DEVELOPMENT LIFE EXPECTANCY MODEL IN CENTRAL JAVA USING ROBUST SPATIAL REGRESSION WITH M-ESTIMATORS

HASBI YASIN[*], ARIEF RACHMAN HAKIM, BUDI WARSITO

Department of Statistics, Faculty of Sciences and Mathematics, Diponegoro University, Indonesia

Abstract: Spatial regression model is used to determine the relationship between the dependent and independent variables with spatial influence. In case only independent variables are affected, Spatial Cross Regressive (SCR) Model is formed. Spatial Autoregressive (SAR) occurs when the dependent variables are affected, while Spatial Durbin Model (SDM) exists when both variables exhibit effects. The inaccuracy of the spatial regression model can be caused by outlier observations. Removing outliers in the analysis changes the spatial effects composition on data. However, using robust spatial regression is one way of overcoming the outliers in the model. Moreover, the typical parameter coefficients, which are robust against the outliers, are estimated using M-estimator. The research develops the life expectancy model in Central Java Province through Robust-SCR, Robust-SAR, and Robust-SDM to reduce spatial outliers' effect. The model is developed based on educational, health and economic factors. According to the results, M-estimator accommodates the outliers' existence in the spatial regression model. This is indicated by an increase in R^2 value and a decrease in MSE caused by the change in the estimating coefficient parameters. In this case, Robust-SDM is the best model since it has the biggest R^2 value and the smallest MSE.

Keywords: life expectancy; spatial regression; robust; spatial outliers.

2010 AMS Subject Classification: 93A30.

[*]Corresponding author

E-mail address: hasbiyasin@live.undip.ac.id

1. INTRODUCTION

According to Indonesia's Central Bureau of Statistics (BPS), Life Expectancy (LE) evaluates the government's performance in improving the population welfare and health. Based on the Central Bureau of Statistics, Indonesian LE increases annually. For instance, the country's LE in 2017 was 71.06 years. Yogyakarta and West Sulawesi provinces have the highest and lowest LE, respectively. Moreover, Central Java province holds the second position with an increase in LE from 74.02 in 2016 to 74.08 years in 2017 [1]. The enhancement of LE in Central Java Province is linked to educational, health, and economic factors. Regression analysis is a statistical method used to determine the factors influencing an increase in LE.

Regression analysis examines the relationship between dependent and independent variables [2]. Essentially, the methods used to deal with the spatial date are determined by the existence of location effects. In case spatial data are forcefully analyzed using classical linear regression, the assumptions of homogeneity and independence from errors are violated. For this reason, Spatial Regression Analysis is often used in studies [3],[4]. However, different models such as Spatial Cross Regressive (SCR), Spatial Autoregressive (SAR, and Spatial Durbin Model (SDM) are formed depending on the spatial effects [5].

In some cases, the outliers making parameter estimation appear bias. There are two outliers divisions, including Global Outliers whose value is significantly different from others, and Spatial Outlier, which is a spatially referenced object with relatively different non-spatial attributes [6]. Therefore, a robust regression method is used to analyze the data contaminated by outliers [7]. M-estimator, the most ordinary method theoretically and computationally, is one of the estimation methods on robust regression [8], and Moran's scatterplot detects outliers [9]. Therefore, Robust Spatial Regression is used to model Life Expectancy in Central Java. This study has a significant contribution in detecting the factors affecting life expectancy in Central Java Province, based on robust spatial regressions model.

2. MATERIAL AND METHODS

A. Life Expectancy in Central Java Province

The study was conducted in Central Java Province in Indonesia. There are several aspects with significant effects on Life Expectancy (LE), including educational, health, and economic factors. Educational factors are explained using the "Average Length of School (ALS)" variable. The "Percentage of Households with Clean and Healthy Living Behavior (PCHLB)" and the "Number

