
**MODELING LIFE EXPECTANCY IN CENTRAL JAVA
USING SPATIAL DURBIN MODEL**

Arief Rachman Hakim, Hasbi Yasin, Agus Rusgiyono

Department of Statistics, Faculty of Science and Mathematics, Diponegoro University

e-mail: arief.rachman@live.undip.ac.id

DOI: 10.14710/medstat.12.2.152-163

Article Info:

Received: 29 July 2019

Accepted: 19 December 2019

Available Online: 30 December
2019

Keywords:

*Life Expectancy, SDM,
Spatial Regression*

Abstract: Central Java in 2017 was one of the provinces with high life expectancy, ranking second. Life expectancy of Central Java Province in 2017 is 74.08% per year. The fields of education, health and socio-economics, are several factors that are thought to influence the life expectancy in an area. To find out what factors that the regression analysis method can use to find out what factors influence the life expectancy. But in observations found data that have a spatial effect (location) called spatial data, a spatial regression method was developed such as linear regression analysis by adding spatial effects. One form of spatial regression is Spatial Durbin Model (SDM) which has a form like the Spatial Autoregressive Model (SAR). The difference between the two if in the SAR model the effect of spatial lag taken into account in the model is only on the response variable (Y) but in the SDM method, effect of spatial lag on the predictor variable (X) and response (Y) are also taken into account. Selection of the best model using Mean Square Error (MSE), obtained by the MSE value of 1.156411, the number mentioned is relatively small 0, which indicates that the model is quite good.

1. INTRODUCTION

Central Java is the second highest province with high level of average Life Expectancy (LE) in Indonesia in 2017, with 74.08% per year. This figure was higher than that of 2016, which was 74.02% per year. The increase of LE in Central Java Province is attributable to several factors, which is interesting to identify to receive some factors supporting the increase of LE in Central Java province. It will be beneficial to use the results of the analysis to model the pilot in other areas with low level of average LE. Linear regression analysis is one of the appropriate statistical methods that allow us to identify any factors contributing to the increase in LE.

Linear regression analysis is a statistical method to look at the relationship between response variables and predictor variables in a linear manner, which is expected to predict the response variable if the predictor variable is known. However, the observation frequently discovers data that has a spatial effect (location) known as spatial data, making the data

analysis using the regression method inaccurate. Therefore, a spatial regression model was developed to use data modelling with spatial effects. This study used LE data and several factors that are assumed to influence LE in Central Java, which includes regencies and cities. The data contains information about location, where each regency and city has different characteristics. Linear regression method is considered insufficient to model and analyze objects in this study, and thus we need a method that can accommodate location (spatial) factors in the model. Therefore, this study used the spatial regression method for analysis.

This study used the Spatial Durbin Model (SDM) as one form of spatial regression, which is developed by Anselin (1988). This method is the development of the Spatial Autoregressive Model (SAR) with the same basic principle using area-based spatial data in its weighting. Its weighting principle uses the Contiguity matrix with the concept of contact between observed locations. The SDM model was selected because it has a basic form like SAR. The difference between the two is that the SAR model only takes into account the effect of spatial lag on the response variable (Y), while the SDM method also takes into account the effect of spatial lag on predictor variables (X) and response (Y).

Spatial regression may be used in the case of Life Expectancy (LE). The research by Tri K (2014) using the spatial lag regression method indicates that education, health, and socio-economic factors have a significant effect on LE in Central Java Province in 2012. Another study related to spatial regression by Ramadani (2013) used the SDM method for the Factors Affecting Toddler Malnutrition in Central Java. Therefore, the use of SDM methods in this study is expected to explain the factors that increase Life Expectancy for each regency and city in Central Java in accordance with the characteristics of each region.

2. LITERATURE REVIEW

2.1 Life Expectancy (LF)

Statistics Indonesia in 2018 defines that LE at an age x is the average life year to be lived by someone who has successfully reached age x in a given year and in the prevailing mortality situation in his community. Some factors are thought to have significantly influenced LE, namely education, health, and economic factors. The education factor is determined by the variable of average length of schooling. The high level of education of a person will also guarantee on going improvements in the level of technology used by society (Atmanti, 2005). The health factor will be represented by the variable of percentage of households that behave cleanly and healthily. The research also uses the number of *posyandu* (Integrated Service Post) variables.

