrease by 0.1759. The coefficient of WY is positive, then a positive correlation is formed. On the other hand, the Poverty Gap Index is positive, so when the value rises by one unit, the percentage of poor people will also increase by 2.611407.

5. **Conclusion**

Thus, the percentage of poor people by regency/city in East Java is influenced by Human Development Index (HDI), Poverty Gap Index (P1), Poverty Severity Index (P2), and also spatial effects. The best
model for this case is Spatial Autoregressive (SAR) model using the concept of distance. It shows that poverty in East Java is related to spatial effect. Therefore, we can conclude that poverty in neighboring regions is interrelated. By SAR model, each regency/city has different constants, depending on its spatial weighting matrix, so that estimation of the percentage of poor people generated from the model obtained will be more accurate.

Acknowledgments
We thank our lecturers and colleagues from Department of Statistics, IPB University for sharing insight and expertise that greatly assisted the research. We are also immensely grateful to BPS-Statistics Indonesia for the data used in this research.

References
[1] Henninger N and Snel M 2002 *Where are the Poor? Experience with the Development and Use of Poverty Maps* (Arendal: World Resources Institute and UNEP/GRID)
[2] Rahayu N S 2017 *Geographically Weighted Panel Regression For Modelling The Percentage Of Poor People In Jawa Tengah Province [Thesis]* (Master Program Department of Statistics Institut Teknologi Sepuluh Nopember Surabaya)
[3] Tobler W R 1970 *A Computer Movie Simulating Urban Growth in the Detroit Region Econ. Geogr. (Supplement Proceedings, Int. Geogr. Union. Comm. Quant. Methods)* 46 234–40
[4] Crandall M and Weber B 2004 Local Social and Economic Conditions, Spatial Concentration of Poverty, and Poverty Dynamics *Am. J. Agric. Econ.* 86 1276–81
[5] Anselin L 1988 *Spatial Econometrics: Method and Models* (Boston: Kluwer Academic Publisher)
[6] Elhorst J P 2017 Spatial Panel Data Analysis *Encyclopedia of GIS* ed S Shekhar, H Xiong and X Zhou (Cham, Switzerland: Springer International Publishing,) pp 2050–8
[7] Setiawati A K and Setiawan 2012 *Permodelan Penduduk Miskin di Jawa Timur dengan Pendekatan Ekonometrika Panel Spasial J. Sains dan Seni POMITS* 1 183–7
[8] Metzopandi V, Fitriani R and Sumarminingsih E 2014 *Metode Regresi Panel Spasial pada Pemodelan Tingkat Kemiskinan Di Kabupaten / Kota Provinsi Jawa Timur J. Mhs. Stat.* 2 325–8
[9] Nandori E S 2012 Spatial Econometric Analysis of Poverty in Northern Hungary Between 2000 and 2010 *J. WEI Bus. Econ.* 1 81–91
[10] Karbasi A and Najkar N 2013 *Poverty Analysis in Iran : Spatial Regression Approach* *Int. J. Agric. Crop Sci.* 6 1436–40
[11] Pramono G and Marsisno W 2018 *Availibility of Infrastructure for Poverty in Indonesia : Spatial Panel Data Analysis Econ. Financ. Indones.* 64 157–80
[12] Sameti M and Faramand S *Spatial Analysis of Income Inequality , Poverty and Economic Growth in the Euro-Med Zone* (Faculty of Administrative Sciences and Economics, University of Isfahan, Hezar-Jerib ST, Istahan, Iran)
[13] David A, Guilbert N, Hamaguchi N, Higashi Y, Hino H, Leibbrandt M and Shifa M 2018 *Spatial Poverty and Inequality in South Africa: A Municipality Level Analysis AFD Res. Pap. Ser. No 2018-66* January
[14] South S J and Crowder K 2010 Neighborhood Poverty And Nonmarital Fertility: Spatial And Temporal Dimensions *J. Marriage Fam.* 72 89–104
[15] Barufi A M, Haddad E and Paez A 2012 *Infant mortality in Brazil, 1980-2000: A spatial panel data analysis BMC Public Health* 12 181
[16] Bhattarai M and Narayanamoorthy A 2003 *Impact of irrigation on rural poverty in India: An aggregate panel-data analysis Water Policy* 5 443–58
[17] Heshmati A, Maasoumi E and Wan G 2015 *Poverty Reduction Policies and Practices in Developing Asia* (Atlanta, USA: Co-publication of the Asian Development Bank and Springer International Publishing AG)
[18] Farrow A, Larrea C, Hyman G and Lema G 2005 Exploring the spatial variation of food poverty in Ecuador Food Policy 30 510–31
[19] Purwaningsih T 2014 Analyze Influence of Spatial Weighted Matrix in Spatial Panel Data Model [Thesis] (Department of Statistics, Graduate School IPB University, Bogor)
[20] Baltagi B H 2005 Econometrics Analysis of Panel Data (England: John Wiley and Sons, LTD)
[21] Arifatin D 2018 Spatial Analysis of Panel Data to Assess Factors Influencing Foreign Tourists in Indonesia 2011-2015 [Thesis] (Department of Statistics, Graduate School IPB University, Bogor)
[22] Lee L and Yu J 2010 Estimation of Spatial Panels Foundations and Trends® in Econometrics vol 4, ed W H Greene (USA: now Publishers Inc.)
[23] Ge X, Zhou Z, Zhou Y, Ye X and Liu S 2018 A spatial panel data analysis of economic growth, urbanization, and NOx emissions in China Int. J. Environ. Res. Public Health 15
[24] Munandar A 2017 Analisis Regresi Data Panel Pada Pertumbuhan Ekonomi Di Negara – Negara Asia J. Ilm. Ekon. Glob. Masa Kini 8 59–67
[25] Almudita N N 2012 Model Estimasi Permintaan Produk FAST Moving Consumer Goods (FMCG) [Final Project] (Universitas Indonesia, Depok)
[26] Elhorst J P 2010 Spatial Panel Data Models Handbook of Applied Spatial Analysis (Springer International Publishing) pp 377–407
[27] Draper N and Smith H 1992 Applied Regression Analisys [Indonesian Version] (Jakarta: Gramedia Pustaka Utama)
[28] Walpole R E and Myers R H 1995 Probability and Statistics for Engineers and Scientists [Indonesian Version] (Bandung: Penerbit ITB)
[29] [BPS Statistics East Java] 2017 Poverty [Accessed: 2018-10-30]. [Available: jatim.bps.go.id]
[30] [BPS Statistics Indonesia] 2017 Statistik Indonesia 2017 [Accessed: 2018-10-30]. [Available: bps.go.id]