of Integrated Health Post (IHP)" explain health factors. "Percentage of Poor Population (PP)" and "Adjusted Per Capita Expenditure (APCE)" describes economic factors. The study used secondary data from the Central Java Province catalog in 2018 issued by the Jawa Tengah Province BPS-Statistics [1], and Health Profile of Central Java Book Profile 2017 issued by *Dinas Kesehatan Provinsi Jawa Tengah* [10]. The observation unit includes 35 regencies and cities in Central Java Province.
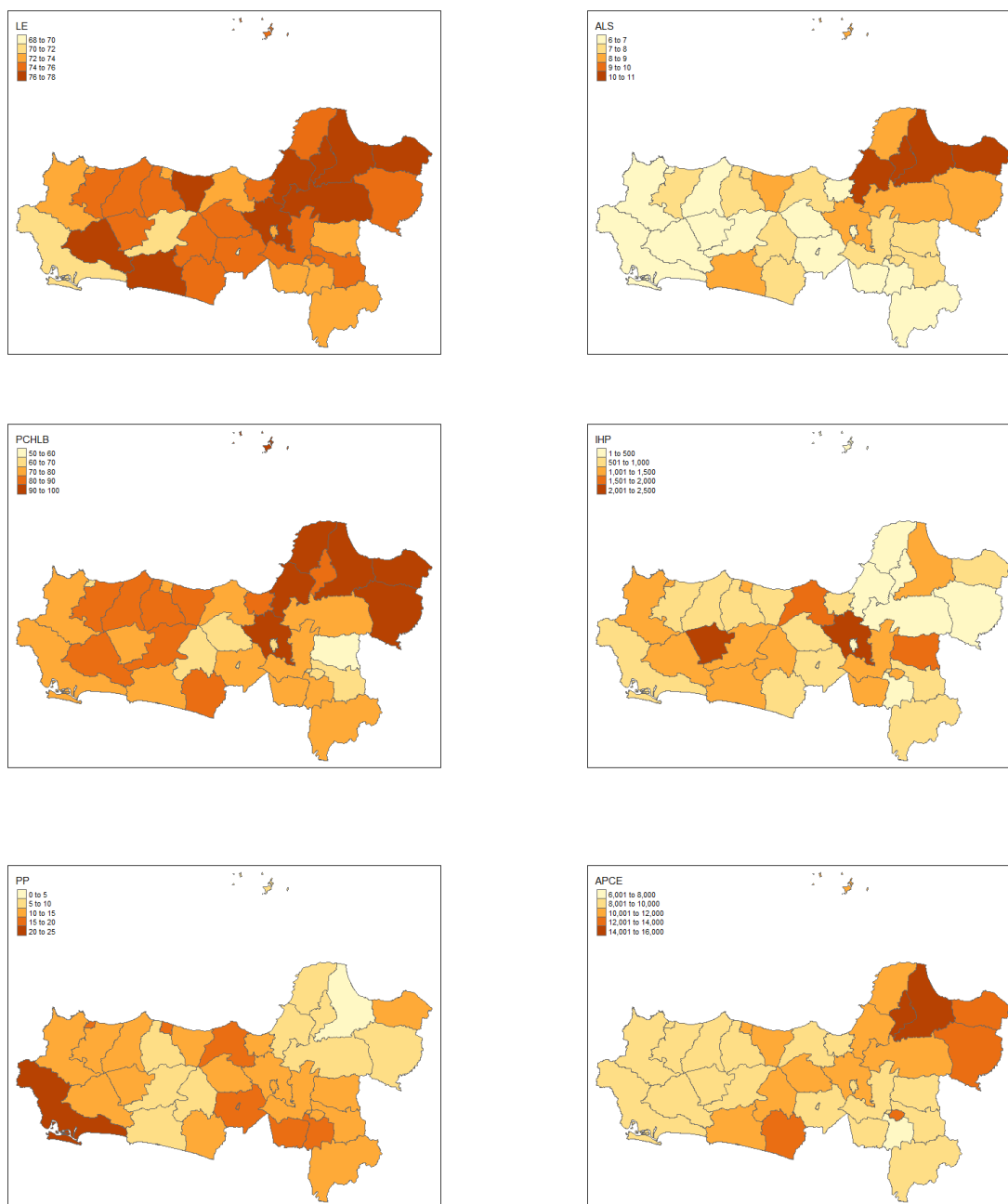
Figure 1. Spatial data distribution of each research variable

Figure 1 shows an overview of the spatial data distribution from each variable, while Table 1 indicates the data's description. Each region has different characteristics and form groups. Therefore, a spatial regression is needed in developing life expectancy models in Central Java Province.

Table 1. Description of Research Variable

Statistics	LE (Year)	ALS (Year)	PCHLB (%)	IHP (Unit)	PP (%)	APCE (Thousand IDR)
Min	68.61	6.18	59.69	164.00	4.62	7,785
1st Qu	73.42	6.71	71.60	600.00	9.33	9,238
Median	74.46	7.29	78.30	901.00	12.42	9,813
Mean	74.63	7.58	79.22	925.60	12.49	10,414
3rd Qu	75.90	8.26	88.06	1160.50	14.09	11,379
Max.	77.49	10.50	97.25	2195.00	20.32	14,921

B. *Moran's I Test to detect the spatial effect*

Spatial autocorrelation is the correlation between variable observations related to the location or an analytic distinction of location and attributes based on point distribution. Therefore, Moran's I methods were used to determine whether there is autocorrelation or spatial dependence between locations [11]–[13].