Another factor is the availability of health facilities as an important factor to support health services, both in the form of treatment and prevention of disease. Clean and healthy living behaviour in the household has a significant influence on LE and the role on the level of public health. The health status of an environment is determined by the health of individuals, families and communities the way they behave. To determine the economic factors, the study used the percentage of the poor population and the adjusted expenditure per capita. The poor are residents with an average monthly per capita expenditure below the poverty line (BPS, 2018). Expenditures per capita are costs incurred for the consumption of all household members for a month divided by the number of household members. Expenditure can reveal patterns of household consumption in general. The level of economic welfare of the population can also be measured by the amount of household expenditure or consumption.

2.2. Spatial Regression

Spatial Regression is developed from linear regression method which includes the location impact. According to Fatati (2017), spatial regression is a statistical analysis method that evaluates the relationship between one variable with several other variables by providing spatial effects at several locations as the center of observation. The Spatial Regression Method was developed by Anselin (1988). This method is based on spatial area, which means that the weighting is done using spatial area data, with a weighting matrix in the form of a Contiguity matrix that uses the concept of contact between observed locations. The equation in SDM is similar with the Spatial Autoregressive Model (SAR) equation. The difference is that the SAR method affects the spatial lag only on the response variable (Y) that will be taken into account in the model, but the spatial lag of the SDM takes into account both the predictor variable (X) and response (Y).

According to LeSage (1999) in Rahmawati (2015), the general model of spatial regression is as follows:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (1)$$

$$\mathbf{u} = \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma_{\varepsilon}^2 \mathbf{I}_n)$$

where \mathbf{y} : response variable vector of size $n \times 1$
 ρ : coefficient of spatial lag parameters of the response variable
 \mathbf{W} : spatial weighting matrix ($n \times n$)
 \mathbf{X} : matrix of predictor variables with size of x ($k + 1$)
 $\boldsymbol{\beta}$: vector coefficient of regression parameter with size of $(k + 1) \times 1$
 λ : coefficient of spatial parameter error
 \mathbf{U} : error vector which has a spatial effect of size $n \times 1$
 $\boldsymbol{\varepsilon}$: vector error of size $n \times 1$

The general spatial regression model equation (1) also produces several other models, namely:

1. The value of $\rho = 0$ and $\lambda = 0$, is called the regression liner model or commonly known as the classic linear regression model forming the equation of:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

2. The value of $\rho \neq 0$ and $\lambda = 0$, is known as the Spatial Autoregressive Model (SAR) regression forming the equation of:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

3. The value of $\rho = 0$ and $\lambda \neq 0$, is known as the Spatial Error Model (SEM) regression forming the equation of:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}$$

4. The value of $\rho \neq 0$ and $\lambda \neq 0$, is called Spatial Autoregressive Moving Average (SARMA) with the form:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}$$

2.3. Spatial Durbin Model (SDM)

Basically, the mathematical equation of SDM is similar to that of SAR as in equation (2). However, the only difference of SDM is that it takes into account the spatial lag of both the predictor variable (X) and the response variable (Y). According to LeSage & Pace (2009), the SDM model has the following form of equation

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \boldsymbol{\alpha} \mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

or it can be written as follows:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \boldsymbol{\alpha} \mathbf{1}_n + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

with

$$\mathbf{Z} = [\mathbf{1}_n \quad \mathbf{X} \quad \mathbf{W}\mathbf{X}]$$

$$\boldsymbol{\delta} = \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\theta} \end{bmatrix}$$

where $\boldsymbol{\alpha}$: constant parameter vector measuring $1 \times n$

$\boldsymbol{\theta}$: spatial lag parameter vector predictor variable ($k \times 1$)

$\mathbf{1}_n$: vector containing number 1 measuring $n \times 1$

Maximum Likelihood Estimation (MLE) can be used in the process of estimating the Spatial Durbin Model regression parameters. According to LeSage (1999) in Ramadani 2013 it has the natural logarithmic function for estimating ρ :

$$l(\rho | \mathbf{y}) = C - \frac{n}{2} \ln \left([\mathbf{e}_0 - \rho \mathbf{e}_a]^T [\mathbf{e}_0 - \rho \mathbf{e}_a] \right) + \ln |\mathbf{I}_n - \rho \mathbf{W}|$$

where C is a constant that does not depend on the parameter ρ . The parameter ρ , which is in the range of $\frac{1}{\lambda_{\min}} < \rho < \frac{1}{\lambda_{\max}}$ with λ of the eigen value of the standardized \mathbf{W} matrix. Thus, parameter $\boldsymbol{\delta}$ is formulated with:

$$\hat{\boldsymbol{\delta}} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T (\mathbf{I}_n - \hat{\rho} \mathbf{W}) \mathbf{y} \quad (3)$$

so the estimated parameter $\boldsymbol{\delta}$ in equation (3) can be written as follows:

$$\hat{\boldsymbol{\delta}} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{y} - \hat{\rho} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{W} \mathbf{y} \quad (4)$$

2.4. SDM Model Fit Test

Ramadani (2013) used the following procedure for testing the suitability of the SDM model:

a. Hypothesis

$$H_0 : \rho = \beta_j = \theta_j = 0, \text{ with } j = 1, 2, \dots, k$$

$$H_1 : \rho \neq 0 \text{ or at least there is } \beta_j \neq 0 \text{ } \theta_j \neq 0 \text{ with } j = 1, 2, \dots, k$$

b. Test Statistics

$$F = \frac{SSR/k}{SSE/(n-k-1)} = \frac{MSR}{MSE} \quad (5)$$

c. The H_0 test criteria are rejected if $F > F_{(\alpha, k, n-k-1)}$ or p-value $< \alpha$

$$\begin{aligned} SSR \text{ (Regression Sum of Square)} &= \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \\ SSE \text{ (Residual Sum of Square)} &= SS_T - SS_R = \mathbf{y}^T \mathbf{y} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} \\ SST \text{ (Total Sum of Square)} &= \mathbf{y}^T \mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \end{aligned}$$

2.5. Test of The Parameters' Significance

Ramadani (2013) used Wald test for parameter significance test on SDM modeling. She used the following procedure for parameter testing:

a. Hypothesis

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

b. Test statistic

$$Wald_{\rho} = \frac{\hat{\rho}^2}{Var(\hat{\rho})} \quad (6)$$

$Var(\hat{\rho})$ = the diagonal element of the variance matrix that corresponds to ρ ,

To test the β parameter:

a. Hypothesis

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, k$$

b. Test statistics

$$Wald_{\beta} = \frac{\hat{\beta}_j^2}{var(\hat{\beta}_j)} \quad (7)$$

$Var(\hat{\beta}_j)$ = diagonal elements of the variance matrix corresponding to β .

The parameter testing of θ is as follow:

a. Hypothesis

$$H_0 : \theta_j = 0$$

$$H_1 : \theta_j \neq 0, j = 1, 2, \dots, k$$

b. Test statistics

$$Wald_{\theta} = \frac{\hat{\theta}_j^2}{Var(\hat{\theta}_j)} \quad (8)$$

$Var(\hat{\theta}_j)$ = diagonal element of the variance matrix corresponding to θ

The testing criteria H_0 is rejected if the value of $Wald > \chi_{\alpha,1}^2$.

2.6. Spatial Weighting Matrix

To determine weights between observed locations of the SDM method, we can use a spatial weighting matrix with the concept of neighbourhood relations of each location. Kosfeld, as quoted by Wuryandari, et al. (2014), explained that there are three ways of weighting. Rook contiguity weighting is determined based on the intersection between the sides of each location, Bishop contiguity weighting is determined based on confusion angles, and Queen contiguity weighting is determined based on intersections of both sides and angles.

| | | |
|------------|------------|------------|
| Location 1 | Location 2 | Location 3 |
| Location 4 | Location 5 | Location 6 |
| Location 7 | Location 8 | Location 9 |

Figure 1 Queen Contiguity Illustration

2.7. Moran's I Test

Moran's test is used to test the effect of spatial autocorrelation to see the correlation between observed values and location on the same variable. A systematic pattern in spreading a variable indicates the occurrence of that spatial autocorrelation. According to Wuryandari, et al. (2014), Moran's test can be defined as:

$$Z_{hitung} = \frac{I - E(I)}{\sqrt{Var(I)}} \quad (9)$$

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

$$E(I) = I_0 = -\frac{1}{n-1}$$

$$Var(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2 - 1) S_0^2} - [E(I)]^2$$

$$c_{ij} = (x_i - \bar{x})(x_j - \bar{x})$$

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

$$S_1 = \frac{1}{2} \sum_{i \neq j}^n (w_{ij} + w_{ji})^2 ; S_2 = \sum_{i \neq j}^n (w_{i0} + w_{0i})^2$$

$$w_{i0} = \sum_{i=1}^n w_{ij} ; w_{0i} = \sum_{j=1}^n w_{ji} .$$

The measurement of model compatibility uses R2 and Mean Square Error (MSE), according to Montgomery and Runger (2003). Adjusted R^2 is formulated as follows:

$$R_{adj,k} = 1 - \frac{SSE / (n - k - 1)}{SST / (n - 1)} \quad (10)$$

The value of $R_{adj,k}^2$ will only increase when the variable added to the model is a new variable that reduces the average square error. The next model fit size is MSE, whose calculation is done with the following formula:

$$MSE = 1 - \frac{SSE}{n - k - 1}$$

$$MSE = \frac{SS_E}{n - k - 1} \quad (10)$$

If the obtained MSE gets smaller to almost zero, it is possible to say that the regression model is getting better.