C. *Spatial Regressions*

The spatial regression method is used for spatial data types with a location effect (spatial effect). There are two types of spatial effects; spatial dependency and spatial heterogeneity. In spatial dependency, observations at one location depend on each other. The basis for the spatial regression method development is the classical linear regression method. The development is based on the influence of place or spatial on the analyzed data. Generally, the spatial regression model can be written as follows [14],[15]:

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

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

where $\boldsymbol{y}$ is the vector of the dependent variable, ρ spatial lag coefficient of the dependent variable, $\boldsymbol{W}$ is the spatial weights matrix arranged based on contiguity, specifically queen and rook contiguity [16]. $\boldsymbol{X}$ is a matrix of independent variables, α is a constant coefficient, $\boldsymbol{\beta}$ is the vector of regression parameter, $\boldsymbol{\theta}$ is the spatial lag coefficient of independent variables, λ is the

spatial lag coefficient of error, $\boldsymbol{u}$ is the residual vector with a spatial effect, and $\boldsymbol{\varepsilon}$ is a vector of error model.

Generally, Spatial Cross Regressive (SCR) Model is formed when independent variables are affected. Spatial Autoregressive (SAR) occurs when only dependent variables are affected, while the Spatial Durbin Model (SDM) model is formed where both variables exhibit effects. The formula of these models are shown below:

- Spatial Cross Regressive (SCR or SLX)　　: $\boldsymbol{y} = \alpha + \boldsymbol{X\beta} + \boldsymbol{WX\theta} + \boldsymbol{\varepsilon}$
- Spatial Autoregressive (SAR)　　　　　　 : $\boldsymbol{y} = \rho\boldsymbol{Wy} + \alpha + \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$
- Spatial Durbin Model (SDM)　　　　　　　 : $\boldsymbol{y} = \rho\boldsymbol{Wy} + \alpha + \boldsymbol{X\beta} + \boldsymbol{WX\theta} + \boldsymbol{\varepsilon}$

Parameter estimation of these models is achieved by Maximum Likelihood Estimation (MLE) methods [15].

D. *Robust Spatial Regressions*

Robust regression reduces the impact of outliers on the parameter estimation in the analysis [17]. This approach is also applied in spatial regression models to analyze contaminated data and provide outliers-resistant results. In this study, the robust M-Estimator method is applied to the SCR, SAR, and SDM models to overcome outliers using Least Square Estimation methods[18], [19]. The Tukey Bisquare weighting function is used as shown below:

$$b(u_i) = \begin{cases} \left[1 - \left(\dfrac{u_i}{c}\right)^2\right]^2 & , if\ |u_i| \leq c \\ 0, & if\ |u_i| > c \end{cases}$$

where c is the tuning constant of the Tukey bisquare weighting (or *biweight*) estimator, c = 4.685 [19]. The algorithm details used to estimate the parameter of Robust SCR, Robust SAR, and Robust SDM are shown in Table 2-4.

Table 2. Parameter Estimation of Robust SCR using M-Estimator

---

Algorithm 1: Parameter Estimation of Robust SCR using biweight estimator

---

Input:

Observations Data: $\{y; X\}$

Spatial Weight Matrix (W)

Output:

1. Initialization of parameter SCR by Ordinary Least Square Estimator

$$y = \alpha + X\beta + WX\theta + \varepsilon$$

$$y = [\mathbf{1}_n \quad X \quad WX] \begin{bmatrix} \alpha \\ \beta \\ \theta \end{bmatrix} + \varepsilon$$

$$y = Z\delta + \varepsilon$$

$$\delta^{(0)} = (Z^T Z)^{-1} Z^T y$$

2. **Repeat**:

$$\delta^{(t+1)} = \left(Z^T B^{(t)} Z\right)^{-1} Z^T B^{(t)} y$$

Where: $\quad B^{(t)} = diag\left(b_1^{(t)}, b_2^{(t)}, \dots, b_n^{(t)}\right)$

$$b_i^{(t)} = \begin{cases} \left[1 - \left(\frac{u_i^{(t)}}{4.685}\right)^2\right]^2 & , if \left|u_i^{(t)}\right| \leq 4.685 \\ 0, \ if \left|u_i^{(t)}\right| > 4.685 \end{cases}$$

$$u^{(t)} = \frac{\varepsilon^{(t)}}{s}$$

$$s = \frac{median\left|\varepsilon^{(t)} - median\left(\varepsilon^{(t)}\right)\right|}{0.6745}$$

$$\varepsilon^{(t)} = y - Z\delta^{(t)}$$

3. **Until** δ Convergence

4. **Return** $\begin{bmatrix} \alpha \\ \beta \\ \theta \end{bmatrix}$

Table 3. Parameter Estimation of Robust SAR using M-Estimator

---

Algorithm 2: Parameter Estimation of Robust SAR using biweight estimator

---

Input:

Observations Data: $\{\boldsymbol{y}; \boldsymbol{X}\}$

Spatial Weight Matrix $(\boldsymbol{W})$

Spatial lag coefficient from the SAR model (ρ)

Output:

1. Initialization of parameter SAR by Ordinary Least Square Estimator

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \alpha + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{y} - \rho \boldsymbol{W}\boldsymbol{y} = [\mathbf{1}_n \quad \boldsymbol{X}] \begin{bmatrix} \alpha \\ \boldsymbol{\beta} \end{bmatrix} + \boldsymbol{\varepsilon}$$

$$(\boldsymbol{I} - \rho \boldsymbol{W})\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\delta}^{(0)} = (\boldsymbol{Z}^T \boldsymbol{Z})^{-1} \boldsymbol{Z}^T (\boldsymbol{I} - \rho \boldsymbol{W})\boldsymbol{y}$$

2. **Repeat**:

$$\boldsymbol{\delta}^{(t+1)} = \left(\boldsymbol{Z}^T \boldsymbol{B}^{(t)} \boldsymbol{Z}\right)^{-1} \boldsymbol{Z}^T \boldsymbol{B}^{(t)} (\boldsymbol{I} - \rho \boldsymbol{W})\boldsymbol{y}$$

Where: $\quad \boldsymbol{B}^{(t)} = diag\left(b_1^{(t)}, b_2^{(t)}, \dots, b_n^{(t)}\right)$

$$b_i^{(t)} = \begin{cases} \left[1 - \left(\frac{u_i^{(t)}}{4.685}\right)^2\right]^2 & , if \left|u_i^{(t)}\right| \leq 4.685 \\ 0, \; if \left|u_i^{(t)}\right| > 4.685 \end{cases}$$

$$\boldsymbol{u}^{(t)} = \frac{\boldsymbol{\varepsilon}^{(t)}}{s}$$

$$s = \frac{median\left|\boldsymbol{\varepsilon}^{(t)} - median(\boldsymbol{\varepsilon}^{(t)})\right|}{0.6745}$$

$$\boldsymbol{\varepsilon}^{(t)} = (\boldsymbol{I} - \rho \boldsymbol{W})\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{\delta}^{(t)}$$

3. **Until $\boldsymbol{\delta}$** Convergence

4. **Return** $\begin{bmatrix} \alpha \\ \boldsymbol{\beta} \end{bmatrix}$

Table 4. Parameter Estimation of Robust SDM using M-Estimator

---

Algorithm 3: Parameter Estimation of Robust SDM using biweight estimator

---

Input:

Observations Data: $\{y; X\}$

Spatial Weight Matrix (W)

Spatial lag coefficient from the SDM model (ρ)

Output:

1. Initialization of parameter SDM by Ordinary Least Square Estimator

    $$y = \rho Wy + \alpha + X\beta + WX\theta + \varepsilon$$

    $$y - \rho Wy = [\mathbf{1}_n \quad X \quad WX]\begin{bmatrix}\alpha\\\beta\\\theta\end{bmatrix} + \varepsilon$$

    $$(I - \rho W)y = Z\delta + \varepsilon$$
    $$\delta^{(0)} = (Z^T Z)^{-1} Z^T (I - \rho W)y$$

2. **Repeat**:

    $$\delta^{(t+1)} = \left(Z^T B^{(t)} Z\right)^{-1} Z^T B^{(t)} (I - \rho W)y$$

    Where: $B^{(t)} = diag\left(b_1^{(t)}, b_2^{(t)}, \dots, b_n^{(t)}\right)$

    $$b_i^{(t)} = \begin{cases} \left[1 - \left(\frac{u_i^{(t)}}{4.685}\right)^2\right]^2 & ,if \left|u_i^{(t)}\right| \leq 4.685 \\ 0, \ if \ \left|u_i^{(t)}\right| > 4.685 \end{cases}$$

    $$u^{(t)} = \frac{\varepsilon^{(t)}}{s}$$

    $$s = \frac{median\left|\varepsilon^{(t)} - median(\varepsilon^{(t)})\right|}{0.6745}$$

    $$\varepsilon^{(t)} = (I - \rho W)y - Z\delta^{(t)}$$

3. **Until** δ Convergence

4. **Return** $\begin{bmatrix}\alpha\\\beta\\\theta\end{bmatrix}$

The life expectancy model is developed using M-Estimator in robust spatial regression through the following steps:

1. Detection of spatial dependence with Moran's I test.
2. Parameter estimation of the spatial regression model (non-robust model)
3. Using Moran's Scatterplot to detect outliers [6].
4. Using M-estimator to estimate Robust Spatial Regression based on **Algorithm 1-3**, according to the model used.
5. Statistical inference of spatial regression model [20], [21].
6. Interpret model

3. MAIN RESULTS

A. Detection of Spatial Dependence with Moran's I Test

Moran's I test is used to detect the spatial dependence. The test was conducted with the *spdep* package in software R [22]. The *queen* method was used to form the spatial weights matrix, as shown in Table 5.