3. RESEARCH METHODS

3.1. Research Variables and Analysis

This study used data obtained from the catalogue issued by BPS 2018, namely *Katalog Jawa Tengah dalam Angka 2018* and books published by the Central Java Health Office in the form of Health Profile of Central Java Province in 2018. The variable used was one response variable (y) and five predictor variables (x). The details of the data variables used are explained below:

- y : Life expectancy at birth
- x_1 : Average length of schooling
- x_2 : Percentage of households with clean and healthy life
- x_3 : Number of *posyandu* (Integrated Service Post)
- x_4 : Percentage of Poor Population
- x_5 : Adjusted Per Capita Expenditure

In this study, the analysis covers the following steps:

1. Determining the spatial weighting matrix based on queen contiguity.
2. Performing Moran's I test to find out the spatial autocorrelation.
3. Estimating parameters with Spatial Durbin Model regression.
4. Performing model compatibility test and parameter significance test.
5. Calculating adjusted coefficient of determination (*Adjusted R²*) and MSE (Mean Square Error).

4. RESULTS AND DISCUSSION

4.1. Summary of Research Variables

A summary description of the research data used for the creation of the Spatial Durbin Model Life Expectancy (LE) by regency / city in Central Java Province in 2017, is shown in Table 1.

Table 1 Summary of Research Variables

| Variabel | N | Min | Max |
|--|----|-------|-------|
| Life expectancy at birth | 35 | 68.61 | 77.49 |
| Average length of schooling | 35 | 6.18 | 10.50 |
| Percentage of households with clean and healthy life | 35 | 59.69 | 97.25 |
| Number of <i>posyandu</i> (Integrated Service Post) | 35 | 164 | 2195 |
| Percentage of Poor Population | 35 | 20.32 | 4.62 |
| Adjusted Per Capita Expenditure | 35 | 14921 | 7785 |

Table 1 explains that the lowest level of LE in regencies/cities in Central Java Province in 2017 was at 68.61 per year and the highest was at 77.49 per year. The lowest average length of schooling in regencies/cities in Central Java Province in 2017 the lowest was at 6.18 years and the highest was in Semarang City with 10.5 years. The lowest percentage of households with clean and healthy living behavior in regencies/cities in Central Java Province in 2017 was 59.69% and the highest was 97.25%. The lowest number of *posyandu* in regencies/cities in Central Java Province in 2017 was 164 *posyandu* and the highest was 2195 *posyandu*. The lowest percentage of poor people in regencies/cities in Central Java Province in 2017 was at 4.62% and the highest at 20.32%. Expenditures per capita are adjusted in regencies/cities in Central Java Province in 2017 with the lowest amounting to IDR 7,785,- and the highest with IDR 14,921,-

4.2. Moran's I Test

The first step before calculating the estimated SDM parameters was spatial autocorrelation test. The calculation of Moran's I and Z values refers to equation (9), which was then processed with the help of R series 3.5.1 software. The complete results from the calculation of the Moran's I and Z values are presented in Table 2. The test covers the following steps:

Hypothesis:

$H_0: I = 0$ (no spatial autocorrelation between locations)

$H_1: I \neq 0$ (there is a spatial autocorrelation between locations)

Statistics Test:

$$E(I) = -\frac{1}{35-1} = -\frac{1}{34} = -0.02941176$$

$$S_0 = \left(\frac{1}{7} + \frac{1}{7} + \dots + \frac{1}{7} \right) = 35$$

$$S_1 = 18.5902$$

$$S_2 = 151.0808$$

$$Var(I) = \frac{35^2(18,5902) - 35(151,0808) + 3(35)^2}{(35^2 - 1)(35)^2} - [-0,0294]^2$$

$$Var(I) = 0,013$$

Tabel 2 Moran's I Test Results

| Variable | Moran's I Value | Z | Decision |
|----------|-----------------|---------|-------------------|
| y | 0.57522320 | 5.2978 | H_0 is rejected |
| x_1 | 0.19489697 | 1.9503 | H_0 is accepted |
| x_2 | -0.12105565 | -0.7779 | H_0 is accepted |
| x_3 | 0.06377452 | 0.8069 | H_0 is accepted |
| x_4 | 0.35389376 | 3.2687 | H_0 is accepted |
| x_5 | 0.17116800 | 1.7432 | H_0 is accepted |

H_0 is rejected if the value of / Z-count $|Z_{\alpha/2}| = 1.96$. Table 2 indicates that the value of |Z-count| of y variable and $x_4 < Z_{\alpha/2} = 1.96$, which means that H_0 is rejected. This indicates the existence of spatial autocorrelation between locations in both the response variable and the predictor variable, because the response variable and at least one of the predictor variables shows the existence of spatial autocorrelation between locations.

4.3. Regression Parameter Estimation of Spatial Durbin Model (SDM)

The SDM regression parameter estimation refers to equation (4) and the Wald value successively refers to equation (6, 7, 8) and it was computed using R series 3.6.1 software, the results of which is as listed in Table 3.

Table 3 Estimation Value of Spatial Durbin Model Parameter

| Parameter | Koefisien | Wald |
|------------------|--------------------------|------------------------|
| ρ | 0.63839 | 5.59767 |
| <i>Intercept</i> | 23.70977 | 1.80201 |
| β_1 | 0.61799 | 2.84664 |
| β_2 | 0.02843 | 1.32565 |
| β_3 | -2.6744×10^{-6} | 2.868×10^{-5} |
| β_4 | -0.19592 | 7.05707 |
| β_5 | -2.6278×10^{-4} | 1.44171 |
| θ_1 | -0.84814 | 0.70375 |
| θ_2 | -0.01468 | 0.07086 |
| θ_3 | 0.00138 | 1.38198 |
| θ_4 | 0.11522 | 0.42340 |
| θ_5 | 6.111×10^{-4} | 1.54393 |

Table 3 highlights that the SDM regression model can be formed as follows:

$$\hat{y}_i = 0.63839 \sum_{j=1}^n w_{ij} y_j + 23.70977 + 0.61799 x_{i1} + 0.02843 x_{i2} - 2.6744 x_{i3} \cdot 10^{-6} \\ - 0.19592 x_{i4} - 2.6278 x_{i5} \cdot 10^{-4} - 0.84814 \sum_{j=1}^n w_{ij} x_{j1} - 0.01467 \sum_{j=1}^n w_{ij} x_{j2} \\ + 0.00138 \sum_{j=1}^n w_{ij} x_{j3} + 0.11522 \sum_{j=1}^n w_{ij} x_{j4} + 6.111 x_{i5} \cdot 10^{-4} \sum_{j=1}^n w_{ij} x_{j5}$$

4.4. Fit Test of SDM Regression Model

To see the usability of the model as a whole, it is necessary to do a model compatibility test based on the following test:

Hypothesis

$$H_0 : \rho = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \theta_1 = \theta_2 = \theta_3 = \theta_4 = \theta_5 = 0$$

$$H_1 : \rho \neq 0 \text{ or there is at least one } \beta_j \neq 0, \theta_j \neq 0, j=1, 2, 3, 4, 5$$

The equation (5) which was computed with R software, obtained a value of $F = 8.793601$ where the value of F is more than that obtained from the $F_{(0,05;5;35)} = 2,49$ resulted from F table with the provisions of $F_{(\alpha,k,n-k-1)}$. Thus, it was decided that H_0 was rejected at the level of $\alpha = 5\%$. Therefore, it is possible to say that the SDM model with five predictor variables is suitable to describe the relationship between the response variable and the predictor variable.

4.5. Significance Test of SDM Regression Model Parameters

Parameter Significance Test was performed to see whether the parameters of each variable partially have an effect or not statistically. The test was carried out as follows:

Hypothesis

$$H_0 : \rho, \beta_j, \theta_j = 0, \text{ where } j = 1, 2, 3, 4, 5$$

$$H_1 : \rho, \beta_j, \theta_j \neq 0, \text{ where } j = 1, 2, 3, 4, 5$$

Table 3 indicating that the calculation of the Wald value refers to equation (6), (7), and (8) which was then computed with the help of software R, with test criteria if the value $Wald > \chi^2_{(\alpha,1)}$ (where $\chi^2_{(\alpha,1)}$ was resulted from the chi-square table) then H_0 is rejected.