Table 5. Moran's I test for detection of spatial dependence

Variable	Moran's I	Z-value	p-value
LE	0.5752	-0.0294	0.0000*
ALS	0.4850	4.4498	0.0000*
PCHLB	0.3122	2.9544	0.0016*
IHP	-0.0954	-0.5709	0.7160
PP	-0.3767	3.5126	0.0002*
APCE	0.4665	4.2892	0.0000*

From Table 5, spatial dependence exists in dependent (ALS, PCHLB, PP, and APCE) and independent (LE) variables. Therefore, life expectancy model development in Central Java Province uses SCR, SAR, or SDM models.

H. YASIN, A.R. HAKIM, B. WARSITO
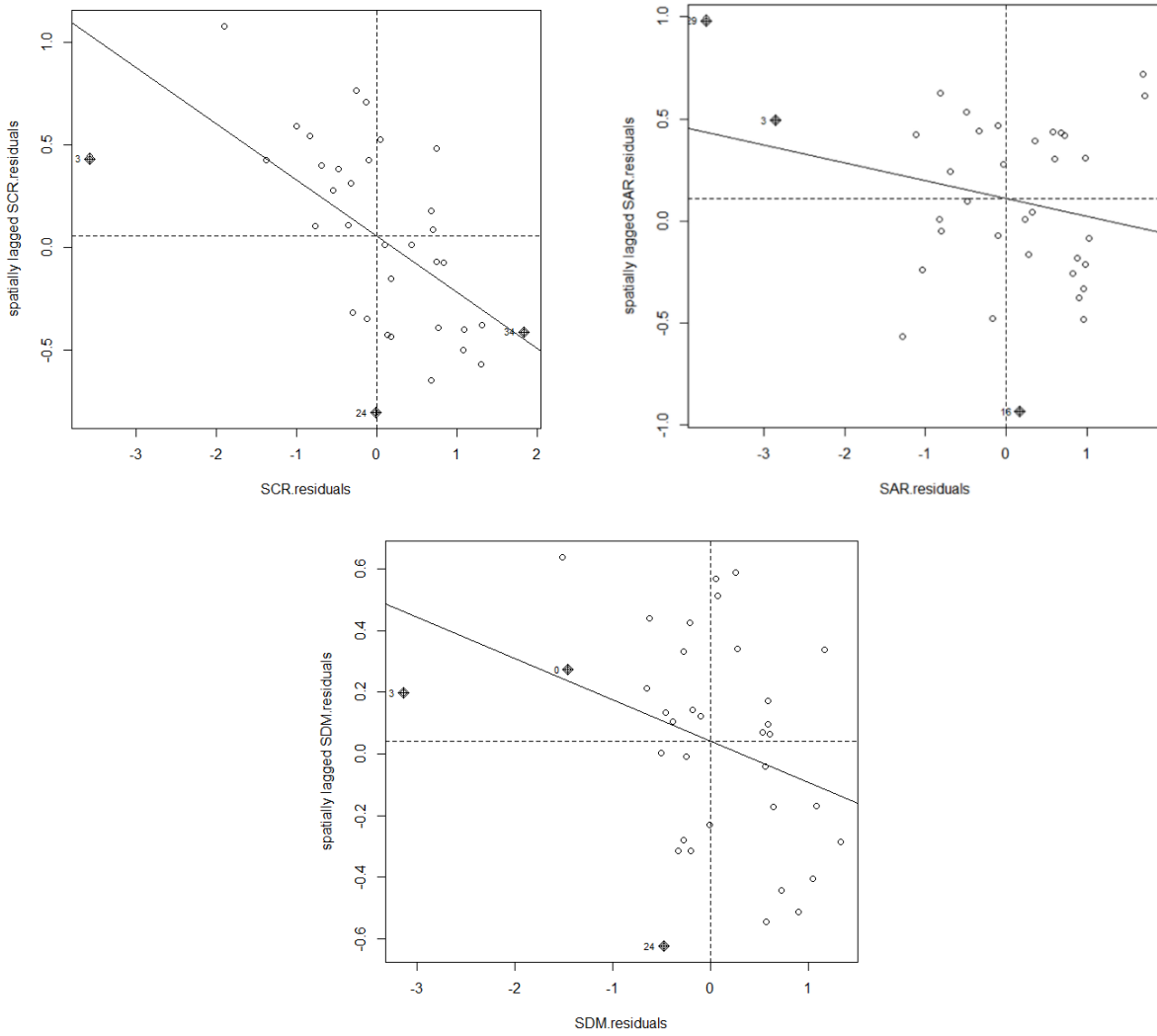
Figure 2. Moran Scatterplot of the residual model (SCR, SAR, SDM)