The test reveals that the spatial lag parameters (ρ) and parameters β_4 have a significant effect on the model because the value of the $Wald > \chi^2_{0.05,1} = 3.841$, while the parameters $\beta_1, \beta_2, \beta_3, \beta_5, \theta_1, \theta_2, \theta_3, \theta_4$, and θ_5 statistically do not have a significant effect on the model because the value of Wald obtained $< \chi^2_{0.05,1} = 3.841$. Thus, it can be said that the spatial lag parameters and the percentage of poor population variables have a significant influence on LE in Central Java Province. As for the average length of school variables, the percentage of households with clean and healthy living, the number of *posyandu*, per capita expenditure adjusted and all spatial lag parameters of the predictor variables do not have a significant effect on LE in Central Java Province.

Calculation of Adjusted R^2 value by referring to equation 10 and computing with R software, obtained a value of 0.6963. This figure indicates that LE is influenced by the average length of schooling, the percentage of households with clean and healthy living, the

number of *posyandu*, the percentage of poor population and adjusted per capita expenditure as much as 69.63%, while the remaining 30.37% cannot be explained by the model. In other words, 30.37% is influenced by other factors outside the model. Calculation of Mean Square Error (MSE) value resulted from the SDM regression model by referring to equation 11 and computing with R software resulted in a value of 1.15641. The value was small enough and was almost 0 indicating that the formed SDM model is a good model.

5. CONCLUSION

Overall, the SDM model with five predictor variables is an appropriate model to illustrate the relationship between the response variable and the predictor variable. Spatial lag parameters and the percentage of poor population variables have a statistically significant effect on LE in Central Java Province. While the average length of school variables, the percentage of households behaving clean and healthy living, the number of *posyandu*, per capita expenditure adjusted and all spatial lag parameters of the predictor variables did not have a significant effect on LE in Central Java Province. Life expectancy is influenced by the average length of schooling, the percentage of households behaving clean and healthy living, the number of *posyandu*, the percentage of poor population and per capita expenditure as much as 69.63% and 30.37% affected by other factors outside the model. The value of MSE of 1.156411 approaching 0 indicates that the form SDM model is an appropriate model.

REFERENCES

- Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- Atmanti, H. D. 2005. Investasi Sumber Daya Manusia Melalui Pendidikan. *Jurnal Dinamika Pembangunan*, Vol 2, No. 1, pp. 30–39.
- Badan Pusat Statistik (BPS). 2018. *Jawa Tengah Dalam Angka 2018*. Semarang: Badan Pusat Statistik Jawa Tengah.
- Fatati, I. F., Wijayanto, H., and Soleh, A. M. 2017. Analisis Regresi Spasial dan Pola Penyebaran pada Kasus Demam Berdarah Dengue (DBD) di Provinsi Jawa Tengah. *Media Statistika*, Vol. 10, No. 2, pp. 95–105.
- Ramadani, I. R., Rahmawati, R., and Hoyyi, A. 2013. Analisis Faktor-faktor yang Mempengaruhi Gizi Buruk Balita di Jawa Tengah dengan Metode Spatial Durbin Model. *Jurnal Gaussian*. Vol. 2, No. 4, pp. 333–342.
- LeSage, J. and Pace, R. K. 2009. *Introduction to Spatial Econometrics*. New York: Taylor and Francis Group.
- Rahmawati, R., Safitri, D., and Fairuzdhiya, O. U. 2015. Analisis Spasial Pengaruh Tingkat Pengangguran terhadap Kemiskinan di Indonesia (Studi Kasus Provinsi Jawa Tengah). *Media Statistika*. Vol. 8, No. 1, pp. 23–30.
- Montgomery, D. C. and Runger, G. C. 2003. *Applied Statistics and Probability for Engineers*. New York: John Wiley and Sons, Inc.
- Tri K, A. I. 2014. *Aplikasi Regresi Spasial untuk Pemodelan Angka Harapan Hidup (AHH) di Provinsi Jawa Tengah*. Yogyakarta: Universitas Negeri Yogyakarta. Unpublished Thesis.

Wuryandari, T., Hoyyi, A., Kusumawardani, D., and Rahmawati, D. 2014. Identifikasi Autokorelasi Spasial Pada Jumlah Pengangguran di Jawa Tengah Menggunakan Indeks Moran. *Media Statistika*, Vol. 7, No. 1, pp. 1–10.