Table 6. Spatial Regression Model of Life Expectancy

Model	Variable	*Parameter*	*Coeff*	p-value	MSE	R^2
SCR	Intercept	α	69.9357	0.0000*	1.4615	72.90%
	ALS	β_1	1.1180	0.0113*		
	PCHLB	β_2	0.0461	0.1220		
	IHP	β_3	0.0006	0.2549		
	PP	β_4	-0.1708	0.0280*		
	APCE	β_5	0.0002	0.4987		
	W_ALS	θ_1	-0.2159	0.7926		
	W_PCHLB	θ_2	-0.0875	0.1532		
	W_IHP	θ_3	0.0027	0.0074*		
	W_PP	θ_4	-0.0242	0.8785		
	W_APCE	θ_5	-0.0001	0.8363		
SAR	W_LE	ρ	-0.5430	0.0105*	1.5198	65.95%
	Intercept	α	109.7200	0.0000*		
	ALS	β_1	0.6666	0.0369*		
	PCHLB	β_2	0.0353	0.1525		
	IHP	β_3	0.0004	0.3691		
	PP	β_4	-0.1941	0.0021*		
	APCE	β_5	0.0000	0.9644		
SDM*	W_LE	ρ	-0.7343	0.0084*	1.0656	80.24%
	Intercept	α	123.5800	0.0000*		
	ALS	β_1	0.8804	0.0026*		
	PCHLB	β_2	0.0354	0.0853		
	IHP	β_3	0.0006	0.0953		
	PP	β_4	-0.1458	0.0048*		
	APCE	β_5	0.0002	0.1354		
	W_ALS	θ_1	0.2686	0.6451		
	W_PCHLB	θ_2	-0.0631	0.1393		
	W_IHP	θ_3	0.0031	0.0000*		
	W_PP	θ_4	-0.1775	0.1368		
	W_APCE	θ_5	-0.0002	0.5448		

B. *Life Expectancy model using Spatial Regression Model*

In this research, three spatial regression models will be used; SCR, SAR, and SDM. Based on the *spatialreg* R Package [23], the results are as shown in Table 6. According to the parameter estimation (Table 6), "Average Length of School (ALS)" and "Percentage of Poor Population (PP)" variables affect the calculation of life expectancy using SCR, SAR, and SDM models. Moreover, the independent variable with spatial dependence, which significantly affects life expectancy

modeling through SCR and SDM models, is "lag of Number of Integrated Health Post (W_IHP)." Based on Table 6, using the smallest MSE and the largest R^2, SDM is the best spatial regression model to describe the life expectancy in Central Java Province.

Table 7. Robust Spatial Regression Model of Life Expectancy

Model	Variable	*Parameter*	*Coeff*	p-value	MSE	R^2
Robust	Intercept	A	77.0732	0.0000*	0.6739	85.41%
SCR	ALS	β_1	0.7340	0.0122*		
	PCHLB	β_2	0.0544	0.0068*		
	IHP	β_3	0.0008	0.0291*		
	PP	β_4	-0.2296	0.0000*		
	APCE	β_5	0.0002	0.2728		
	W_ALS	θ_1	-0.2949	0.6042		
	W_PCHLB	θ_2	-0.1071	0.0097*		
	W_IHP	θ_3	0.0031	0.0000*		
	W_PP	θ_4	-0.1940	0.0939		
	W_APCE	θ_5	-0.0002	0.6793		
Robust	W_LE	P	-0.5430	0.0026*	0.6449	75.57%
SAR	Intercept	A	110.7448	0.0000*		
	ALS	β_1	0.4289	0.0742		
	PCHLB	β_2	0.0310	0.0913		
	IHP	β_3	0.0004	0.2554		
	PP	β_4	-0.1748	0.0004*		
	APCE	β_5	0.0001	0.5286		
Robust	W_LE	P	-0.7343	0.0107*	0.4438	90.25%
SDM*	Intercept	A	130.2673	0.0000*		
	ALS	β_1	0.5260	0.0347*		
	PCHLB	β_2	0.0423	0.0134*		
	IHP	β_3	0.0007	0.0121*		
	PP	β_4	-0.1881	0.0000*		
	APCE	β_5	0.0003	0.0451*		
	W_ALS	θ_1	0.1777	0.7212		
	W_PCHLB	θ_2	-0.0800	0.0236*		
	W_IHP	θ_3	0.0034	0.0000*		
	W_PP	θ_4	-0.3451	0.0015*		
	W_APCE	θ_5	-0.0003	0.3635		

C. Detection of outliers

The method used to detect outliers in spatial data is looking at the residual models' Moran scatter plot [6]. The Moran Scatterplot of residual SCR, SAR and SDM models are shown in Figure 2. Several observations indicate outliers' existence (influence measures), including in SCR, SAR, and SDM models. Therefore, a robust spatial regression model is needed to reduce outliers influence and increase accuracy.

D. Robust Spatial Regression to model life expectancy

Robust spatial regression model parameters are estimated based on Algorithms 1, 2, and 3, as detailed in Table 2-4. The results are shown in Table 7. From Table 7, the robust method increases the significance level of the model parameters, which in turn raises the significant variables. This is compared to Table 6, where the significant variables in SCR and SDM models escalate from 3 to 6 and 8, respectively. Moreover, the method results increase model accuracy and decrease MSE. Therefore, this method increases R^2 in the SCR model from 72.90% to 85.41%, in the SAR model from 65.95% to 75.57%, and in the SDM model from 80.24 to 90.25%. From Tables 6 and 7, using the smallest MSE and the largest R^2, Robust SDM is the best spatial regression model to describe the life expectancy in Central Java Province.

E. Model Interpretation

From the tests carried out, the best spatial regression method to develop 2017 life expectancy in Central Java Province is the Robust Spatial Durbin Model (Robust SDM). The model formed is as follows:

$$\hat{y} = -0.7343 Wy + 130.2673 + [ALS \quad PCHLB \quad IHP \quad PP \quad APCE] \begin{bmatrix} 0.5260 \\ 0.0423 \\ 0.0007 \\ -0.1881 \\ 0.0003 \end{bmatrix}$$

$$+ [W\_ALS \quad W\_PCHLB \quad W\_IHP \quad W\_PP \quad W\_APCE] \begin{bmatrix} 0.1777 \\ -0.0800 \\ 0.0034 \\ -0.3451 \\ -0.0003 \end{bmatrix}$$

Based on this model, it can be explained that:

- An increase in the Average Length of School (ALS) by1 year increases the Life Expectancy by 0.5260 years, assuming other variables are constant.
- An increase in Households' Percentage with Clean and Healthy Living Behavior (PCHLB) by 1% raises the Life Expectancy by 0.0423 years, assuming other variables are constant.

- An increase in the Number of Integrated Health Post (IHP) by 1 unit increases the Life Expectancy by 0.0007 years, assuming other variables are constant.

- An increase in the Percentage of Poor Population (PP) by 1% reduces Life Expectancy by 0.1881 years, assuming other variables are constant.

- An increase in Adjusted Per Capita Expenditure (APCE) by 1,000 IDR increases the Life Expectancy by 0.0003 years, assuming other variables are constant.

- The spatial lag coefficient of the variable Y (ρ) is (-0.7343), meaning the Life Expectancy of each district/city has an influence of (-0.7343) times the average Life Expectancy of each neighboring district/city.

- The spatial lag coefficient of the variable W_ALS (θ_1) is 0.1777 which means mean Life Expectancy increases by 0.1777 times the average ALS of each neighboring district/city.

- The spatial lag coefficient of the variable W_PCHLB (θ_2) is (-0.0800), meaning Life Expectancy decreases by 0.0800 times the average PCHLB of each neighboring district/city.

- The spatial lag coefficient of the variable W_IHP (θ_3) is 0.0034, meaning Life Expectancy increases by 0.0034 times the average IHP of each neighboring district/city.

- The spatial lag coefficient of the variable W_PP (θ_4) is (-0.3451), which implies that Life Expectancy decreases by 0.3451 times the average PP of each neighboring district/city.

- The spatial lag coefficient of the variable W_APCE (θ_1) is (-0.0003), which decreases Life Expectancy by 0.0003 times the average APCE of each neighboring district/city.

4. CONCLUSION

The use of robust spatial regression methods in modeling life expectancy increases the model accuracy by about 10%. Robust SD is the model for developing life expectancy in Central Java. Therefore, the life expectancy rate in a district/city is also determined by the surrounding regions' spatial effect. It is affected by Average Length of School (ALS); Percentage of Households with Clean and Healthy Living Behavior (PCHLB), Number of Integrated Health Post (IHP), Percentage of Poor Population (PP), and Adjusted Per Capita Expenditure (APCE), whether spatially or not. The Percentage of Poor Population (PP) and the spatial lag of the Number of Integrated Health Post (IHP) are the most significant factors determining life expectancy in Central Java Province.

ACKNOWLEDGMENT

AUTHORS' CONTRIBUTIONS

Hasbi Yasin leads this research and does the conceptualization, methodology, software, investigation, data curation, formal analysis, writing-original draft preparation, writing-review, and editing. Arief Rachman Hakim, project administration, and funding acquisition. Budi Warsito, writing-review and editing, resources, and supervision. All authors read and approved the final manuscript.

CONFLICT OF INTERESTS

The authors declared no conflict of interest.

REFERENCES

[1] Badan Pusat Statistika Provinsi Jawa Tengah, Provinsi Jawa Tengah Dalam Angka (Jawa Tengah Province in Figures) 2018. Semarang: Badan Pusat Statistik Provinsi Jawa Tengah, 2018.

[2] D.C. Montgomery, G.C. Runger, Applied Statistics and Probability for Engineers, vol. 30, no. 1. Wiley, New York, 1998.

[3] S.F. Higazi, D.H. Abdel-Hady, S.A. Al-Oulfi, Application of spatial regression models to income poverty ratios in middle delta contiguous counties in egypt, Pak. J. Stat. Oper. Res. 9 (1) (2013), 93–110.

[4] A. Karim, A. Faturohman, S. Suhartono, D.D. Prastyo, B. Manfaat, Regression Models for Spatial Data: An Example from Gross Domestic Regional Bruto in Province Central Java, J. Ekon. Pembang. Kaji. Masal. Ekon. Pembang. 18 (2) (2017), 213-224.

[5] L. Anselin, Spatial data analysis with gis : an introduction to application in the social sciences, Technical Report 92-10, 1992.

[6] S. Shekhar, C.T. Lu, P. Zhang, A Unified Approach to Detecting Spatial Outliers, Geoinformatica, 7 (2) (2003), 139–166.

[7] K. Khotimah, K. Sadik, A. Rizki, Kajian regresi kekar menggunakan metode penduga-mm dan kuadrat median terkecil. Indones. J. Stat. Appl. 4 (1) (2020), 97-115.

[8] C. Chen, Robust Regression and outlier detection with the ROBUSTREG procedure. In: Proceedings of the

Twenty-Seventh Annual SAS Users Group International Conference. SAS Institute Inc., Cary (2002).

[9] Q. Cai, H. He, H. Man, Spatial outlier detection based on iterative self-organizing learning model, Neurocomputing, 117 (2013), 161–172.

[10] Dinas Kesehatan Provinsi Jawa Tengah, Profil Kesehatan Provinsi Jawa Tengah 2017, vol. 3511351, no. 24. Dinas Kesehatan Provinsi Jawa Tengah, 2017.

[11] M.F. Goodchild, Spatial Autocorrelation. Norwich: Geo Books, Norwich, 1989.

[12] T. Wuryandari, A. Hoyyi, D. S. Kusumawardani, D. Rahmawati, Identifikasi Autokorelasi Spasial Pada Jumlahpengangguran Di Jawa Tengah Menggunakan Indeks Moran," Media Stat. 7 (1) (2014), 1–10.

[13] H. Yasin, R. Saputra, Pemetaan Penyakit Demam Berdarah Dengue Dengan Analisis Pola Spasial Di Kabupaten Pekalongan, Media Stat. 6 (1) (2013), 27–36.

[14] L. Anselin, Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht, 1988.

[15] J.P. LeSage, R.K. Pace, Introduction to spatial econometrics. Taylor & Francis, New York, 2009.

[16] J.P. LeSage, The Theory and Practice of Spatial Econometrics. University of Toledo, New York. 1999.

[17] N.R. Draper, H. Smith, Applied Regression Analysis. John Wiley and Sons, New York. 1998.

[18] J. Nahar, S. Purwani, Application of Robust M-Estimator Regression in Handling Data Outliers, in 4th ICRIEMS Proceedings, 2017, pp. 53–60.

[19] J. Fox, S. Weisberg, An R Companion to Applied Regression, 3rd ed. SAGE Publications, California. 2018.

[20] P.M. Robinson, S. Thawornkaiwong, Statistical inference on regression with spatial dependence, J. Econometrics. 167 (2012), 521–542.

[21] R. Bivand, G. Piras, Comparing Implementations of Estimation Methods for Spatial Econometrics, J. Stat. Softw. 63 (18) (2015), 1–36.

[22] R. Bivand, spdep: Spatial Dependence: Weighting Schemes, Statistics, R package version 1.1-5. 2020.

[23] R. Bivand, spatialreg: Spatial Regression Analysis, R package version 1.1-5. 2019.